Channing: Journal of English Language Education and Literature Vol. 9, No. 1, Halaman: 1 – 17, April 2024

Measuring EFL Learners' Ability in Writing Correlation between Linguistic Language Variables and Writing Proficiency

Masrul[1]*, Melyann Melani[2]
[1]Universitas Pahlawan Tuanku Tambusai
[2]State Islamic University Sjech M. Djamil Djambek Bukittinggi
* E-mail: masrulm25@mail.com

This study was administered to examine the correlation between linguistic language variables and writing proficiency among English as a Foreign Language (EFL) learners in grades 1-4 and grade 8. 100 students were participating, whose writing proficiency was evaluated based on five aspects: content, fluency, sentence formation, usage, and mechanics. Meanwhile, students' linguistic skill was measured using five indicators: cognitive, composition, vocabulary, style, and sociocultural. The relationship between linguistic language skills and writing ability was examined using a developmental scale model, and which results indicated a positive correlation in all grades. The findings of this study add to the body of knowledge regarding the relationship between linguistic language skills and writing ability and its implications.
Keywords: Academic Essay; Assessment; EFL Learners; Writing Ability

INTRODUCTION

Writing plays a critical role in academic success, and students put in considerable effort to improve their English writing skills. To accurately assess writing abilities, it is important to define writing as a domain and to have a clear understanding of what is being tested (Jahin, 2012). Despite the vast amount of research and theories surrounding writing, the concept is not always precisely defined (Bachman, 1990). Writing ability, composing, writing process, and writing performance are terms that are often used interchangeably, but they refer to different aspects of writing. In this study, writing ability refers to an individual's innate traits or characteristics that enable them to construct written material (Burdick et al., 2013). However, there is currently no established developmental scale that measures writing ability, hence writing ability has been measured based on the quality of written products. Similarly, essays can provide valuable insights into writing ability just as oral reading passages can provide insights into a reader's cognitive abilities (Goodman, 2014).

To some extent, writing ability can be defined as a similarly intrusive personal characteristic, while writing quality is the characteristics of a product, composition, or status obtained from a particular writing opportunity (Burdick et al., 2013). Learners already have schemata, vocabulary mastery, and writing style that make their writing circumstances differ which also account for different writing abilities. Writing ability also relates to the collective processes of cognitive and socio-cultural composition, such as the implementation of working memory, cognitive functioning, or the process of planning and revising (Kellogg, 2008). A considerable amount of literature has been published on the writing process on cognitive approach (e.g. Allinson & Hayes, 1996; Bereiter & Scardamalia, 1987; Flower & Hayes, 1980). Writing is the production of ideas expressed in written form, where writers must gather and analyze as much information as possible before starting to write (Merjen et al., 2019). In addition, writing ability emphasizes the final result or written product as qualified by the use of words and linguistic forms such as grammar, syntax, and mechanics (Nunan, 1999; Raimes, 1983; Tribble, 2009).

Kellogg (2008) proposed a three-step model of writing ability emphasizing the role of working memory in writing development where writers must retrieve vocabulary from their long-term memory and store it temporarily in working memory, transform the vocabulary into a written text, and monitor and evaluate their work through cognitive functioning. The first step, Knowledge-Telling, involves the representation of information in working memory before the text is generated. The second step, Knowledge-Transforming, is where the text representation becomes more detailed. The last step, Knowledge-Crafting, is where writers are required to maintain and manipulate the text representation in their working memory and to construct a text that can be imagined by the reader. The author's, reader's, and text's representations must be stored in the limited capacity of working memory and monitored continuously (Traxler & Gernsbacher, 1993). Although the three putative developmental stages have some empirical support, there is currently no effective method to quantify individual growth through these stages over the long term.

The current writing assessment applies national and international standardized examinations as a part of quality analysis in instructional evaluation. Several national assessments of writing ability (Applebee, 1986; Rea et al., 1991; Engelhard et al., 1991) and global assessments (Gorman et al., 1988) use learners' essay writing. Large writing assessments are a very important test for candidates with direct results for placement of education (Engelhard, 2009). Therefore, it is important to measure the EFL learners' ability to write academic essays using a developmental scale, since their writing ability can be changed or developed in many ways.

An analytical scale has been widely used by teachers and researchers and it can be used for much more sophisticated and comprehensive purposes. Analytical scales can measure different evaluation levels. The scales can be used in prediction, placement, exemption, or growth, or if we need guidance for informal diagnosis or feedback with valid scores (Cooper, 1977). It is essential to put the present study into discussions on computer analyses since the Common Core State Standards, the Partnership for Assessment of Readiness for College and Careers (PARCC), and the Smarter Balanced Assessment Consortium (SBAC) are aligned and the National Assessment of Educational Progress (NAEP) (Burdick et al., 2013). Writing assessment incorporates the development of new summative high-stakes writing assessments which involve educators in the development of the tests.

Computer analysis assessment helps promptly estimate writing skills based on different situations, including learning contexts, providing rapid, numerous, and consistent evaluation (Graham et al., 2011). Formative and summative writing assessments at the national level are costly and time inefficient. On the other side, the ability to perform a particular writing activity is frequently used in high-stakes evaluations (Beck & Jeffery, 2007). Computer-based programs, such as Intelligent Essay Assessor can facilitate this need (Landauer et al., 1998). The empirical analysis of students' writing abilities based on their essays raises several assessment challenges that are hard to solve using the present assessment. The first issue relates to the standard evaluation processes, such as the use of the traditional Likert scale for essays that are non-dichotomous assessments (Cooper, 1977). Another issue relates to inter-rater reliability. It will be difficult to assess students' work due to complex examinations of writing ability (Engelhard, 2009).

