

ANOVA-Analysis: The Effect of Vocabulary Mastery and Reading Comprehension in Paraphrasing Skills

Ardianto Ardianto¹, Hadirman Hadirman², Agustini Buchari³, Zainuddin Soga⁴, Khairul Huda⁵ and Irmawati Irmawati⁶

Abstract

The research aims to examine ANOVA statistic about the effect of three main aspects: First, the influence of Vocabulary Mastery on students' ability in paraphrasing the paragraph. Second, the influence of Reading Comprehension Ability on students' ability in paraphrasing the paragraph. The last is the correlation between Mastery Vocabulary and Reading Ability Comprehension of student's ability in paraphrasing the paragraph. This research subjects involved 250 students at State Madrasah Aliyah Model Plus Skills of Manado. The results of the research indicates that there is a significant difference in how the level of vocabulary and reading mastery affects students' ability to paraphrase, and there is a significant difference in how the level of vocabulary mastery affects students' ability to read comprehension.

Keywords: ANOVA, Statistic, Vocabulary

INTRODUCTION

Understanding and restating other people's thoughts according to the sender's previous meaning is a complex matter. Misunderstandings often occur because the reader or interlocutor interprets some words or terms differently. In the context of written language, this shows that readers are less able to understand the content of the reading and lack understanding of vocabulary and understanding of sentence connotations. Therefore, the vocabulary mastery factor has a significant influence; it even becomes a prerequisite for understanding messages conveyed, both orally and in writing and allows someone to receive and convey broader and more complex information. Vocabulary learning is the heart of second and/or foreign language acquisition as it enables learner to achieve all form oral and written communication (Tuêce K°se & Mede, 2016). Therefore, to be able to communicate using language, both spoken and written, sufficient vocabulary is required.

Vocabulary mastery significantly impacts all aspects of language skills, such as listening, speaking, reading, and writing (Susilawati and Suhardi 2016). When we discuss the ability to paraphrase, this is closely related to the four language skills because the ability to rephrase a text into a paraphrase depends on how wealthy a person's vocabulary is. The larger the vocabulary a person has, the more fluently can receive and convey information and re-express it in a new written form (a paraphrase). Students' ability to paraphrase, namely the ability to restate the message in their own words, can be used as a primary indicator of their understanding of the reading content. A basic understanding of reading texts is very dependent on mastering adequate vocabulary. Therefore, both vocabulary mastery and the ability to read with comprehension significantly affect students' ability to paraphrase.

¹ Corresponding Author, Associate Professor in Indonesian Language Education at Faculty of Education and Teacher Training, State Islamic Institute of Manado, Manado, Indonesia, E-mail: ardianto@iain-manado.ac.id.

² Lecturer at Faculty of Ushuluddin, Adab and Dakwah, State Islamic Institute of Manado, Manado, Indonesia, E-mail: hadirman@iain-manado.ac.id

³ Teacher of Japanese Language Education, State Madrasah Aliyah Model Plus Skills of Manado, Manado, Indonesia, E-mail: agustinybuchari@gmail.com

⁴ Lecturer in Arabic Language Education, Faculty of Education and Teacher Training, State Islamic Institute of Manado, Manado, Indonesia, E-mail: z.soga@iain-manado.ac.id

⁵ Master of Arts in Interdisciplinary Islamic Studies, State Islamic University of Sunan Kalijaga, Yogyakarta, Indonesia, E-mail: huda.uinsk@gmail.com

⁶ Lecturer in English Language Studies, Hasanuddin University, Makassar, Indonesia, Email: irmawatiimarah@gmail.com

The focus of this research is to analyze three main aspects, namely; how much influence the level of vocabulary mastery has on students' ability to paraphrase, how much influence does the level of reading ability and comprehension have on students' ability to paraphrase and, how the interaction between vocabulary mastery and reading ability with comprehension influences students' paraphrasing ability.

The analysis results of these three research focuses can significantly contribute to improving the quality of Indonesian language learning in schools or madrasas. Efforts to improve Indonesian language teaching must begin with understanding who the students are learning and identifying obstacles related to their learning process.

Vocabulary Mastery

Language learning, including Indonesian, focuses on achieving four main language skills aspects: listening, reading, speaking, and writing. As stated by Garrison (1977, 558), these aspects of language skills are closely related. He also stated that if this is the case, increasing ability in one or more of these aspects will positively impact other language skills. On this basis, to improve writing skills, students must develop speaking, listening, and reading skills. However, it is not enough to master these aspects; Vocabulary mastery also has a crucial role in appropriately expressing feelings and thoughts.

The importance of vocabulary mastery by each individual reflects their intellectual growth and is a manifestation of real efforts to improve their understanding of words (Dakhi and Fitria 2019; McCrimmon 1968). Improving vocabulary mastery cannot be separated from a person's intellectual development because the two are closely related (Hariati 2020). One of the best methods to expand vocabulary is to understand words context. A person can recognize new words in certain situations. Other times, they may try to formulate the meaning of a word by using synonyms or providing concrete examples to explain the meaning more clearly. A good understanding of vocabulary is an important foundation for effective communication and continued intellectual development.

A vocabulary teaching approach that emphasizes context is also a concern for Gorman. Gorman (as quoted in Ramli, 2002, p. 220) stated that understanding the meaning of words through context, providing word definitions, explanations, illustrations, and the affixation process (adding prefixes or suffixes) is more effective than just referring to a dictionary. Based on this view, Ramli (Ramli 2002) concluded that the best approach to help students master vocabulary is to use these words in context. By practicing this habit, students can master new vocabulary more quickly and efficiently. To increase students' understanding of figurative meanings, Heilman (1972) proposed several strategies, including:

Matching Denotative Meaning with Connotative Meaning: One practical approach is to link a word's denotative meaning (concrete meaning) with the related connotative meaning (figurative meaning). In this way, students can more easily understand how figurative meanings relate to the basic meanings they know.

