

---

## A REGRESSION ANALYSIS OF THE EFFECT OF SYNTACTIC COMPETENCE ON ARABIC TEXT READING SPEED

Putri Syarifatul Mahfudloh<sup>1</sup>, Abdul Basith<sup>2</sup>

<sup>1,2</sup>Universitas KH Mukhtar Syafa'aat Banyuwangi; Indonesia

Correspondence E-mail; rifabq706@gmail.com

Submitted: 11/02/2026

Revised: 23/02/2025

Accepted: 24/03/2026

Published: 10/04/2026

---

### Abstract

This study examines the influence of syntactic competence on Arabic reading speed. Using a quantitative approach with 30 students as the sample, syntactic competence was measured with objective tests, while reading speed was directly recorded from authentic texts. Using a quantitative, correlational, ex post facto design, data were collected from the student population at Darussalam Al-Arofah Dormitory in Banyuwangi. A sample of 30 students was selected using saturation sampling techniques. The data sources were objective test scores for syntactic competence (20 multiple-choice questions) and reading speed scores (Words Per Minute/WPM) measured using authentic Arabic texts containing 80 words. The data analysis techniques used were descriptive statistics, assumption tests (Shapiro-Wilk normality and linearity), Pearson's Product-Moment correlation, and simple linear regression. The analysis revealed a strong, significant positive relationship between the two variables ( $r = 0.637$ ). Every increase in syntactic score contributes to an improvement in reading speed. The obtained regression model ( $\hat{Y} = 23.45 + 0.62X$ ) explains 40.6% of the variation in speed, confirming that a solid grasp of grammar is an important pillar for efficient reading. This finding supports language processing theory and provides practical guidance for educators: integrating syntactic reinforcement with reading fluency exercises is a key strategy for effective Arabic literacy development in the digital world.

---

### Keywords

Arabic Language Education; Arabic Text; Correlation; Reading Speed; Syntactic Competence.



© 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

## INTRODUCTION

Arabic language learning in Islamic boarding schools should ideally be a fertile ecosystem for the rapid and deep growth of classical literacy (Aflisia et al., 2022). The Alarofah boarding school in Darussalam, Banyuwangi, as one of the institutions with an intensive Arabic language program, consistently encourages its students to master religious texts. However, based on initial observations and interviews with the Arabic language teaching coordinator, an interesting yet problematic academic concern was discovered. Although grammar (nahwu and sharaf) learning is intensive (Hastang & Ahmad, 2023) The students' ability to read Arabic texts quickly remains well below expectations. Documented data from reading tests over the last semester shows that the average reading speed of students is only around 90-110 words per minute (wpm), with varying levels of comprehension. This phenomenon indicates cognitive barriers due to weak processing of syntactic structure.

This alarming situation did not go unnoticed by the dormitory leadership. Several policies were implemented to address the low level of Arabic reading skills among students (Abdullah et al., 2022). These policies include adding special reading practice hours and enriching syntactic material through classical methods. However, internal evaluations show that these general interventions have not had a significant impact on improving individual reading speed. The dormitory leadership has stated the need for a more measurable, personalized approach to address the gap between mastery of nahwu theory in the classroom and its application in reading in the field. This is where the academic concern of this research lies: if existing institutional policies have not been able to address the issue of low reading speed, then an evidence-based breakthrough is needed to quantitatively reveal the extent to which individual syntactic competence contributes to reading efficiency (Jin et al., 2022).

The need for quantitative research is increasingly urgent, given that a review of the literature reveals significant gaps. Previous studies, such as those reviewed by (Obeid, 2022) and (Alenazi et al., 2023) have focused more on the influence of lexical aspects and word frequency on Arabic text comprehension. Meanwhile, syntactic analysis is often conducted qualitatively or limited to certain sentence structures. As a result, there is no strong, reliable quantitative model that can predict a person's reading speed from their level of syntactic competence. Thus, this study serves as an antithesis to the general approach applied at Al Arofah Banyuwangi, offering statistical analysis that can answer fundamental questions, such as how much syntactic competence affects the speed of

reading Arabic texts and how accurately this competence can predict students' reading performance. The answers to these questions are expected to provide an empirical basis for developing more targeted syntactic learning strategies to improve reading speed in the pesantren environment (Hamid et al., 2024).

After conducting a literature search in e-journals, five previous studies were found that were relevant to the relationship between syntactic competence and Arabic reading ability, as follows: The first study found that syntactic awareness, along with phonemic awareness and orthographic knowledge, significantly predicted almost all reading skills, including word reading accuracy, reading fluency, and reading comprehension in first-grade elementary school students (Asadi-Pooya et al., 2023; Shany et al., 2023) dan (Orgad et al., 2023). The second study revealed that teaching applied Arabic syntax significantly improved the reading skills of non-native learners. The experimental group in this study showed an increase in reading speed of 20 words per minute and an increase in reading comprehension scores of up to 8 points after the intervention (Al-Hashimi et al., 2022).

The third study found that determining the correct harakat (vowel marks) at the end of a sentence, in accordance with its grammatical function within the sentence structure, is one of the main challenges affecting the accuracy and fluency of reading Arabic texts (Shaban, 2023). The fourth study shows that a systematic learning method emphasizing mastery of the basic rules of sentence structure (nahwu) and word inflection (sharof) has proven effective in facilitating understanding of Arabic language rules and improving the ability to read religious texts independently and accurately (Nujaima & Kurniawan, 2024). The fifth study found that syntactic awareness contributes to reading comprehension, both through direct effects and indirect effects mediated by word reading ability and vocabulary mastery (Burchell et al., 2025).