The Writing Ability Developmental Scale addresses the shortcomings of other writing assessment methods, particularly regarding academic essay writing. This innovative scale employs a standardized set of components to create a clear and consistent developmental progression (Burdick et al., 2013). The computer-based assessment evaluates a range of writing features that are considered key indicators of writing proficiency (Attali & Burstein, 2006; Kellogg, 2008). However, the major limitation of this method is that it requires several compositions to be produced near one another to generate an accurate and comprehensive measurement of a student's writing ability over time.

There have been some Rasch models of essay ratings. Andrich (1973) introduced a Poisson modeling approach that measures the errors in writing products. In other studies, De Gruijter (1984) presented simple models for raters' effects: the additive model and a nonlinear model. The nonlinear model is based on Choppin's pair Rasch model (Chopin, 1982). A recent Rasch model extension was proposed by (Linacre, 1989) which provided multiple facets. The FACETS model is an extension of Rasch measurement models (Wright & Masters, 1982; Wright & Stone, 1979) that can be used in writing assessments.

The attributes used in the Writing Ability Developmental Scale were selected based on crucial writing qualities that serve as indicators of various contributions to the organization of writing. This approach was influenced by Kellogg's (2008) hypothesis that working memory plays a crucial role in the development of writing. For example, syntactic complexity in writing is often seen as a measure of a writer's ability to retain language structures in working memory (Chenoweth & Hayes, 2003; Kellogg, 2008). This is expressed as a percentage of the total words in a composition, as well as the ratio of different parts of speech. Several lexical aspects of writing are proxies for a writer's ability to retrieve and retain language from both long-term memory and short-term/working memory (Chenoweth & Hayes, 2003; Kellogg, 2008). Compositional lexical properties were measured in three ways: lexical density, lexical diversity, and lexical sophistication (Burdick et al., 2013). Lexical density is calculated based on the number of words in the composition. Lexical diversity is represented by the number of general word types. Lexical sophistication is defined as the average length and frequency of words and is a more sophisticated measure of writing ability. The length of a composition is not just a pointless indicator of writing ability, instead, it measures fluency. Shorter, low-scoring answers often lack important features that contribute to both writing quality and document length, such as the development of supporting points. Hence, the relationship between document length and human scores reflects the writer's ability to effectively organize and regulate the writing process (Klobucar et al., 2012).

The Lexile framework is used in Korea to assess English-Lectio Quotient (E-LQ) when reporting a reader's Lexile measure (Fitzgerald et al., 2015). The Lexile framework is becoming increasingly

popular in America and it is widely used in schools (Copeland et al., 2013). This measure represents student writing ability on a similar scale as reading proficiency, providing teachers with an easy method for measuring and tracking student writing growth over time.

In this study, research analysis with a Developmental Scale was used to assess essay writing performance and examine the relationship between language variables and writing ability linguistically. This study assessed the overall writing ability when used occasionally, rather than only at the time of composition.

METHODOLOGY

2.1. Research Design and Research Questions

Participants were required to write six essays twice a week during each of the three weeks of December. The computer analysis was performed on each of the six configurations of each student, and the Bayesian algorithm was applied to the data of students' scores.

The following research questions were addressed: (a) Is there any relationship between variables in writing proficiency: namely content, fluency, sentence formation, usage, and mechanics in students of 1st – 4th grades? (b) Is there any relationship between variables in linguistic language: cognitive, compositions, vocabulary, style, and socio-cultural in students in 8th grade students? (c) Does writing proficiency as measured by indicators of content, fluency, sentence formation, usage, and mechanics in students in grades 1-4 have a relationship with linguistic language as measured by indicators of cognitive, compositions, vocabulary, style, and sociocultural in students in grade 8?

2.2. Participants

A total of 100 EFL students selected using a convenience sampling technique from one school in Malang, Indonesia, took part in this study. The students were in grades 1-6 (n = 48) and 7-8 (n= 52). Fifty percent of the students are female, while 40% are male. The data were collected by 17 teachers.

2.3. Prompts

Researchers adopted a standardized setting for students to write in the study, such as students limiting themes. There were 18 indicators used in this study published on the NAEP website and made available to the authors. They were grouped into three grade levels, each having two story signs (N), informative (I), or persuasive (P).

The National Assessment of Educational Progress (NAEP) texts are grouped as follows. (a) Narrative Writing: This involves composing personal stories and essays that are rich in creativity and imagination. The theme of the writing evaluation challenges writers to use their unique perspectives to engage their audience. (b) Informative Writing: This type of writing focuses on conveying information to the reader, to share knowledge, ideas, and messages. Students are tested on their ability to write about a specific topic in various formats, such as reports, reviews, letters, and others. (c) Persuasive Writing: The primary purpose of persuasive writing is to convince the reader to take a specific action, through the use of reasons, examples, and comparisons. In a persuasive writing assessment, students must not only write to friends, newspaper editors, or potential employers but also be able to participate in discussions and refute opposing viewpoints.

2.4. Prompt administration

One of the authors of the study trained the instructor to give instructions during a 20-minute group session. Before rapid management, packages were adjusted to each class's requirements and gender type. For each writing session, the instructor used the same standardized set of instructions as specified in the NAEP instructions.

2.5. Measures

The attitudes towards the scale, individual differences, and the problem-solving process were assessed using various parameters. Essays were evaluated based on content and organization, style, sentence formation, usage, and mechanics. A four-point rating scale was used to classify the writing from "poor" (1) to "very good" (4). The evaluation criteria included: analytical, evaluative, or creative thinking; organization and coherence, including effective transitions; control of lexical level and word structure; vocabulary range and sophistication; level of detail and elaboration; use of relevant analogies, illustrations, anecdotes, or examples; and mastery of grammar, spelling, and punctuation. The average scores of the six pieces of writing were calculated and classified into 1 (low) to 6 (high). The combination of these scores comprehensively reflects students' writing skills and their general ability to write effectively.