Create a Definition in Your Own Words: Students are asked to define the given figurative meaning using their own words. It helps students to internalize figurative meanings in a more personal and in-depth way.

Identifying Words with Figurative Meanings in Texts: Students are invited to identify words used in figurative meanings in the texts they read. It helps them recognize and understand figurative meanings in authentic contexts.

Reading Comprehension

Reading is one of three significant skills for students in the school environment. These three skills include reading, mathematics and communication abilities (Cronbach 1970, 510). The inability to achieve competency in these aspects can be a serious obstacle in completing various tasks within and outside formal education. Therefore, every teacher must monitor their students' progress in this aspect closely. Frandsen (1957, 289) also emphasized that reading is considered a process that involves understanding, rational thinking, and even the ability to think creatively. Hence, the ability to understand or "comprehension" is very crucial in this

process. Comprehension is absorbing words and the ability to analyze, interpret, and relate information obtained from text. Therefore, comprehension is one of the most critical elements in reading skills.

Comprehension, sometimes called language decoding, is essential for reading skills. According to Luria (1982, 169), this process is the opposite of oral speech. Every word has homonyms because it can have different meanings. Therefore, a selection process is needed every time the word is used to determine the correct meaning and concretize the meaning of a word. This selection is a choice between various possible alternative meanings. The alternative is determined first by the context in which the word is used. Furthermore, understanding words involves not only selecting the relevant meaning from its possible uses, as usually found in a dictionary. A word can have a special meaning only in a particular context. More than understanding a word in a text is required to know its meaning and reference.

Paraphrasing

Paraphrasing is a re-expression of a concept using the same language, without changing the meaning, but with a slightly different delivery (Kridalaksana 1993, 154). It also includes providing literal meaning with different words without changing the original idea in the existing text (Winterowd 1975, 197). In other words, paraphrasing involves restating the author's ideas or reconveying specific ideas about what has been read or heard using different words and sentences (D'Angelo 1980, 492). Paraphrasing is the same as writing, where someone tries to re-express other people's ideas in their own words and sentences. It is important to note that a paraphrase is not a direct copy of another person's sentences. To paraphrase well, one must truly understand the text one has read (Heilman 1972).

In this context, Singer and Donlan (1980, p. 136) explain that writing is a creative act, while copying is not. Writing requires writers to translate their thoughts into a distinctive and unique language, whereas copying does not require such creativity. Thus, writing at the peak of one's abilities can be an exhilarating experience, while copying at the peak of one's abilities is simply humbling.

FRAMEWORK OF THINKING

The framework of thinking developed in this research is described below.

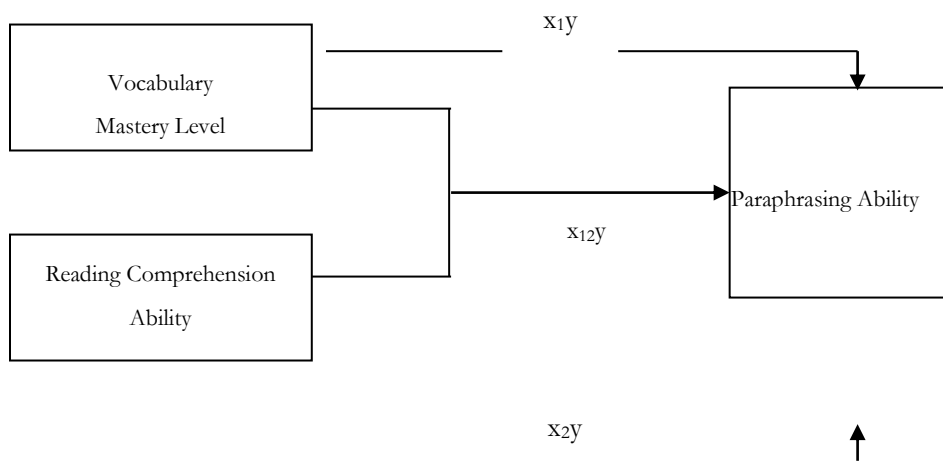


Figure 1. Research Model

METHODOLOGY

This research involved 250 students of Madrasah Aliyah Negeri Model 1 Plus Skills Manado who were at grade XI level during the 2022-2023 academic year. For statistical analysis, the sample of 250 students was then grouped based on two factors, namely the level of vocabulary mastery (high and low) and reading

$$6) RJK_A = \frac{JK_A}{K_{A-1}}$$

$$7) RJK_B = \frac{JK_B}{K_{B-1}}$$

$$8) RJK_{A \times B} = \frac{JK_{A \times B}}{(KA - 1)(K_{B-1})}$$

$$9) DK_{DS} = \sum (n_{i-1})$$

$$10) \text{Rasio F} = \frac{RJK_a}{RJK_d}$$

RJK_a is the average sum of squares between cells

RJK_d is the average of the sum of squares in the cell

Hypothesis testing is carried out by comparing the F_{count} value with the F_{table} value at a significance level 0.05. The criterion for this test is to reject H₀ (null hypothesis) if the calculated F value is greater than the F value in Ftable with certain degrees of freedom v1 and v2, according to the predetermined significance level, namely 0.05. In other words, if F_{count} > F_{Fv1.v2table}, then H₀ will be rejected.