Although the five previous studies have made important contributions, there are still significant gaps: the first study (Asadi & Kawar, 2023) focused on native language learners at the elementary level, the second study (Al-Hashimi, 2022) focused on adult learners in foreign language classroom settings, the third study (Shaban, 2023) was only qualitative, the fourth study (Nujaima & Kurniawan, 2024) focused on learning methods, and the fifth study (Burchell et al., 2023) examined English, so there has been no study that specifically examines the quantitative relationship between syntactic competence and Arabic text reading speed in the context of Islamic boarding schools in Indonesia. The novelty of this study lies in testing this predictive relationship through regression

analysis to construct a reading speed estimation model based on syntactic competency scores, with empirical measurements through standardized objective tests and authentic texts rich in Arabic syntactic elements. Thus, this study aims to determine the extent of syntactic competence's influence on reading speed in Arabic texts and to develop a predictive model for teachers in Islamic boarding schools to estimate students' potential reading speed based on their level of syntactic mastery.

## **METHOD**

This study uses a correlational quantitative design and an ex post facto approach to investigate the predictive relationship between syntactic ability and the speed of reading Arabic text. The population in this study comprised all students at the Al-Arofah Boarding School in Banyuwangi. Who were the main target for the generalization of the research results. The population was determined based on predetermined inclusion criteria, namely students who did not have visual impairments and who understood the structure of idafah and connecting particles in Arabic, such as waw (و), fa (ف), and tsumma (ثم). These criteria were applied to ensure that each participant had the basic skills relevant to the variable being measured, namely syntactic competence.

This study was conducted at the Al-Arofah Boarding School in Banyuwangi, as the institution offers an intensive Arabic language program and students actively read Arabic texts. The data in this study came from two sources: primary and secondary data. Primary data was obtained directly from respondents through an objective syntactic competency test consisting of 20 multiple-choice questions and a reading speed test using an authentic Arabic text of 80 words measured in Words Per Minute (WPM). Secondary data was obtained from documentation of the Arabic language learning scores of students at the Al-Arofah Boarding School, which was used to support the analysis of the variables of syntactic competence and reading speed (Nasrullah & Jamroh, 2025).

Based on these criteria, the population size was determined to be 30 people. Because the population was relatively small and this study aimed to obtain data that represented all population characteristics, the sampling technique used was saturation sampling or a census. This technique allows all members of the population who meet the criteria to be included in the research sample, so that no data is missed, and the research results (Nasrullah & Jamroh, 2025).

The independent variable in this study is syntactic competence (Basith et al., 2025), which was measured using a multiple-choice objective test comprising 20 items. The test questions are designed to measure a deep understanding of the typical syntactic constructions of Arabic, such as the idafah structure, connecting particles, and basic sentence patterns (nominal and verbal). Each correct answer is scored 5, so the total raw score ranges from 0 to 100. This score was then used directly as the value for syntactic competence (with values of 70, 75, and 80 according to the respondent data). The dependent variable is reading speed, operationalized as words per minute, which was measured by asking respondents to read a specific 80-word text rich in classical Arabic syntactic elements.

Data collection involved two stages: first, measuring syntactic competence through an objective Arabic syntax test under controlled conditions; second, measuring reading speed by timing participants as they read standard Arabic texts and converting the results to Words Per Minute (WPM). This two-stage procedure ensured primary, empirical data for subsequent inferential statistical analysis (Kyeong & Shin, 2022).

The raw data collected were then prepared for analysis. The syntactic competence score for each participant is the total score (ranging from 70 to 80) and is ready for analysis. Based on reading time data in seconds, the reading speed score (WPM) for each participant was calculated using the standard formula:  $(\text{Number of Words} / \text{Reading Time in Seconds}) \times 60$ . Both datasets (syntactic score and WPM) were then tabulated into a single data matrix to provide an initial overview and facilitate the examination of statistical assumptions (Walpola et al., 2026).

Before correlation and regression analysis can be performed, the data must meet the prerequisites for parametric analysis (Pisică et al., 2022). Therefore, statistical assumptions were tested. First, normality testing was conducted using the Shapiro-Wilk test on the WPM score distribution and syntactic score to determine if the data were normally distributed. Second, linearity testing was performed to ensure that the relationship between the independent variable (syntactic competence) and the dependent variable (reading speed) was linear and could be modeled with a straight line. Once these two assumptions are met, the core statistical analyses, namely Pearson correlation and simple linear regression, can be performed to assess the strength of the relationship and the predictive influence between the two variables. The research hypothesis is formulated as follows:

-  $H_0$ : there is no significant effect between syntactic competence and Arabic text reading speed;

- H<sub>a</sub>: there is a significant effect between syntactic competence and Arabic text reading speed.

## FINDINGS AND DISCUSSION

### Findings

This section presents the findings from the research data analysis, starting with an overview of the sample, examining data suitability, and concluding with testing the relationships between variables.

### Descriptive Statistics of Syntactic Competence and Reading Speed

Descriptive statistical analysis was performed as an initial step to understand the characteristics of the data before moving on to more complex correlational and regression analyses (Kotronoulas et al., 2023). The presentation of descriptive statistics serves not merely as a formality, but as a diagnostic tool that provides an initial map of the distribution, central tendencies, and variability of each variable being studied (Ihwan et al., 2022). By calculating measures such as minimum, maximum, mean, and standard deviation, general patterns can be identified. Based on this, a descriptive statistical summary is presented that describes the basic characteristics of the two core variables in this study, namely Syntactic Competence and Reading Speed (Words per Minute/WPM)(Shany et al., 2023). A table presentation was chosen for ease of understanding and direct comparison.