2.6. Data analysis

A descriptive analysis was first performed to calculate the central tendencies (mean, mode, median, etc.) and measure the data distribution (standard deviation, variance, etc.). The results of the

descriptive analysis are presented in Table 1. After that, Pearson's correlation analysis was carried out to determine the relationship between the indicators in grades 1-4 and grade 8. To further investigate the impact of writing proficiency in grades 1-4 on cognitive abilities, composition skills, vocabulary, style, and socio-cultural factors at grade 8, multiple linear regression analysis was performed. In multiple linear regression analysis, the relationship between two or more independent variables (x) (content, fluency, sentence formation, usage, and mechanics) and a dependent variable (y) (cognitive ability, composition skill, vocabulary, style, and socio-cultural factors) was examined. Before conducting the multiple linear regression analysis, a normality test was performed to ensure that the data met the requirements of the analysis. A good multiple linear regression model should be free from multicollinearity, heteroscedasticity, and autocorrelation issues and should also meet the normality assumption. The assumption test was also performed using multiple linear regression analysis.

RESULT AND DISCUSSION

Result

3.1. Descriptive Statistics

Table 1. Descriptive statistics of research variables

No	Variable	N	Mean	StDev	Minimum	Maximum
1	Cognitive	100	3.117	1.314	1.100	6.000
2	Compositions	100	3.040	1.249	1.100	6.000
3	Vocabulary	100	3.134	1.306	1.100	6.000
4	Style	100	3.404	1.256	1.100	6.000
5	Socio-cultural	100	3.501	1.346	1.100	6.000
6	Content	100	2.6690	0.7935	1.1000	4.0000
7	Fluency	100	2.6450	0.7924	1.1000	4.0000
8	Sentence formation	100	2.7130	0.8830	1.0000	6.0000
9	Usage	100	2.5460	0.8683	1.1000	5,3000
10	Mechanics	100	2.6540	0.7804	1.1000	4.0000

3.2. Pearson’s correlation

Correlation is a frequently used statistical analysis in identifying the relationship between variables in a study. Correlation only explains the strength of the relationship and does not measure the causal relationship between the variables. The results of the Pearson's correlation test performed in this study are presented in Table 2.

Table 2. Correlation between X variables

		Content	Fluency	Sentence formation	Usage	Mechanics
Content	Pearson Correlation	1	.098	.012	-.103	-.147
	Sig. (2-tailed)		.331	.908	.307	.143
Fluency	Pearson Correlation	.098	1	.001	-.094	.071
	Sig. (2-tailed)	.331		.993	.353	.484
Sentence formation	Pearson Correlation	.012	.001	1	-.082	.116
	Sig. (2-tailed)	.908	.993		.419	.252
Usage	Pearson Correlation	-.103	-.094	-.082	1	.075
	Sig. (2-tailed)	.307	.353	.419		.459
Mechanics	Pearson Correlation	-.147	.071	.116	.075	1
	Sig. (2-tailed)	.143	.484	.252	.459	

a. Listwise N=100

The correlations are shown by the significance value of each, in which each correlation pair shows a significance value or p-value > 0.05 which indicates that the two variables do not have a linear relationship.

Table 3. The correlation between Y variables

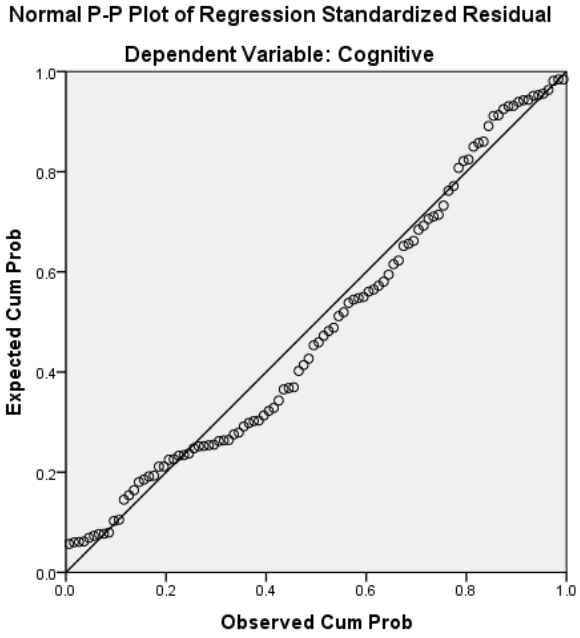
		Cognitive	Compositions	Vocabulary	Style	Sociocultural
Cognitive	Pearson Correlation	1	.057	-.046	.158	-.029
	Sig. (2-tailed)		.576	.652	.115	.772
Compositions	Pearson Correlation	.057	1	-.178	.209	-.004
	Sig. (2-tailed)	.576		.076	.037	.971
Vocabulary	Pearson Correlation	-.046	-.178	1	.035	-.017
	Sig. (2-tailed)	.652	.076		.729	.865
Style	Pearson Correlation	.158	.209	.035	1	.022
	Sig. (2-tailed)	.115	.037	.729		.827
Sociocultural	Pearson Correlation	-.029	-.004	-.017	.022	1
	Sig. (2-tailed)	.772	.971	.865	.827	

a. Listwise N=100

Table 3 shows the p-value for each pair of variables greater than 0.05, indicating that there is no significant linear relationship between the variables except for the p-value of the correlation between composition and style, as their p-value is lesser than 0.05. Hence, a significant linear relationship between the two variables was found, suggesting that changes in the composition and style are likely to affect each other.

3.3. Model 1: The Influences of Content, Fluency, Sentence Formation, Usage and Mechanics on Cognitive.

The normality assumption test states that the regression is considered to be normally distributed if the graphing data that illustrates the real statistics follows a diagonal line.