Test Validity and Reliability

The test instruments used in this research were a vocabulary mastery test (X1) and a reading comprehension test (X2). Both of these tests have been tested for validity and reliability. The data for variable Y (paraphrasing task) is only instructional, where research subjects are only given instructions to paraphrase the text provided.

The validity of the vocabulary mastery and reading comprehension test instruments was tested using the Pearson correlation technique using the Statistical Package for the Social Sciences (SPSS) version 25 application. To determine the critical value in the r-value correlation table, the linear interpolation formula was used, as explained by Chapra & Canale (1997, p. 232). In this case, the critical value of the correlation table for N = 248 is 0.126. A test item is considered valid if the calculated correlation value (r-count) is greater than or equal to the predetermined critical value of the correlation table (r-table), namely 0.126. Based on these criteria, an analysis was conducted to determine which items were valid and which were invalid for vocabulary mastery and reading comprehension skills, as presented in Table 2.

Table 2. Test Validity Test Results

Vocabulary Mastery Test Validity Coefficient		Validity Coefficient of Reading Comprehension Ability Test	
Test	r-value	Test	r-value
1	0,449	1*	-0,016
2	0,868	2	0,326
3	0,406	3	0,976
4	0,629	4	0,976
5	0,891	5	0,985
6	0,479	6	0,489
7	0,712	7	0,985
8	0,854	8*	-0,071
9	0,232	9*	0,005
10	0,613	10	0,968
11	0,610	11	0,981

Vocabulary Mastery Test Validity Coefficient		Validity Coefficient of Reading Comprehension Ability Test	
Test	r- value	Test	r-value
12	0,164	12	0,959
13	0,140	13*	0,067
14	0,950	14*	0,015
15	0,380	15*	0,006
16	0,609	16	0,961
17	0,146	17	0,954
18	0,935	18	0,931
19	0,838	19	0,855
20	0,871	20	0,912
21	0,958	21	0,530
22	0,420	22	0,960
23	0,886	23	0,922
24	0,909	24*	-0,091
25*	-0,107	25*	-0,119
26	0,946	26*	-0,059
27	0,879	27	0,681
28	0,175	28*	0,109
29	0,971	29	0,761
30	0,720	30	0,980

*) Test items are invalid so they are discarded

The test items used to measure vocabulary mastery and reading comprehension ability have undergone a reliability test using the Kuder-Richardson 20 (KR-20) formula, as explained by Nurgiyantoro (2001, p. 122). The results of reliability calculations using the KR-20 formula show that $r_{KR-20} = 0.959$ for vocabulary mastery and $r_{KR-20} = 0.982$ for reading comprehension ability. These correlation values, namely 0.959 and 0.982, are considered strong or substantial correlations, according to the criteria stated by Best (1978, p. 260). Thus, in terms of reliability, these two tests can be considered reliable and can be used effectively in this research.

Level of Difficulty and Differentiating Power of Test Items

A good test is a test that has the right level of difficulty, neither too difficult nor too easy. Therefore, when using tests as research instruments, it is necessary to check the test's difficulty level. Apart from that, each item in the test must also have good discriminating power to ensure that the test can differentiate between individuals with different abilities well.

Using the Chung-The Fan Table (Fan, 1952) to analyze the difficulty level of test items is essential in this research. This table not only provides information about the extent to which test items are considered difficult or easy but also provides insight into the discriminating power of test items and the proportion of subjects who answered correctly. In this context, only the percentage of items categorized as complex, medium and easy will be presented, making it possible to understand the difficulty level in more detail.

Table 3. Degree of Test Item Difficulty

Difficulty Level	Vocabulary Mastery Test		Reading Comprehension Ability Test	
	Frequency (f)	Percentage (%)	Frequency (f)	Percentage (%)
Hard (Sk)	3	10,34	3	15.00
Medium (Sd)	16	55,17	16	80.00
Easy (Md)	10	34,49	1	5.00
Total	26	100		

The data contained in Table 3 illustrates that, in general, around 55.17% of the test items used in measuring vocabulary mastery in the subjects of this research were classified in the 'medium' category. Only 3 test items, or around 10.34%, fall into the 'difficult' or 'difficult' category, while the remaining 34.49% fall into the 'easy' category.

Likewise, around 80.00% of the test items are included in the 'medium' category in measuring the variable reading comprehension ability. There are only 3 test items, or around 15.00%, which fall into the 'difficult' category, and only 1 test item, or around 5.00%, which falls into the 'easy' category. This data shows that the

tests used to measure students' vocabulary mastery and reading comprehension abilities tend to have difficulty in the 'medium' category.

Table 4. Differentiating Power of Test Items

Differentiating Power	Vocabulary Mastery Test		Reading Comprehension Ability Test	
	Frequency (f)	Percentage (%)	Frequency (f)	Percentage (%)
Very good (Bs)	23	79.31	19	95.00
Good (Bk)	2	6.90	1	5.00
Medium (Sd)	0	0.00	0	0.00
Poor (Br)	4	13.79	0	0.00
Total	29	100	20	100.00

The interpretation of the data in Table 3 is strengthened by the data in Table 4, which provides information regarding the differentiating power of the test items. Around 79.91% of the items used in measuring vocabulary mastery had a discriminating power that was categorized as 'very good.' Likewise, in measuring reading comprehension ability, all items, as much as 100%, show good discriminating power. The majority, namely 95.00%, falls into the 'very good' category, while the remaining 5% falls into the 'good' category. Thus, the questions used in measuring vocabulary mastery and the questions used in measuring reading comprehension ability in this study can be considered suitable for use, considering their good discriminating power.