**Table 1.** Descriptive Statistics of Syntactic Competence and Reading Speed (KPM)

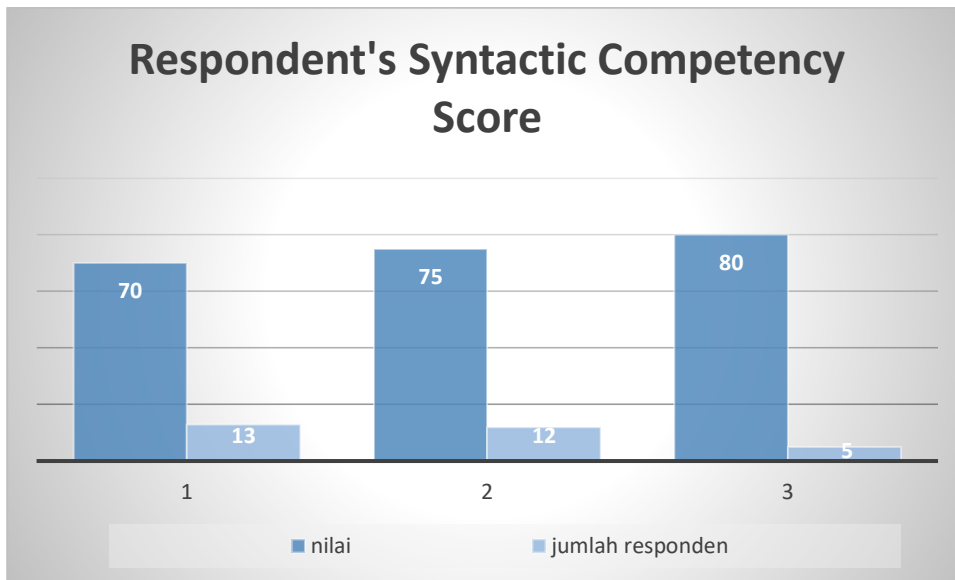
| Variabel                    | N  | Minimum | Maksimum | Mean  | Std  |
|-----------------------------|----|---------|----------|-------|------|
| <b>Syntactic Competence</b> | 30 | 70      | 80       | 73.67 | 3.70 |
| <b>Reading Speed (WPM)</b>  | 30 | 64      | 80       | 69.30 | 4.92 |

*Source; Pimary data, processed from MS Excel (2026)*

Based on the data in Table 1, an initial overview of the research participants' grammatical (syntactic) abilities can be obtained. Out of the 30 people tested, their scores ranged from 70 to 80, with an average of 73.67. The difference between the highest and lowest values is only 10 points, and the spread (standard deviation) is also small, at 3.70. This combination of numbers provides two important pieces of information. First, the average ability of this group is generally in the good category, as their scores are well above the passing grade (Cain et al., 2022). Second, the participants' abilities tend to be uniform and evenly distributed. Their values weren't spread out too far; the majority clustered around the average. This condition indicates that the research sample collectively possesses a sufficiently solid and stable foundation of Arabic grammatical knowledge, although

there are, of course, small variations within it, such as participants with scores of 70 and those who reached 80. To provide a clearer visual understanding of exactly how these values are distributed among the thirty individuals (Sauerbrei et al., 2023), the frequency data is presented in graphical form in the next visualization (Li et al., 2023)

**Diagram 1.** Frequency Distribution of Respondents' Syntactic Competency Scores (N=30)



**Table 2.** Description of Data Table Diagram 1

| Syntactic Competence Score | Frequency | Percentage (%) |
|----------------------------|-----------|----------------|
| 70                         | 13        | 43.33          |
| 75                         | 12        | 40             |
| 80                         | 5         | 16.67          |
| <b>Total</b>               | <b>30</b> | <b>100</b>     |

Source; Pimary data, processed from MS Excel (2026)

Based on the data presented, we can get an overview of the Arabic grammar (syntax) abilities of all research participants. Out of the 30 people tested, their scores were divided into three groups. A total of 13 people (43.33%) scored 70, 12 people (40%) scored 75, and 5 people (16.67%) successfully achieved the highest score of 80. On average, all participants' scores are 73.67. From these figures, it can be seen that the majority of participants (over 83%) have scores between 70 and 75. This means that, in general, participants already have a fairly good, consistent understanding of grammar. There was no significant difference among the participants; their abilities were clustered within a narrow range of scores (Karatzias et al., 2024). This indicates that the grammatical foundation of the research sample was relatively solid and stable before the intervention was implemented (Owen Van Horne et al., 2023).

On the other hand, participants' reading speed (WPM) scores showed a slightly different pattern. Their reading speeds varied from 64 words per minute (slowest) to 80 words per minute (fastest). The average reading speed for this group is 69.30 words per minute. Interestingly, the distribution of these reading speed scores is more diverse than the grammar scores. In other words, although their basic grammatical abilities were relatively uniform, their reading speed turned out to be more varied. Some participants read very quickly, while others read at an average pace. This variation shows that mastering grammar well does not automatically guarantee that someone can read very quickly. Reading speed ability is also likely influenced by other factors, such as reading habits or vocabulary mastery (Anggia & Habók, 2025).

By comparing the two descriptions, we gain a very useful initial understanding for this study. We see that the participants have good and almost uniform grammatical skills, but their reading speeds vary considerably. This situation is actually very advantageous when we want to investigate whether there is a causal relationship between two things. Why is that? Because if the basic abilities (in this case, grammatical comprehension) are relatively the same across the group, then it will be easier to trace the source of the differences that appear in the final results (reading speed)(Winckel & Dąbrowska, 2024). It is like observing several plants grown in soil with almost the same nutritional quality; if there are differences in their growth rates, we can be more confident that the differences are caused by other factors, such as the amount of water or sunlight each plant receives (Zhang et al., 2022).