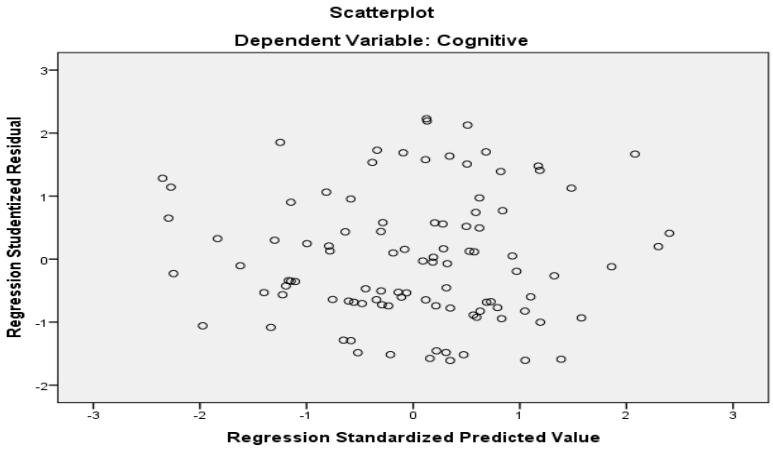
Multicollinearity test: There are no signs of multicollinearity if the tolerance value > 0.1 and VIF < 10.00.

Table 4. Coefficients of model 1

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.316	.986		3.364	.001		
	Content	-.068	.172	-.041	-.397	.692	.959	1.043
	Fluency	-.083	.170	-.050	-.488	.627	.975	1.026
	Sentence formation	.207	.153	.139	1.355	.179	.978	1.023
	Usage	.051	.156	.034	.326	.745	.970	1.031
	Mechanics	-.184	.175	-.109	-1.048	.297	.951	1.051

a. Dependent variable: cognitive

Heteroscedasticity test: There is no heteroscedasticity if the scatter plot image has no discernible pattern (curly, widening, and then narrowing) and the dots are scattered near the outer number 0 on the Y axis.



Autocorrelation test: There is no autocorrelation symptom if the Durbin-Watson value lies between du to (4-du).

Durbin Watson (1.776) <(4 - 1.7804 = 2.2196) and Durbin Watson (1.776)> DL (1.5710) It means there is no autocorrelation.

Table 5. Model Summary of Model 1

Model	R	R Square	Adjusted Square	RStd. Error of the Estimate	Durbin-Watson
1	.179[a]	.032	-.019	1.32673	1.776
a.	Predictors: (constant), mechanics, fluency, sentence formation, usage, content.				
b.	Dependent variable: cognitive.				

3.3.1. Multiple Linear Regression Analysis

The partial and simultaneous effect of variable X on Y was also examined.

a. Partial Regression Coefficient Test

The regression coefficient shows if the independent variable (X) has a meaningful influence on the dependent variable (Y).

As presented in Table 4, the partial regression coefficient is explained as follows.

1. Content does not partially affect students' cognitive scores.
2. Fluency does not partially affect students' cognitive scores.

3. Sentence formation does not partially affect students' cognitive scores.
4. Usage does not partially affect the cognitive rate of students.

b. Simultaneous test

Table 6. ANOVA of model 1

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.481	5	1.096	.623	.683[b]
	Residual	165.460	94	1.760		
	Total	170.941	99			
a.	Dependent variable: cognitive					
c.	Predictors: (constant), mechanics, fluency, sentence formation, usage, content.					

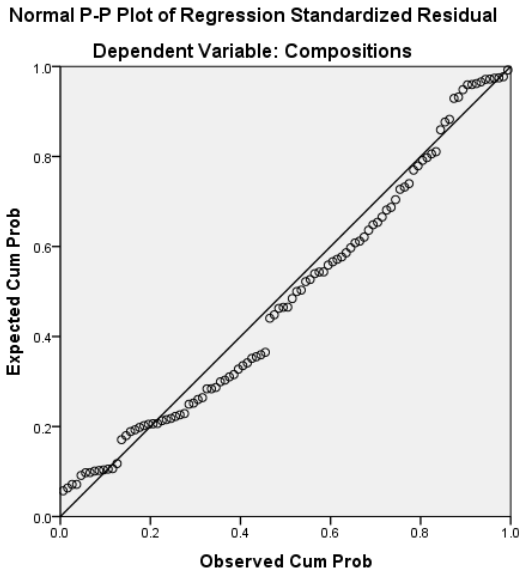
Content (X1), fluency (X2), sentence formation (X3), usage (X4), and mechanics (X5) do not simultaneously affect the cognitive value of students.

c. Coefficient of Determination

The coefficient of determination (R square) represents the proportion of the variation in the dependent variable that is explained by the independent variables. R^2 value smaller than 3.2%, suggests that the independent variables have a weak effect on the dependent variable. Hence, there are other factors influencing the dependent variable, or the independent variables themselves are not strong predictors.

3.4. Model 2: The Influences of Content, Fluency, Sentence Formation, Usage, and Mechanics on Compositions Value

Normality assumption test: The regression is considered to be normally distributed if the graph shows a diagonal line.



Multicollinearity test: There are no signs of multicollinearity if the tolerance value > 0.1 and VIF < 10.00.