FINDINGS AND ANALYSIS

Data Description

The score data used in this analysis were converted to T-scores. The decimal number scores have been converted to whole numbers to simplify analysis. This section presents some basic statistics, including the highest value, lowest value, mean, and standard deviation. In addition, we also provide information about the distribution of scores for each research variable. To organize the frequency distribution of the data, we used Sturges' formula to determine the class range, the number of class intervals, and the length of the corresponding class intervals (Sudjana 2005).

Distribution of Vocabulary Mastery Score Data

Based on calculating the transformation of vocabulary mastery score data into a T-score using the collected primary data, an average score of 48.86 was obtained, with a standard deviation of 10.171. For the T-score data on student vocabulary mastery measured from 250 students, the highest score was 59, while the lowest was 36.

Meanwhile, using the Sturges formula, the class range is 23, the number of interval classes is 8, and the length of the interval classes is 3. Thus, the frequency distribution table for vocabulary mastery scores can be described in detail in Table 5 below.

Table 5. Frequency Distribution of Vocabulary Mastery Scores

No.	Class Intervals	Absolute Frequency	Relative Frequency
1.	36 – 38	93	37,2 %
2.	39 – 41	15	6%
3.	42 – 44	0	0%
4.	45 – 47	0	0%
5.	48 – 50	9	3,6%
6.	51 – 53	2	0,8%
7.	54 – 56	3	1,2%
8.	57 – 59	128	51,2%
Total		250	100%

We can determine students' vocabulary mastery tendencies based on the frequency distribution of scores listed in Table 5 above. This tendency is divided into two categories, namely high vocabulary mastery and low

vocabulary mastery. The classification of vocabulary mastery into high and low is based on the following criteria: ≤ 47 is included in the low category, while ≥ 48 is included in the high category. By referring to these criteria, we can conclude that 142 respondents (56.8%) were in the high category, while 108 respondents (43.2%) were in the low category. The percentage of variable classification of students' vocabulary mastery in the high and low categories is shown in Table 6 below.

Table 6. Percentage of Vocabulary Mastery Categories

Category	Score	Total	Percentage (%)
High	48 – 59	142	56,8
Low	36 – 47	108	43,2
Total		250	100

Data Distribution of Reading Comprehension Ability Scores

Reading comprehension score data shows an average value of 42.22 with a standard deviation 10.03. The score of students' reading comprehension ability obtained from the measurement results shows that the highest score is 53, while the lowest score is 29. Using Sturges' rule, we can determine the class range of 24, the number of interval classes is 9, and the length of the interval classes is 3. Based on the results, the frequency distribution of students' reading comprehension ability score data can be presented in Table 7 below.

Table 7. Distribution of Reading Comprehension Ability Scores

No.	Class Intervals	Frequency Absolute	Frequency Relatively
1.	27 – 29	29	11,6%
2.	30 – 32	74	29,6%
3.	33 – 35	6	2,4%
4.	36 – 38	0	0%
5.	39 – 41	0	0%
6.	42 – 44	0	0%
7.	45 – 47	5	2%
8.	48 – 50	42	16,8%
9.	51 – 53	94	37,6%
Total		250	100%

Based on the data in Table 7 above, we can determine the trend in students' reading comprehension abilities. This trend includes the classification of reading comprehension ability into high and low. This classification is based on specific criteria, namely, ≤ 42 is included in the low category, and ≥ 43 is included in the high category. By referring to these criteria, it can be concluded that as many as 141 respondents (56.4%) had high reading comprehension abilities, while 109 respondents (43.6%) had low reading comprehension abilities. The percentage of categories of students' reading comprehension ability variables can be seen in Table 8 below.

Table 8. Percentage of Reading Comprehension Ability Categories

Category	Score	Total	Percentage (%)
High	42 – 53	141	56,4
Low	27 – 38	109	43,6
Total		250	100

The table above shows that the number of respondents for each category is almost equal, with a difference of only around 12.8%. Based on the distribution of score data on the independent variable X1 (vocabulary mastery) and the independent variable X2 (reading comprehension ability), the grouping of score data for the two independent variables can be seen in Table 9 below.

Table 9. Number of Subjects in High and Low Categories for Variables X1 and X2

Category	Variables	
	Vocabulary mastery (x ₁)	Reading comprehension (x ₂)
High	142	141
Low	108	109
Total	250	250

Based on the data in Table 9, there is a balance between the number of respondents in the high and low categories for these two variables. For variable X1, there were 142 respondents (56.8%) in the high category and 108 respondents (43.2%) in the low category out of 250 respondents. Meanwhile, for variable X2, there were 141 respondents (56.4%) in the high category and 109 respondents (43.6%) in the low category out of 250 respondents. This data shows that the percentage of respondents based on high and low categories is almost equal. This data will be used in Anova statistical calculations for testing research hypotheses.

Distribution of Paraphrasing Ability Score Data

Based on the T score in the ability to paraphrase, it can be observed that students obtained the highest score of 63, while the lowest score achieved by respondents was 36. An average score of around 50 was obtained by analyzing the collected data with a standard deviation of 10.02. Using Sturges' rule, it can be identified that the class range is 27, with a total class interval of 8 and a class interval length of around 4. Details of the frequency distribution of students' paraphrasing ability scores can be seen in Table 10 below.