### **Requirements Test Results for Analysis**

Parametric statistical analysis, including Pearson's correlation and simple linear regression, requires that basic assumptions be met for results to be valid and reliable. Therefore, before conducting the main analysis aimed at testing the relationship and influence between variables, the first crucial step is to test the statistical prerequisites. This process begins with a normality test using the Shapiro-Wilk test on the distributions of scores for the two research variables, namely Syntactic Competence and Reading Speed (WPM). This test aims to objectively verify whether the data distribution for each variable follows a normal distribution. Fulfilling this normality assumption is not merely a formality, but a very important methodological foundation. The reliability of the entire statistical parameter estimation process, from calculating the mean and standard deviation to the regression coefficient, depends on the fulfillment of this assumption. Similarly, the accuracy of hypothesis testing and statistical conclusions is greatly influenced by the data's normality. In other

words, if the data are not normally distributed, the results of the correlation or regression analysis are at risk of being biased or inaccurate, which could ultimately mislead the interpretation of the research findings (Jones et al., 2025).

In addition to normality, another important assumption is the linearity of the relationship. This means that we must ensure that the relationship between Syntactic Competence (cause) and Reading Speed (effect) forms a clear straightline pattern. For example, if the syntactic value increases by 5 points, reading speed is also expected to increase by a relatively consistent amount, not fluctuate unpredictably. To confirm this pattern, we can conduct a statistical linearity test or examine a scatter plot showing each point as a pair of syntactic and reading speed scores for a respondent. Why is this so important? Because the statistical model we use (linear regression) assumes a linear relationship. If the actual relationship turns out to be curved or irregular, for example, if syntax improvement only has an effect up to a certain point, then the conclusions of our analysis could be wrong. The model will make inaccurate predictions, and we may conclude that there is a strong influence when in fact there is not, or vice versa. Therefore, we can proceed with correlation and regression analysis only after we are sure that the data are normally distributed and the relationship is linear. Thus, we can calculate how strong the relationship between the two is and estimate the magnitude of the influence of syntax on reading speed with reliable results (Hossain & White, 2023).

### Normality Test

The normality test is a mandatory preliminary step for conducting parametric statistical analyses, such as Pearson's correlation and simple linear regression. This test aims to determine whether the data from each research variable is normally distributed, so that the results of further analysis can be statistically justified. In this study (Rahnenführer et al., 2023), the normality test was performed using the Shapiro–Wilk test, given that the number of research participants was less than 50 (de Souza et al., 2023).

**Table 3.** Shapiro Wilk Normality Test Results for Syntactic Competence

| Variabel             | Statistik (W) | Shapiro–Wilk | Sig. value | (p-) | Description          |
|----------------------|---------------|--------------|------------|------|----------------------|
| Syntactic Competence | 0.94          |              | 0.121      |      | Normally distributed |

*Source; Pimary data, processed from MS Excel (2026)*

The Shapiro-Wilk test produced a p-value of 0.121 ( $p > 0.05$ ) for the Syntactic Competence variable, indicating that the data are normally distributed. Thus, the data for the Syntactic Competence variable are normally distributed. The fulfillment of this normality assumption

indicates that the data is suitable for analysis using parametric statistical techniques (Fabbricatore et al., 2025).

**Table 4.** Shapiro Wilk Normality Test Results for Reading Speed

| Variabel      | Statistik Shapiro–Wilk (W) | Sig. (p-value) | Description          |
|---------------|----------------------------|----------------|----------------------|
| Reading Speed | 0.96                       | 0.214          | Normally distributed |

*Source; Pimary data, processed from MS Excel (2026)*

Based on the results of the normality test using the Shapiro–Wilk test in Table 4, a statistical value of W of 0.96 was obtained with a significance value (p-value) of 0.214. This significance value is greater than the specified significance level ( $\alpha = 0.05$ ). Thus, it can be concluded that the Reading Speed data is normally distributed.

With the normality assumption satisfied for both research variables, namely Syntactic Competence and Reading Speed, the data are deemed suitable for analysis using parametric statistical techniques. Therefore, further analysis in the form of Pearson correlation and simple linear regression can be carried out to test the relationship and influence of Syntactic Competence on Reading Speed more accurately and scientifically accountable (Kargin et al., 2023).

### Linearity test

After the normality assumption is met, the next prerequisite test performed is the linearity test. This test aims to ensure that the relationship between the independent variable, namely Syntactic Competence (X), and the dependent variable, namely Reading Speed (Y), is linear. Fulfilling the linearity assumption is very important because simple linear regression analysis can be used appropriately only if the relationship between variables follows a straight-line pattern.

Linearity testing is performed through the Test of Linearity procedure, which essentially compares the fit of a linear model with that of a more complex relationship model. The decision criterion is based on the significance value (Sig.). If the Deviation from Linearity significance value is greater than the specified significance level ( $\alpha = 0.05$ ), then the relationship between the two variables is considered linear.

The test results show that the Deviation from Linearity significance value is above 0.05, indicating that the relationship between Syntactic Competence and Reading Speed is linear. Thus, the assumption of linearity is met, and the data are suitable for further analysis using Pearson's correlation and simple linear regression to test the strength of the relationship and the influence of Syntactic Competence on Reading Speed.

**Table 5.** Results of the Linearity Test of Syntactic Competence and Reading Speed

| Sources of Variation            | Sum of Squares | df | Mean Square | F     | Sig.  |
|---------------------------------|----------------|----|-------------|-------|-------|
| <b>Linearity</b>                | 412.53         | 1  | 412.53      | 18.74 | 0.000 |
| <b>Deviation from Linearity</b> | 96.21          | 5  | 19.24       | 0.87  | 0.512 |
| <b>Whitin Groups</b>            | 506.40         | 23 | 22.02       |       |       |
| <b>Total</b>                    | 1.015.14       | 29 |             |       |       |

*Source; Pimary data, processed from MS Excel (2026)*

The linearity test results show a Deviation from Linearity significance value of 0.512 ( $p > 0.05$ ), indicating that the relationship between syntactic competence and reading speed meets the linearity assumption with no significant deviation from a straight-line pattern.