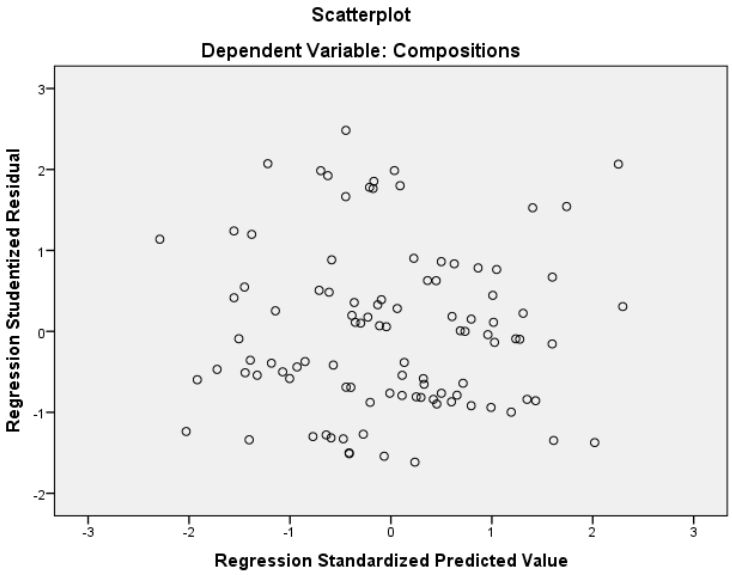
Table 7. Coefficients of model 2

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.025	.937		2.160	.033		
	Content	.022	.163	.014	.133	.894	.959	1.043

Fluency	.067	.162	.043	.413	.680	.975	1.026
Sentence formation	-.027	.145	-.019	-.185	.854	.978	1.023
Usage	.069	.148	.048	.467	.642	.970	1.031
Mechanics	.255	.167	.159	1.529	.130	.951	1.051

a. Dependent variable: compositions

Heteroscedasticity test: There is no heteroscedasticity if the scatter plot image has no discernible pattern (curly, widening, and then narrowing) and the dots are scattered near the outer number 0 on the Y axis.



Autocorrelation test: There is no autocorrelation symptom if the Durbin value Watson is located between du to (4-du).

Durbin Watson (2.034) <(4- 1.7804= 2.2196) and Durbin Watson (2.034)> DL (1.5710) There is no autocorrelation.

Table 8. Model Summary Model 2

Model	R	R Square	Adjusted Square	RStd. Error of the Estimate	Durbin-Watson
1	.175[a]	.031	-.021	1.26199	2.034

a. Predictors: (constant), mechanics, fluency, sentence formation, usage, content.
b. Dependent variable: compositions.

3.4.1 Multiple Linear Regression Analysis:

a. Partial Regression Coefficient Test

Table 7 can be interpreted as follows.

1. Content has no partial influence on students' composition scores.
2. Fluency has no partial influence on students' composition scores.
3. Sentence formation has no partial influence on students' composition scores.
4. Usage has no partial influence on students' composition scores.
5. Mechanics has no partial influence on students' grade compositions.

b. Simultaneous test

Table 9. ANOVA of model 2

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.714	5	.943	.592	.706[b]
	Residual	149.706	94	1.593		
	Total	154.420	99			

a. Dependent variable: compositions
b. Predictors: (constant), mechanics, fluency, sentence formation, usage, content.

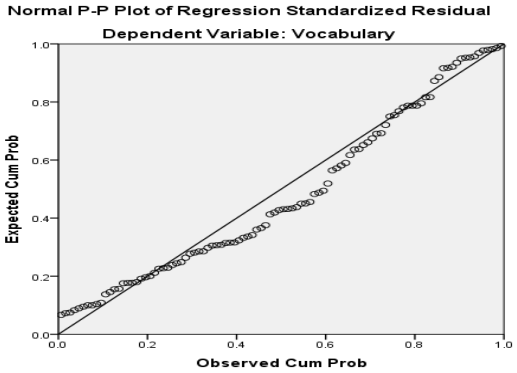
Content (X1), fluency (X2), sentence formation (X3), usage (X4), and mechanics (X5) of students in grades 1-4 do not have simultaneous influences on the composition scores of students in grades 8.

c. Coefficient of Determination

The coefficient of determination (R square) shows the influence of the independent variable (X) on the dependent variable (Y). In other words, R square can be used to predict the contribution of variable X on variable Y. Low R^2 of 3.1% found in this study indicates a weaker influence of variable X on Y is weak.

3.5. Model 3: The Influences of Content, Fluency, Sentence Formation, Usage and Mechanics on Vocabulary Mastery

Normality assumption test: The regression is normally distributed if the dots follow the diagonal line.



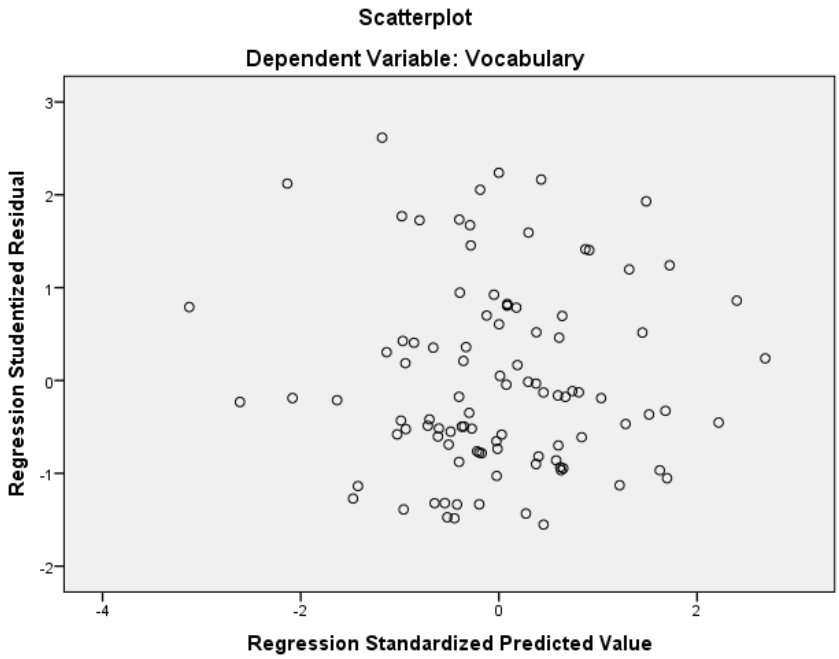
Multicollinearity test: There is no symptom of multicollinearity if the tolerance value > 0.1 and VIF < 10.00.