Table 10. Distribution of Paraphrasing Ability Scores

No.	Class intervals	Frequency Absolut	Frequency Relative
1.	34 – 37	8	3,2%
2.	38 – 41	98	39,2%
3.	42 – 45	0	0%
4.	46 – 49	0	0%
5.	50 – 53	3	1,2%
6.	54 – 57	44	17,6%
7.	58 – 61	89	35,6%
8.	62 – 65	8	3,2%
Total		250	100%

Based on the data in Table 10 above, we can analyze trends in students' paraphrasing abilities which show two categories, namely high ability and low ability. This grouping refers to specific criteria: a score ≤ 49 which is in the low category, and a score ≥ 50 , which is in the high category. By referring to these criteria, we can conclude that out of 250 respondents, 144 respondents (57.6%) had a high ability in paraphrasing, while 106 respondents (42.4%) had a low ability in the same matter. The percentage of categories of students' ability to paraphrase variables can be seen in more detail in Table 11 below.

Table 11 Percentage of Paraphrasing Ability Categories

Category	Score	Total	Percentage (%)
High	50 – 65	144	57,6
Low	34 – 49	106	42,4
Total		250	100

Ability to Paraphrase Based on Level of Vocabulary Mastery

As stated previously, students' ability to paraphrase is closely related to their level of vocabulary mastery. Therefore, this section will explain trends in students' paraphrasing abilities based on their level of vocabulary mastery. Since each variable has two levels, the tendency in students' paraphrasing abilities related to the independent variables being measured will be reflected in a 2 X 2 cross-tabulation (contingency table).

The trend in students' ability to paraphrase, which is influenced by their vocabulary mastery, can be found in Table 12. In this table, students who have a high level of vocabulary mastery, reaching 100%, can paraphrase. On the other hand, only a tiny portion of students with low vocabulary mastery, as many as 9.4%, have high paraphrasing abilities. In contrast, the majority, namely 91.6%, have low paraphrasing abilities.

Table 12. Distribution of Paraphrasing Ability Based on Level of Vocabulary Mastery

Criterion Predictor	Vocabulary Mastery		Total	
	High	Low		
<i>Paraphrase</i>	High	133 (100%)	11 (9,4%)	144 (57,6%)
	Low	0 (0%)	106 (91,6%)	106 (42,4%)
Total		133 (53,2%)	117 (46,8%)	250 (100%)

The data in Table 12 above shows a strong relationship between students' ability to paraphrase and their vocabulary mastery. In other words, students with high vocabulary mastery tend to have high paraphrasing abilities. Furthermore, students with low vocabulary mastery tend to have a low ability to make paraphrases.

Ability to Paraphrase Based on Level of Vocabulary Mastery

The frequency distribution of students' ability to paraphrase based on the variable reading comprehension ability is also shown in Table 13, a 2 x 2 cross-tabulation table (contingency table). The data in this table indicates that students with a high level of reading comprehension ability, namely 100%, also have a high ability to paraphrase. Meanwhile, students with a low reading comprehension ability, only three or around 2.75%, have a high ability to paraphrase. However, this percentage differs from the number of students who fall into the low ability to paraphrase category, which is 106 or around 97.25%.

Table 13. Distribution of Paraphrasing Ability Based on Reading Comprehension Ability Level

Criterion Predictor	Reading Comprehension		Total
	High	Low	
<i>Paraphrase</i> High	141 (100%)	3 (2,75%)	144 (57,6%)
Low	0 (0%)	106 (97,25%)	106 (42,4%)
Total	141 (56,4%)	109 (43,6%)	259 (100%)

The data in Table 13 above strengthens the interpretation that students' reading comprehension ability is closely related to their paraphrasing ability. In this context, students with a high reading comprehension ability tend to have a high paraphrasing ability. On the other hand, students with low reading comprehension ability tend to have low paraphrasing abilities as well.

Testing Statistical Assumptions

A variance analysis test specifies conditions that must be met before it is carried out, namely that the population must be assumed to be normally distributed. After ensuring this assumption is met, the next stage is to conduct checks regarding the homogeneity of population variance. These two assumptions must be checked carefully, namely, whether the study population's characteristics meet the normality and homogeneity criteria. When one of these assumptions is not met, an alternative method or approach must be sought to test the data. Thus, each characteristic related to the population must meet these two assumptions so that the variance analysis test can be carried out correctly.

If one of the two assumptions, namely population normality and population variance homogeneity, is not met, then one method that can be used to fulfil these two assumptions is to transform the research data into a T-score. By carrying out this transformation, the initial data can be changed so that it is closer to a normal distribution and can also help achieve homogeneity of variance between different data groups. However, it is essential to remember that the transformation method should be based on a deep understanding of the data and unmet assumptions.

Apart from transforming data into T-scores, various other methods can be used to overcome incomplete assumptions, such as bootstrapping, nonparametric analysis, or more flexible statistical models. The choice of method will depend on the type of data, research objectives, and the context of the analysis being conducted.

Population Normality

The population characteristics that will be tested for normality include

- three variables,
- namely the independent variable X1 (level of vocabulary mastery),
- the independent variable X2 (reading comprehension ability), and
- the dependent variable Y (ability to paraphrase).

Therefore, the Lilliefors test technique was used to test the normality of each of these variables. The normality test results using the Lilliefors technique can be found in Table 14 for the vocabulary mastery level variable, Table 15 for the reading comprehension ability variable, and Table 16 for the paraphrasing ability variable. By considering the available score data, we can group the data for the vocabulary mastery variable (Variable X1) in the framework of the Lilliefors test, as shown in Table 14.