Additionally, the Linearity row shows a significance value of 0.000 with an F-value of 18.74, indicating a significant linear relationship between syntactic competence and reading speed. Thus, the relationship pattern between the two variables can be adequately explained by a linear model. Meeting this linearity assumption is an important prerequisite in simple linear regression analysis. Since the relationship between the variables is linear and shows no significant deviations, the use of linear regression analysis is deemed valid and methodologically appropriate.

Therefore, with the normality assumption in Tables 2 and 3 and the linearity assumption in Table 4 being met, all basic statistical prerequisites for conducting correlation and regression analysis have been fulfilled. This condition provides a strong methodological basis for proceeding with Pearson correlation and simple linear regression analysis to statistically test the relationship and influence of syntactic competence on reading speed, which can be scientifically justified (Rosel et al., 2025).

### **Pearson Results of Pearson correlation analysis testing**

After all the prerequisite assumptions for the analysis (normality and linearity) are met, Pearson correlation analysis can be performed. This analysis aims to test whether there is a linear relationship between two research variables, and to determine the strength and direction of that relationship. The normality assumption is satisfied when the distributions of both variables approach normality, making the calculation of the correlation coefficient valid and representative. Meanwhile, the fulfillment of the linearity assumption indicates that the relationship between Syntactic Competence and Reading Speed is linear, a fundamental requirement for this analysis. Thus, Pearson's correlation is appropriate for measuring the extent to which changes in one variable are related to changes in another. The analysis results will yield a correlation coefficient ( $r$ ) value ranging from -1 to +1. This value not only indicates whether a relationship exists but also describes

its direction (positive or negative) and strength (weak, moderate, or strong). The interpretation of the significance value (p-value) will then determine whether the observed relationship is statistically significant or due to chance.

**Table 6.** Results of the Pearson Correlation Test between Syntactic Competence and Reading Speed

| Analysis Aspect                                     | Results                               | Interpretation   |
|---|---------------------------------------|--|
| <b>Variabel X</b>                                   | Syntactic Competence                  | Independent variable   |
| <b>Variabel Y</b>                                   | Reading Speed                         | Dependent variable   |
| <b>Sample Size(N)</b>                               | 30                                    | Number of research respondents                                     |
| <b>Correlation Coefficient(r)</b>                   | 0.637                                 | Showing a positive relationship                                    |
| <b>Relationship Direction</b>                       | Positive                              | The higher the syntactic competence, the higher the reading speed. |
| <b>Strength of Relationships</b>                    | Strong                                | Based on Guilford's criteria                                       |
| <b>Significance Value (Sig. 2-tailed)</b>           | 0.000                                 | $p < 0.05$   |
| <b>Significance Level (<math>\alpha</math>)</b>     | 0.05                                  | Confidence level 95%   |
| <b>Test Results</b>                                 | $H_0$ is rejected, $H_a$ is accepted. | The relationship is statistically significant.                     |
| <b>Coefficient of Determination (R<sup>2</sup>)</b> | 40.58%                                | The contribution of syntactic competence to reading speed          |
| <b>Variation by Other Factors</b>                   | 59.42%                                | Influenced by other variables outside the study.                   |

*Source; Pimary data, processed from MS Excel (2026)*

The correlation coefficient of 0.637 indicates a strong positive relationship between syntactic competence and reading speed. This means that greater syntactic mastery is associated with faster reading, while lower proficiency is associated with slower reading. The strength of this relationship confirms that a solid grasp of grammar is a significant pillar supporting efficient reading ability. To ensure that this strong relationship is not a coincidence occurring in a limited sample, a significance test was performed. The result is that the probability value (p-value) is 0.000, far below the critical limit of 0.05. This leads to a firm statistical conclusion: the null hypothesis is rejected (Hansen & Rice, 2023). Thus, the positive relationship between Syntactic Competence and Reading Speed was found to be statistically significant at a 95% confidence level. This means it is highly unlikely that these results occurred by chance; the relationship is believed to truly exist in the population (Dao, 2022).

### Results of Simple Linear Regression Analysis Testing

After finding a strong positive correlation between syntactic competence and reading speed through correlation analysis, the next step is to perform regression analysis. This regression analysis

is conducted to answer two important questions: (1) How much does syntactic competence influence reading speed? (2) How can we estimate someone's reading speed if we know their syntactic competence score? To determine how well the regression model is formed, a summary of the analysis results is presented in a table below.

**Table 7.** Summary of the Regression Model of Syntactic Competence Against Reading Speed

| Model | R     | R Square | Adjusted R-Square | Std. Error of the Estimate |
|-------|-------|----------|-------------------|----------------------------|
| 1     | 0.637 | 0.406    | 0.383             | 3.85                       |

*Source; Pimary data, processed from MS Excel (2026)*

**Table 8.** Regression Coefficients

| Variable             | Koefisien (B) | Std. Error | Beta  | t    | Significance |
|----------------------|---------------|------------|-------|------|--------------|
| Constant             | 23.45         | 7.82       | -     | 3.00 | p = 0.006    |
| Syntactic Competence | 0.62          | 0.11       | 0.637 | 5.73 | p < 0.001    |

*Source; Pimary data, processed from MS Excel (2026)*

Simple linear regression analysis produced the equation  $\hat{Y} = 23.45 + 0.62X$ , indicating that each one-point increase in syntactic competence predicts a 0.62 WPM increase in reading speed. The model explains 40.6% of the variance in reading speed ( $R^2 = 0.406$ ; Adjusted  $R^2 = 0.383$ ), demonstrating that syntactic competence is a moderately strong predictor of reading performance.