Table 10. Coefficients of Model 3

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.750	.972		3.859	.000		
	Content	-.074	.169	-.045	-.437	.663	.959	1.043
	Fluency	.115	.168	.070	.682	.497	.975	1.026
	Sentence formation	.139	.151	.094	.920	.360	.978	1.023
	Usage	-.158	.154	-.105	-1.029	.306	.970	1.031
	Mechanics	-.262	.173	-.156	-1.515	.133	.951	1.051

a. Dependent variable: Vocabulary

Heteroscedasticity test: There is no heteroscedasticity if the scatter plot image has no discernible pattern (curly, widening, and then narrowing) and the dots are scattered near the outer number 0 on the Y axis.



Autocorrelation test: There is no autocorrelation symptom if the Durbin-Watson value lies between du to (4-du).

DU Durbin Watson (1.502) <(4- 1.7804= 2.2196). There is no autocorrelation.

Table 11. Model summary of model 3

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.219[a]	.048	-.003	1.30813	1.502

a. Predictors: (constant), mechanics, fluency, sentence formation, usage, content.

b. Dependent variable: vocabulary.

3.5.1. Multiple Linear Regression Analysis:

a. Partial Regression Coefficient Test

The partial regression coefficient is explained as follows.

1. Content has no partial influence on students' vocabulary scores.
2. Fluency has no partial influence on students' vocabulary scores.
3. Sentence formation has no partial influence on students' vocabulary scores.
4. Usage has no partial influence on students' vocabulary scores.
5. Mechanics has no partial influence on students' vocabulary scores

b. Simultaneous Test

Table 12. ANOVA of model 3

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8.090	5	1.618	.946	.455[b]
	Residual	160.854	94	1.711		
	Total	168.944	99			

a. Dependent variable: vocabulary.

b. Predictors: (constant), mechanics, fluency, sentence formation, usage, content.

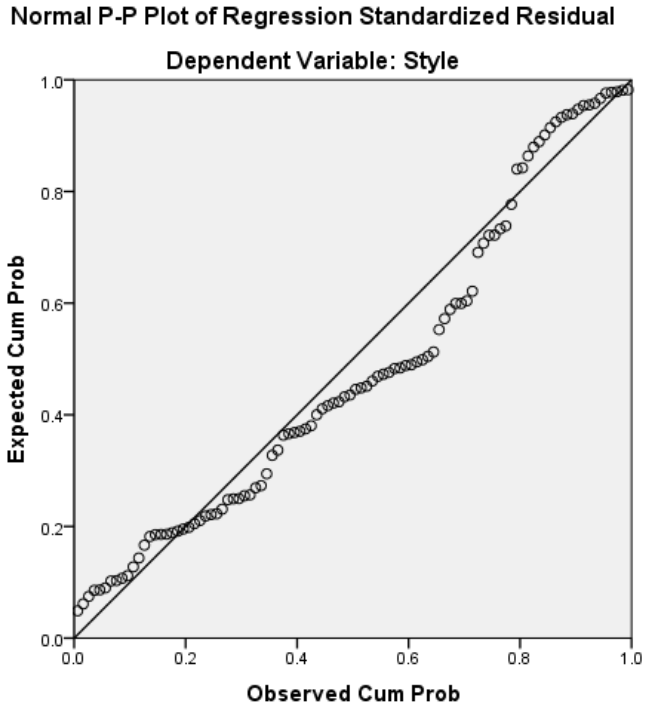
Content (X1), fluency (X2), sentence formation (X3), usage (X4), and mechanics (X5) of students in grades 1-4 do not have simultaneous influences on the vocabulary of students in grades 8.

c. Coefficient of Determination

The coefficient of determination (R square) shows the influence of the independent variable (X) on the dependent variable (Y). In other words, R square can be used to predict the contribution of variable X on variable Y. A Low R^2 of 4.8% found in this study indicates a weaker influence of variable X on Y is weak.

3.6. Model 4: Relationship of Content, Fluency, Sentence Formation, Usage and Mechanics to Style Value

Normality assumption test: The regression is normally distributed if the dots follow the diagonal line.



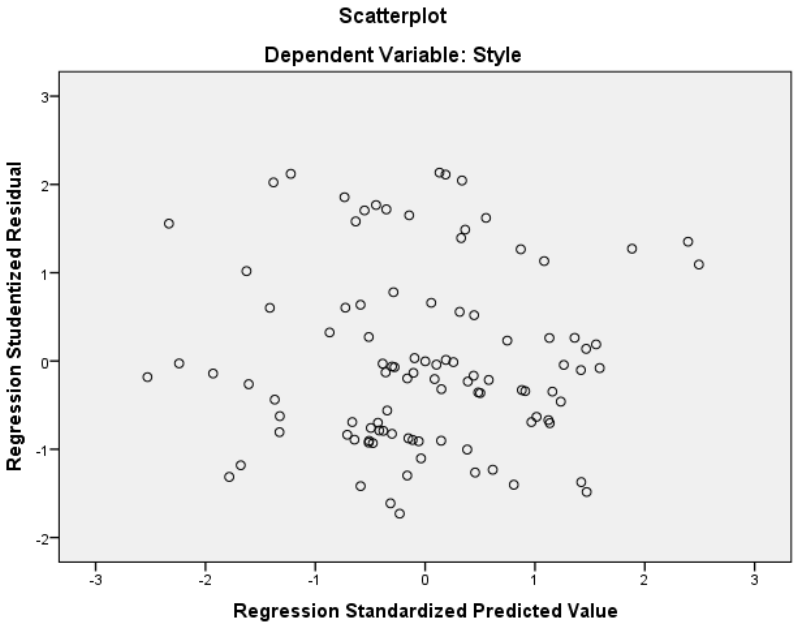
Multicollinearity test: There is no sign of multicollinearity if the tolerance value > 0.1 and VIF < 10.00.