Table 14. Normality Test Results for Vocabulary Mastery Data

Xi	Fi	Zi	F(zi)	S(zi)	F(zi) - S(zi)	Information
8	39	-1.29	0.0985	0.1560	-0.0575	
9	2	-1.17	0.1210	0.1640	-0.0430	
10	52	-1.06	0.1446	0.3720	-0.2274	
12	15	-0.83	0.2033	0.4320	-0.2287	Maximum
19	9	-0.03	0.4880	0.4680	0.0200	
21	2	0.20	0.5793	0.4760	0.1033	
24	3	0.55	0.7088	0.4880	0.2208	
26	41	0.77	0.7794	0.6520	0.1274	
27	4	0.89	0.8133	0.6680	0.1453	
28	83	1.00	0.8413	1.0000	-0.1587	

The results of the Lilliefors test with a sample of 250 obtained a maximum L value of 0.2287, while the L0.95 (250) table was 0.0560 with a significance level of $\alpha = 0.05$. Therefore, it can be concluded that the maximum Lcount (0.2287) is greater than the L0.95 (250) table value (0.0560). Thus, at the significance level $\alpha = 0.05$ (95%), we have sufficient evidence to state that the population characteristics of the vocabulary mastery variable (X1) are generally not distributed.

Based on the available score data, we can group the data for the reading comprehension ability variable (Variable X2) in the framework of the Lilliefors test, as shown in Table 15. In this table, the maximum L-count was found to have a value of 0.3801. Meanwhile, the L0.95 (250) table value, with a significance level of $\alpha = 0.05$, is 0.0560. Thus, it can be concluded that the maximum L-count (0.3801) is greater than the L0.95 (250) table value (0.0560). At the significance level $\alpha = 0.05$ (95%), we have sufficient evidence to state that the population characteristics of the reading comprehension ability variable (X2) are generally not distributed.

Table 15. Normality Test for Reading Comprehension Ability Variables

Xi	Fi	zi	F(zi)	S(zi)	F(zi) - S(zi)	Information
0	29	-2.06	0.0197	0.1160	-0.0963	
1	39	-1.94	0.0262	0.2720	-0.2458	
2	35	-1.82	0.0344	0.4120	-0.3776	
3	1	-1.71	0.0436	0.4160	-0.3724	
4	5	-1.59	0.0559	0.4360	-0.3801	Maximum
14	1	-0.40	0.3446	0.4400	-0.0954	
15	4	-0.28	0.3897	0.4560	-0.0663	
16	6	-0.16	0.4364	0.4800	-0.0436	
17	36	-0.04	0.4840	0.6240	-0.1400	
18	50	0.08	0.5319	0.8240	-0.2921	
19	23	0.19	0.5754	0.9160	-0.3406	
20	21	0.31	0.6217	1.0000	-0.3783	

Based on the available score data, we can group the data for the paraphrasing ability variable (Variable Y) in the framework of the Lilliefors test, as shown in Table 16. In this table, the maximum L-count was found to have a value of 0.2452. Meanwhile, the L0.95 (250) table value, with a significance level of $\alpha = 0.05$, is 0.0560. Thus, it can be concluded that the maximum L-count (0.2452) is greater than the L0.95 (250) table value

(0.0560). At the significance level $\alpha = 0.05$ (95%), we have sufficient evidence to state that the population characteristics of the paraphrasing ability variable (Y) are not normally distributed.

Table 16. Normality Test for Paraphrasing Ability Variables

Xi	Fi	zi	F(zi)	S(zi)	F(zi) - S(zi)	Information
12	8	-1.36	0.0869	0.0320	0.0549	
13	44	-1.22	0.1112	0.2080	-0.0968	
14	44	-1.07	0.1423	0.3840	-0.2417	
15	10	-0.92	0.1788	0.4240	-0.2452	Maximum
23	3	0.25	0.5987	0.4360	0.1627	
24	10	0.39	0.6517	0.4760	0.1757	
25	13	0.54	0.7054	0.5280	0.1774	
26	21	0.68	0.7518	0.6120	0.1398	
27	34	0.83	0.7967	0.7480	0.0487	
28	36	0.98	0.8365	0.8920	-0.0555	
29	19	1.12	0.8686	0.9680	-0.0994	
30	8	1.27	0.8980	1.0000	-0.1020	

Based on the results of normality testing for the three population characteristics, namely X1 (vocabulary mastery level), X2 (reading comprehension ability), and Y (paraphrasing ability), it can be concluded that all three population characteristics are not generally distributed at a significance level of $\alpha = 0.05$. It indicates that the data for these three variables have a significantly non-symmetric distribution and do not follow a typical distribution pattern.

Homogeneity of Population Variance

Testing for homogeneity of population variance was carried out by applying the Barlett method, which was chosen because the population analyzed consisted of more than two groups. The results of this test significantly reveal that the three populations related to the research variables do not have homogeneity of variance within the significance level $\alpha = 0.05$, with degrees of freedom of 2 minus one, according to the number of subject groups involved in this research.

The results of testing population normality and homogeneity of population variance in this study show that the population does not follow a normal distribution and does not have a balanced homogeneity of population variance. Therefore, the data on each variable has transformed T-scores to meet the requirements of ANOVA statistical testing. This transformation automatically fulfils the assumptions of normality and homogeneity of population variance required in subsequent parametric statistical tests.

Hypotheses Tested

There are three hypotheses tested based on the research problem, namely:

The paraphrasing abilities of students who have a high level of vocabulary mastery and students who have a low level of vocabulary mastery are significantly different.