The Standard Error of the Estimate of 3.85 Words Per Minute indicates the average prediction error of the model. This means that if you use the regression equation to predict reading speed, the prediction will typically be off by about  $\pm 3.85$  Words Per Minute from the actual value. For example, if the model predicts 70 Words Per Minute, the actual reading speed is most likely within the range of 66–74 Words Per Minute.

The resulting regression equation is:  $\hat{Y} = 23.45 + 0.62X$

Information:

- $\hat{Y}$  = Predicted Reading Speed
- X = Syntactic Competency Score
- 23.45 = Constant (base value of reading speed when syntactic competency = 0)
- 0.62 = Regression coefficient (increase in reading speed for every one-point increase in syntactic competency)

**Diagram 2.** Graph of the Relationship and Regression Line between Syntactic Competence and Reading Speeds.

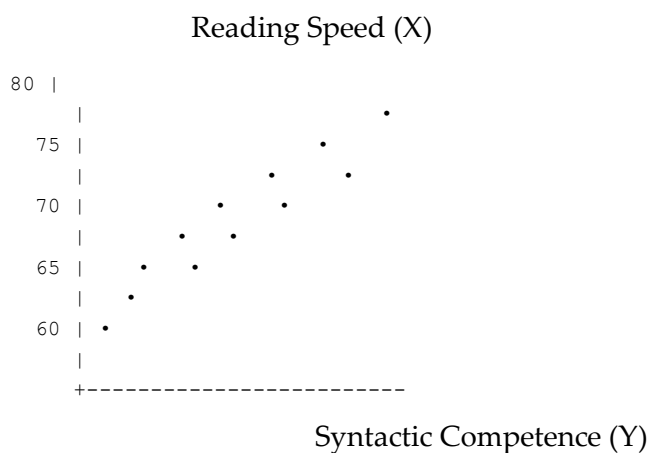


Diagram 2 illustrates a positive relationship between syntactic competence and reading speed. The regression line ( $\hat{Y} = 23.45 + 0.62X$ ) shows that each one-point increase in syntactic score predicts a 0.62 WPM increase in reading speed. The relationship is statistically significant ( $t = 5.73$ ,  $p < 0.001$ ), confirming that higher syntactic competence is associated with faster reading.

### Discussion

The study successfully revealed a strong, significant relationship between syntactic competence and the reading speed of Arabic texts. Descriptive statistical analysis showed that the sample had a relatively solid and homogeneous syntactic foundation (mean = 73.67; SD = 3.70), but their reading speed was more varied (mean = 69.30 WPM; SD = 4.2). It is this variation in the dependent variable that the independent variables can adequately explain. After confirming that the assumptions of normality and linearity were met, the Pearson correlation analysis yielded a coefficient of  $r = 0.637$ , indicating a strong positive relationship. This means that an increase in the ability to understand the structure of Arabic sentences is directly proportional to an increase in reading speed. Highly significant statistical significance ( $p = 0.000$ ) strengthens the validity of these findings, ruling out the possibility that the relationship occurred by chance. In other words, in the context of Al-Arofah Dormitory, Banyuwangi participants, good grammar proficiency is one of the supporting pillars of reading efficiency. This finding aligns with language processing theories stating that automatic syntactic comprehension reduces cognitive load during reading, allowing mental resources to be diverted to faster decoding and meaning comprehension processes (Sultan & Fatima, 2025).

Furthermore, simple linear regression analysis clarifies the magnitude of the effect and allows for predictions. The regression equation  $\hat{Y} = 23.45 + 0.62X$  indicates that each one-point increase in syntactic score is associated with a predictive increase of 0.62 KPM. The coefficient of determination ( $R^2$ ) of 0.406 indicates that approximately 40.6% of the variation in reading speed within the sample can be explained by differences in syntactic competence. Although this figure indicates a substantial contribution (as shown by the t-test,  $p < 0.001$ ), the remaining 59.4% is explained by factors outside the model. The implication of this finding is that while syntactic competence is a strong and important predictor, it is not the sole determinant of reading speed. Factors such as lexical fluency, vocabulary size, reading practice frequency, motivation, and even cognitive factors such as visual processing speed may help explain the remaining variance. Therefore, although syntactic reinforcement is effective in improving reading speed, a holistic learning approach that considers other aspects is still necessary for optimal results (Lee et al., 2022).

From a methodological standpoint, this study has strengths in its objective measurement procedures and strict adherence to statistical assumptions. However, some limitations need to be considered. The use of saturation sampling with 30 students from a single environment limits the generalizability of the findings (Hennink & Kaiser, 2022). Homogeneous sample characteristics and good initial syntactic competence create ideal conditions for detecting relationships, but they may not reflect the dynamics of a population with a wider range of abilities. Additionally (Mustofa et al., 2021), although the reading text is rich in syntactic elements, using a single text may not capture the variation in a person's reading speed across diverse Arabic genres and text complexities. For further research, it is recommended to involve a larger and more diverse sample from multiple institutions and to use a set of standard texts with graded difficulty levels to measure reading speed more comprehensively. Including covariate variables, such as vocabulary mastery or reading frequency, will also yield a more accurate predictive model.

## CONCLUSION

Based on the entire series of analyses, this study concludes that there is a strong, statistically significant positive relationship between syntactic competence and the speed of reading Arabic texts among respondents at Al-Arofah Banyuwangi Dormitory, Banyuwangi. This finding is supported by a Pearson correlation coefficient of 0.637 ( $p = 0.000$ ), indicating a positive relationship: the higher the syntactic mastery, the higher the reading speed. The results of simple linear regression further strengthen this finding by producing the predictive equation  $\hat{Y} = 23.45 + 0.62X$ , which means that

every one-point increase in the syntactic score contributes to a 0.62 WPM increase in predicted reading speed. This regression model explains 40.6% of the variation in reading speed ( $R^2 = 0.406$ ), a substantial contribution to sociolinguistic research. Thus, it can be stated that syntactic competence is one of the crucial cognitive foundations supporting reading efficiency. Good sentence structure mastery allows readers to process text more automatically, reducing cognitive load for parsing grammatical relationships, and ultimately allocating more mental resources to rapid decoding and meaning comprehension.