Table 13. Coefficients of Model 4

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.053	.898		2.285	.025		
	Content	.034	.156	.021	.216	.829	.959	1.043
	Fluency	-.055	.155	-.035	-.352	.726	.975	1.026
	Sentence formation	.401	.139	.282	2.878	.005	.978	1.023
	Usage	.308	.142	.213	2.170	.033	.970	1.031
	Mechanics	-.175	.160	-.109	-1.099	.275	.951	1.051

a. Dependent variable: style

Heteroscedasticity Test: There is no heteroscedasticity if the scatter plot image has no discernible pattern (curly, widening, and then narrowing) and the dots are scattered near the outer number 0 on the Y axis.



Autocorrelation test: There is no autocorrelation symptom if the Durbin-Watson value lies between du to (4-du).

DU Durbin Watson (1.825) <(4- 1.7804= 2.2196). There is no autocorrelation.

Table 14. Model summary of model 4

Model	R	R Square	Adjusted Square	RStd. Error of the Estimate	Durbin-Watson
1	.346[a]	.120	.073	1.20908	1.825

a. Predictors: (constant), mechanics, fluency, sentence formation, usage, content.
b. Dependent variable: style.

3.6.1. Multiple Linear Regression Analysis:

a. Partial Regression Coefficient Test

Based on the data in Table 13. partial regression coefficient was as follows:

1. Content has no partial influence on students' style values.
2. Fluency has no partial influence on the student's style value.
3. Sentence formation has a partial influence on the student's style value.
4. Usage has a partial influence on student's style value.
5. Mechanics has no partial influence on students' style scores.

b. Simultaneous Test

Table 15. ANOVA model 4

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18.663	5	3.733	2.553	.033[b]
	Residual	137.415	94	1.462		
	Total	156.078	99			

a. Dependent variable: style
b. Predictors: (constant), mechanics, fluency, sentence formation, usage, content

Content (X1), fluency (X2), sentence formation (X3), usage (X4), and mechanics (X5) of students in grades 1-4 do not have simultaneous influences on the writing style of students in grades 8.

Regression equation:

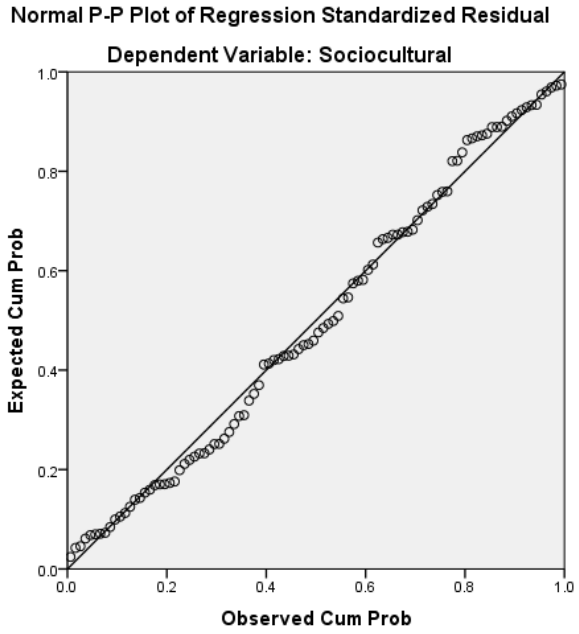
$$Y = 3.316 - 0.068\,X_1 - 0.083X_2 + 0.207X_3 + 0.051X_4 - 0.184X_5$$

c. Coefficient of Determination

The coefficient of determination (R square) shows the influence of the independent variable (X) on the dependent variable (Y). In other words, R square can be used to predict the contribution of variable X on variable Y. A Low R^2 of 1.2% found in this study indicates a weaker influence of variable X on Y is weak.

3.7. Model 5: Relationship of Content, Fluency, Sentence Formation, Usage and Mechanics to Socio-cultural Values

Normality assumption test: The regression is normally distributed if dots follow the diagonal line.



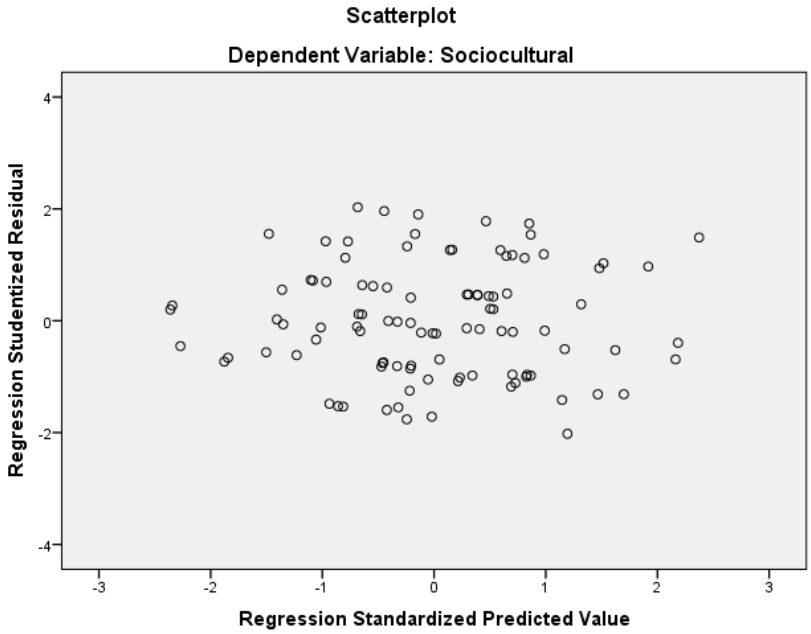
Multicollinearity test: There is no sign of multicollinearity if the tolerance value > 0.1 and VIF < 10.00.