The statistical hypothesis is:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

The paraphrasing abilities of students who have high reading comprehension abilities and students who have low reading comprehension abilities are significantly different. The statistical hypothesis is:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

There is a significant interaction between the influence of the level of vocabulary mastery and reading comprehension ability on students' ability to paraphrase.

If the reading ability level is called A and the reading comprehension level is called B, then the statistical hypothesis is:

$$H_0: \text{interacts } A \times B = 0$$

$$H_1: \text{interacts } A \times B \neq 0$$

This research hypothesis has been substantiated using survey data to test its validity. In this research hypothesis, the main focus is to identify and analyze the impact of two main variables, namely the level of vocabulary mastery and reading comprehension ability, on students' ability to paraphrase. The significance of the influence of these two variables has been tested in detail using the F-test using two-way analysis of variance (with a 2 x 2 factorial design). The results of this test show that F table with degrees of freedom 1 and 249 has a value of 3.88. A complete summary of the results of the ANOVA analysis in the 2 x 2 factorial design can be found in the attached Table 17.

Table 17. Summary of ANOVA between groups of vocabulary mastery (A), reading comprehension ability (B), and their interactions

Sources of Variance	/K	DK	RJK	F _{hitung}	F _{table}
A	20436,51	1	20436,51	2974,75	
B	22849,71	1	22849,71	3326,01	3,88
A x B interaction	- 19978,83	1	- 19978,83	- 2908,13	
Mistakes (in cells)	1691,13	246	6,87		
Total		249			

In Table 17 above can be concluded that (1) there is an F-ratio obtained, which shows that the influence of the level of Vocabulary Mastery is significant at the significance level $\alpha = 0.05$ ($F_{\text{count}} = 2974.75$; $F_{\text{table}} = 3.88$), (2) the F-ratio results obtained also show that the effect of Reading Comprehension Ability is significant at the significance level $\alpha = 0.05$ ($F_{\text{count}} = 3326.01$; $F_{\text{table}} = 3.88$), and (3) however, the F-ratio obtained for The interaction effect between the level of Vocabulary Mastery and Reading Comprehension Ability together is not significant at the significance level $\alpha = 0.05$ ($F_{\text{count}} = - 2908.13$; $F_{\text{table}} = 3.88$).

The results of the ANOVA analysis indicated that the two independent variables, namely Vocabulary Mastery and Reading Comprehension Ability, significantly influenced students' ability to paraphrase as the dependent variable. However, it is important to note that both effects do not occur simultaneously; that is, Vocabulary Mastery has its own influence, as does Reading Comprehension Ability. The insignificant interaction between Vocabulary Mastery (column A) and Reading Comprehension Ability (row B) confirms that these effects do not occur together.

DISCUSSION

Analysis of variance for the first hypothesis test has produced a highly significant F-ratio ($=2974.75$) at the confidence level $\alpha = 0.05$. In other words, the hypothesis that the level of vocabulary mastery significantly impacts students' ability to make paraphrases is accepted, and the null hypothesis is rejected. These findings indicate a substantial difference in the effect between students with high vocabulary mastery and those with low vocabulary mastery on their ability to paraphrase. This conclusion is based on the significant F-ratio value ($= 2974.75$) resulting from data analysis. These results also reflect that differences in mean scores in vocabulary mastery testing did not occur by chance. Students who have good vocabulary mastery tend to be

better able to improve their ability to make paraphrases. On the other hand, students with poor vocabulary skills will need help in carrying out paraphrasing tasks.

Analysis of variance for the second hypothesis test has produced a highly significant F-ratio (=3326.01) at the confidence level $\alpha = 0.05$. It means that the hypothesis which states the ability to understand reading content significantly impacts students' ability to make paraphrases, is accepted, and the null hypothesis is rejected. These findings indicate that there is a significant difference in effect at $\alpha = 0.05$ between the group of students who have a high level of reading ability and the group of students who have a low level of reading ability on their ability to paraphrase, as shown by the significant F-ratio value (=3326.01) resulting from data analysis.

The results of the analysis of variance to test the third hypothesis revealed that the F-ratio value (= -2908.13) did not reach the significance level at $\alpha = 0.05$. Therefore, the hypothesis that there is an interaction between the influence of the level of vocabulary mastery and the influence of the level of ability to understand reading content on students' ability to make paraphrases is rejected, and the null hypothesis is accepted. This conclusion confirms that at the significance level $\alpha = 0.05$, there is not sufficient evidence to support the existence of an interaction between the two types of ability, namely vocabulary mastery (high and low) and reading comprehension ability (high and low), on students' ability to make paraphrase. This finding is strengthened by the F-ratio value (= -2908.13) resulting from data analysis. Therefore, there is no similarity or significant relationship between the effect of vocabulary mastery and the effect of reading comprehension ability on students' ability to make paraphrases.

Analysis of variance to test the third hypothesis has produced an F-ratio value (= -2908.13) which does not reach the significance level at $\alpha = 0.05$. Therefore, the hypothesis which states that there is an interaction between the influence of the level of vocabulary mastery and the influence of the level of ability to understand reading content on students' ability to make paraphrases, is rejected, and the null hypothesis is accepted. This conclusion states that at the significance level $\alpha = 0.05$, there is not sufficient evidence to support the existence of an interaction between the two types of ability, namely vocabulary mastery (high and low) and reading comprehension ability (high and low), on students' ability to make paraphrase. This finding is strengthened by the F-ratio value (= -2908.13) resulting from data analysis. Therefore, it can be concluded that there is no significant relationship or interaction between the effect of vocabulary mastery and the effect of reading comprehension ability on students' ability to paraphrase.