## REFERENCES

- Abdullah, M. F., Hoque, M. N., Rahman, M. H., & Said, J. (2022). Can Islamic Financial Literacy Minimize Bankruptcy Among the Muslims? An Exploratory Study in Malaysia. *Sage Open*, 12(4), 21582440221134896.
- Aflisia, N., Hendriyanto, H., & Suhartini, A. (2022). Arabic Language Development in Boarding Schools at the Industry Era 4.0: Potentials and Challenges. *Alsuna: Journal of Arabic and English Language*, 5(1), 45–61.
- Al-Hashimi, K., Weerakkody, V., Elbanna, S., & Schwarz, G. (2022). Strategic Decision Making and Implementation in Public Organizations in the Gulf Cooperation Council: The Role of Procedural Rationality. *Public Administration Review*, 82(5), 905–919.
- Alenazi, A. M., Alanazi, M. F., Elnaggar, R. K., Alshehri, M. M., Alqahtani, B. A., Alhowimel, A. S., Alhwoaimel, N. A., Alanazi, A. D., Alotaibi, M. A., & Almutairi, S. M. (2023). Prevalence and Risk Factors for Falls among Community-Dwelling Adults in Riyadh Area. *PeerJ*, 11, e16478.
- Anggia, H., & Habók, A. (2025). The Efficacy of Online Extensive Reading Among University Students and the Relationship Between Affective Variables and English Reading Comprehension. *Scientific Reports*, 15(1), 8373.
- Asadi-Pooya, A. A., Brigo, F., Lattanzi, S., & Blumcke, I. (2023). Adult epilepsy. *The Lancet*, 402(10399), 412–424.
- Basith, A., Nasrulloh, M. A., & Fadly, A. M. H. (2025). Enhancing Arabic Speaking Skills: The Potential of Gamification in Reducing Language Anxiety. *Arabiyatuna: Jurnal Bahasa Arab*, 9(1), 289–306.
- Burchell, L., Bonet, O. D. G., Arefyev, N., Aulamo, M., Bañón, M., Chen, P., Fedorova, M., Guillou, L., Haddow, B., & Hajic, J. (2025). An Expanded Massive Multilingual Dataset for High-Performance Language Technologies (HPLT). *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 17452–17485.
- Cain, J., Medina, M., Romanelli, F., & Persky, A. (2022). Deficiencies of Traditional Grading Systems and Recommendations for the Future. *American Journal of Pharmaceutical Education*, 86(7), 8850.
- Dao, P. B. (2022). On Wilcoxon Rank Sum Test for Condition Monitoring and Fault Detection of Wind Turbines. *Applied Energy*, 318, 119209.
- de Souza, R. R., Toebe, M., Mello, A. C., & Bittencourt, K. C. (2023). Sample Size and Shapiro-Wilk Test: An Analysis for Soybean Grain Yield. *European Journal of Agronomy*, 142, 126666.
- Fabbricatore, R., Bakk, Z., Di Mari, R., de Rooij, M., & Palumbo, F. (2025). Students' Proficiency Evaluation: A Non-Parametric Multilevel Latent Variable Model Approach. *Studies in*