The absence of interaction also indicates that the difference in the ability to paraphrase between two groups of students with different levels of reading comprehension ability (B1 and B2) in the group of students with high vocabulary mastery (A1) will not differ significantly from the difference in the ability to paraphrase between the same two groups of students (B1 and B2) in the group of students with low vocabulary mastery (A2).

Another interpretation of the data result analysis that was presented previously is that in making conclusions about this data analysis, there is a level of confidence of around 95% that the conclusions drawn are correct. In other words, if a similar study were conducted in a place with similar conditions to this study population and used the same methods, then it is likely that the study's results would be similar to the results of this study, with an error rate of around 5%.

CONCLUSIONS

Based on the results of hypothesis testing carried out, this research reveals several things that can be concluded: (a) Students with a high level of vocabulary mastery also have a high ability to make paraphrases. On the other hand, students with a low level of vocabulary mastery also have a low ability to paraphrase. In other words, the higher a student's vocabulary mastery, the better the paraphrases they make; (b) Students with a high level of reading comprehension ability also have a high ability to make paraphrases. On the other hand, students with a low reading comprehension ability also have a low ability to paraphrase. In other words,

to make a good paraphrase, someone must understand the content of the reading better. The better a person's reading comprehension, the better the paraphrases he makes; (c) There is no interaction between the influence of vocabulary mastery and reading comprehension on students' ability to make paraphrases. It shows that one of the independent variables, namely vocabulary mastery alone and reading comprehension ability alone, is in the high category, which will trigger students' high ability to paraphrase.

REFERENCES

- Best, John W. 1978. *Research in Education*. New Delhi: Prentice-Hall of India Private Ltd.
- Cronbach, Lee J. 1970. *Essentials of Psychology Testing*. New York: Harper & Row Publisher.
- D'Angelo, Frank J. 1980. *Process and Thought in Composition*. Cambridge: Winthrop Publisher, Inc.
- Dakhi, Saniago, and Tira Nur Fitria. 2019. "The Principles and the Teaching of English Vocabulary: A Review." *JET (Journal of English Teaching)* 5(1): 15.
- Faraj, Avan Kamal Aziz. 2015. "Effective Strategies for Turning Receptive Vocabulary into Productive Vocabulary in EFL Context." *Journal of Education and Practice* 6(27).
- Fillmore, Charles J. 1968. "The Case for Case." In *Universals in Linguistics Theory*, eds. E. Back and R.G. Harms. New York: Holt, Rinehart & Winston.
- Frandsen, Arden M. 1957. *How Children Learn*. New York: McGraw-Hill Book Company, Inc.
- Garrison, Karl C. 1977. "Learning and Basic School Subject." In *Educational Psychology*, ed. Charles E. Skinner. New Delhi: Prentice-Hall of India Private Limited.
- Gersmehl, Phil. 2020. "Symbolics, Syntactics, and Semantics: Teaching a Language of Maps." In *Handbook of the Changing World Language Map*, Cham: Springer International Publishing, 249–74.
- Hadi, Alnoor Sabit Abdelkarim. 2017. "Significance of Vocabulary in Achieving Efficient Learning." *American Scientific Research Journal for Engineering* 29(1): 271–85.
- Halliday, M.A.K. 1970. "Language Structure and Language Function." In *New Horizons in Linguistics*, ed. J. Lyons. Harmondsworth: Penguin, 140–65.
- Hariati, Puji. 2020. "Improving Students' Vocabulary Mastery through Teaching Real Objects." *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal* 3(2): 740–48.
- Heilman, Arthur W. 1972. *Principles and Practices of Teaching Learning*. Columbus: Charles E. Merrill Publishing Company.
- Kieffer, Michael J., and Nonie K. Lesaux. 2012. "Knowledge of Words, Knowledge about Words: Dimensions of Vocabulary in First and Second Language Learners in Sixth Grade." *Reading and Writing* 25(2): 347–73.
- Lakoff, G. 1973. "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts." *Journal of Philosophical Logic* 2: 458–508.
- Luria, Alexander R. 1982. *Language and Cognition*. New York: John Wiley & Sons.
- McCawley, James D. 1973. "Prelexical Syntax." In *Semantic Syntax*, Oxford: Clarendon, 29–42.
- McCrimmon, James M. 1968. *Writing a Purpose*. Boston: Houghton Mifflin Company.
- Rahman, Aam Ali, and Anggi Angraeni. 2020. "Empowering Learners with Role-Playing Game for Vocabulary Mastery." *International Journal of Learning, Teaching and Educational Research* 19(1): 60–73.
- Saud, Mohan Singh. 2023. "Measuring the Academic Vocabulary Size of University Students in Nepal." *Mextesol Journal* 47(1).
- Singer, Harry, and Dan Donlan. 1980. *Reading and Learning from Text*. Boston: Little, Brown and Company.
- Tuće Kose, E. "Influence of Vocabulary Mastery on Reading." *Journal of English Education* Vol. 4 No. 2 December 2018 | 107 P-ISSN: 2459-9719, E-ISSN 2597-7091 E. (2016). Perception of EFL Learner about Using an Online Tool for Vocabulary Learning in EFL Classroom: A Pilot Project in Turkey. *Procedia - Social and Behavioral Sciences*, 232, 362±372. <https://doi.org/10.1016/j.sbspro.2016.10.051>
- Winterowd, W. Ross. 1975. *The Contemporary Writer*. New York: Harcourt Brace-Jovanovisch, Inc.