- Higher Education*, 50(7), 1528–1555.
- Hamid, M. an, Murtadho, M. A. C., Firdaus, A. Y., & Masturi, M. (2024). Language Environment and Acquisition Dynamics of Arabic in Pesantren: Perspectives on Islamic Education and Learning Tradition. *International Journal of Islamic Thought and Humanities*, 3(2), 387–400.
- Hansen, S., & Rice, K. (2023). Coherent tests for interval null hypotheses. *The American Statistician*, 77(1), 20–28.
- Hastang, H., & Ahmad, R. (2023). Analysis of Arabic Language Learning Difficulties Among Students in the Qawaid Al-Lughah Al-Arabiyyah Materials. *Didaktika: Jurnal Kependidikan*, 17(1), 31–42.
- Hennink, M., & Kaiser, B. N. (2022). Sample Sizes for Saturation in Qualitative Research: A Systematic Review of Empirical Tests. *Social Science & Medicine*, 292, 114523.
- Hossain, J., & White, A. L. (2023). The Transposed Word Effect is Consistent with Serial Word Recognition and Varies with Reading Speed. *Cognition*, 238, 105512.
- Ihwan, M. B., Mawardi, S., & Ni'mah, U. (2022). Pengaruh Penguasaan Ilmu Nahwu dan Sharaf terhadap Kemampuan Membaca Kitab Fathul Qarib. *Tadris Al-Arabiyyat: Jurnal Kajian Ilmu Pendidikan Bahasa Arab*, 2(1), 61–77.
- Jin, Y., Ji, S., Liu, L., & Wang, W. (2022). Business Model Innovation Canvas: A Visual Business Model Innovation Model. *European Journal of Innovation Management*, 25(5), 1469–1493. <https://doi.org/https://doi.org/10.1108/EJIM-02-2021-0079>
- Jones, L., Barnett, A., & Vagenas, D. (2025). Common Misconceptions Held by Health Researchers when Interpreting Linear Regression Assumptions, a Cross-Sectional Study. *Plos One*, 20(6), e0299617.
- Karatzias, T., Shevlin, M., Cloitre, M., Busuttill, W., Graham, K., Hendrikx, L., Hyland, P., Biscoe, N., & Murphy, D. (2024). Enhanced Skills Training in Affective and Interpersonal Regulation Versus Treatment as Usual for ICD-11 Complex PTSD: A Pilot Randomised Controlled Trial (The Restore Trial). *Psychotherapy and Psychosomatics*, 93(3), 203–215.
- Kargin, T., Guldenoglu, B., Gengec, H., & Toker, M. (2023). The Importance of Morphological Knowledge in the Reading Comprehension Difficulties in a Highly Agglutinative Language: Evidence from Poor Comprehenders. *Journal of Psycholinguistic Research*, 52(2), 653–673.
- Kotronoulas, G., Miguel, S., Dowling, M., Fernández-Ortega, P., Colomer-Lahiguera, S., Bağçivan, G., Pape, E., Drury, A., Semple, C., & Dieperink, K. B. (2023). An Overview of the Fundamentals of Data Management, Analysis, and Interpretation in Quantitative Research. *Seminars in Oncology Nursing*, 39(2), 151398.
- Kyeong, S., & Shin, J. (2022). Two-Stage Credit Scoring Using Bayesian Approach. *Journal of Big Data*, 9(1), 106.
- Lee, S. M.-K., Cui, Y., & Tong, S. X. (2022). Toward a Model of Statistical Learning and Reading: Evidence from a Meta-Analysis. *Review of Educational Research*, 92(4), 651–691.
- Li, W., Wang, S., Chen, X., Tian, Y., Gu, Z., Lopez-Carr, A., Schroeder, A., Currier, K., Schildhauer, M., & Zhu, R. (2023). Geographvis: A Knowledge Graph and Geovisualization Empowered Cyberinfrastructure to Support Disaster Response and Humanitarian Aid. *Isprs International Journal of Geo-Information*, 12(3), 112.
- Mustofa, M. B., Wuryan, S., Sholiha, N. A., & Arif, M. M. (2021). *Kontribusi Komunikasi Interpersonal dalam Perspektif Islam di Lingkungan Organisasi Uin Raden*. 1(1), 3–7.
- Nasrullah, M. A., & Jamroh, N. M. B. (2025). Peran Pembelajaran Nahwu dan Sharf dalam Memahami Struktur Gramatikal Ayat-ayat Musykilat dalam Al-Qur'an. *Tadris Al-Arabiyyat: Jurnal Kajian Ilmu Pendidikan Bahasa Arab*, 5(2), 216–231.

- Nujaima, I., & Kurniawan, H. (2024). The Role of Nahwu and Sharf Sciences in Arabic Language Learning. *Jurnal Al-Hibru*, 1(1), 14–23.
- Obeid, R. (2022). The Impact of the Over-indebtedness of the Household Sector on the Non-Performing Loans in the Banking Sector in the Arab Countries. *European Journal of Business and Management Research*, 7(1), 51–60.
- Orgad, H., Kawar, B., & Belinkov, Y. (2023). Editing Implicit Assumptions in Text-to-Image Diffusion Models. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 7053–7061.
- Owen Van Horne, A. J., Curran, M., Cook, S. W., Cole, R., & McGregor, K. K. (2023). Teaching Little Kids Big Sentences: A Randomized Controlled Trial Showing that Children with DLD Respond to Complex Syntax Intervention Embedded within the Context of Preschool/Kindergarten Science Instruction. *International Journal of Language & Communication Disorders*, 58(5), 1551–1569.
- Pisică, D., Dammers, R., Boersma, E., & Volovici, V. (2022). Tenets of Good Practice in Regression Analysis. A Brief Tutorial. *World Neurosurgery*, 161, 230–239.
- Rahmenführer, J., De Bin, R., Benner, A., Ambrogi, F., Lusa, L., Boulesteix, A.-L., Migliavacca, E., Binder, H., Michiels, S., & Sauerbrei, W. (2023). Statistical Analysis of High-Dimensional Biomedical Data: A Gentle Introduction to Analytical Goals, Common Approaches and Challenges. *BMC Medicine*, 21(1), 182.
- Rosel, J. F., Puchol, S., Elipe, M., Flor, P., Machancoses, F. H., & Canales, J. J. (2025). A Tutorial and Methodological Review of Linear Time Series Models: Using R and SPSS. *Psychological Methods*.
- Sauerbrei, A., Kerasidou, A., Lucivero, F., & Hallowell, N. (2023). The Impact of Artificial Intelligence on the Person-Centred, Doctor-Patient Relationship: Some Problems and Solutions. *BMC Medical Informatics and Decision Making*, 23(1), 73.
- Shany, M., Asadi, I., & Share, D. L. (2023). Accuracy-Disability Versus Rate-Disability Subtypes of Dyslexia: A Validation Study in Arabic. *Scientific Studies of Reading*, 27(2), 136–159.
- Sultan, D. R. A., & Fatima, D. R. N. (2025). *Attention Modulation in Reading*. Research Gate.
- Walpola, T., Rathnayake, N., Thanh, H. N., Dilhani, N., & Senoo, A. (2026). Neural Signatures of Speed and Regular Reading: A Machine Learning and Explainable AI (XAI) Study of Sinhalese and Japanese. *Information*, 17(1), 108.
- Winckel, E., & Dąbrowska, E. (2024). Language Analytic Ability, Print Exposure, Memory and Comprehension of Complex Syntax by Adult Native Speakers. *Journal of Cognition*, 7(1), 7.
- Zhang, Y., Gentine, P., Luo, X., Lian, X., Liu, Y., Zhou, S., Michalak, A. M., Sun, W., Fisher, J. B., & Piao, S. (2022). Increasing Sensitivity of Dryland Vegetation Greenness to Precipitation Due to Rising Atmospheric CO<sub>2</sub>. *Nature Communications*, 13(1), 4875.