Table 16. Coefficients of model 5

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	4.746	1.010		4.698	.000		
	Content	-.136	.176	-.080	-.775	.440	.959	1.043
	Fluency	-.009	.175	-.005	-.052	.958	.975	1.026
	Sentence formation	-.116	.157	-.076	-.740	.461	.978	1.023
	Usage	.030	.160	.019	.188	.851	.970	1.031
	Mechanics	-.233	.180	-.135	-1.300	.197	.951	1.051

a. Dependent Variable: Sociocultural

Heteroscedasticity Test: There is no heteroscedasticity if the scatter plot image has no discernible pattern (curly, widening, and then narrowing) and the dots are scattered near the outer number 0 on the Y axis.



Autocorrelation test: There is no autocorrelation symptom if the Durbin Watson value lies between due to (4-du).

DU Durbin Watson (1.699) < (4- 1.7804= 2.2196). There is no autocorrelation.

Table 17. Model summary of model 5

Model	R	R Square	Adjusted Square	RStd. Error of the Estimate	Durbin-Watson
1	.175[a]	.031	-.021	1.35989	1.669

a. Predictors: (constant), mechanics, fluency, sentence formation, usage, content.

b. Dependent variable: sociocultural.

3.7.1. Multiple Linear Regression Analysis:

a. Partial Regression Coefficient Test

Table 16 shows the partial regression coefficients that are explained as follows.

1. Content has no partial influence on students' socio-cultural scores.
2. Fluency has no partial influence on students' socio-cultural scores.
3. Sentence formation has no partial influence on students' socio-cultural scores.
4. Usage has no partial influence on students' socio-cultural scores.
5. Mechanics has no partial influence on students' socio-cultural scores.

b. Simultaneous Test

Table 18. ANOVA of model 5

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.497	5	1.099	.594	.704[b]
	Residual	173.833	94	1.849		
	Total	179.330	99			

a. Dependent variable: sociocultural.

b. Predictors: (constant), mechanics, fluency, sentence formation, usage, content.

Content (X1), fluency (X2), sentence formation (X3), usage (X4), and mechanics (X5) of students in grades 1-4 do not have simultaneous influences on the socio-cultural scores of students in grades 8.

c. Coefficient of Determination

The coefficient of determination (R square) shows the influence of the independent variable (X) on the dependent variable (Y). In other words, R square can be used to predict the contribution of variable X on variable Y. Low R^2 of 3.1% found in this study indicates a weaker influence of variable X on Y is weak.

Discussion

This study combined the trait-oriented and situational approaches in writing assessment as proposed by Burdick et al. (2013) and Kellogg (2008). Evaluation of writing ability based on personal characters and the quality of the written text as a characteristic of the product was performed to gain a comprehensive understanding of the student's writing skills. The Lexile scale was employed to evaluate written text and provided a quantitative measure of the quality of writing, while the assessment of students' cognitive processes during the writing task was carried out to understand the thinking processes and strategies used by the students in writing. This study provides a well-rounded assessment of the student's writing abilities.

A strong linear relationship was found between the composition and style variables, and vice versa, between students in grades 1-4 and grade 8, implying the presence of influence of linguistic language on writing abilities across grades. The composition test consisted of eight items, two word-choice items, and one style-related item (Silverman et al., 2015). The significance of word choice in composition assessment highlights the correlation between vocabulary breadth and writing performance. Duin & Graves, (1986) found that oral vocabulary instruction positively impacted written vocabulary, leading to improved writing performance. However, further investigation should be performed to gain more comprehensive findings related to the influence of vocabulary on writing ability.

The results of this study indicated no correlation between the content, fluency, sentence formation, usage, and mechanics scores of students in grades 1-4 and their cognitive, composition, vocabulary, and sociocultural scores in grade 8. However, a strong relationship was found between the aforementioned scores in grades 1-4 and the students' style scores in grade 8, which significantly impacted their assessment of writing style in grade 8. This finding aligns with the conclusion of Goodman (2014) who stated that essays can reflect one's writing ability and language proficiency.

This study initiates the exploration of EFL writing abilities over time using a developmental scale. Further studies can be conducted to gain more comprehensive findings by incorporating measures of second-language proficiency and examining the relationship between EFL writing instruction and performance. It is also necessary to examine whether different approaches to EFL teaching result in unique patterns of writing development. Therefore, the underlying factors that contribute to the growth of EFL writing abilities can be identified.

CONCLUSION

This study sheds light on an important yet under-researched topic. Some researchers have examined the relationship between writing ability and language skills (Berninger & Abbott, 2010; Kim et al., 2011; Olinghouse & Leaird, 2009), while the studies that specifically examined the relationship in the context of EFL writing are limited. The increasing number of EFL students in academic settings and the centrality of writing in education show that this topic needs to be further explored. The results of this study suggest that language skills positively affect EFL students' writing ability and support the idea that language and writing skills are interconnected components of a larger system ((Berninger & Abbott, 2010). More studies should be carried out to fully understand the relationship between language abilities and writing performance using a variety of measures and methods It is also worth researching the progression of EFL students' language abilities and its impact on writing performance.

REFERENCES

Allinson, C. W., & Hayes, J. (1996). The Cognitive Style Index: A Measure of Intuition-Analysis For Organizational Research. *Journal of Management Studies*, *33*(1), 119–135. https://doi.org/10.1111/J.1467-6486.1996.TB00801.X

Andrich, D. (1973). Latent Trait Psychometric Theory in The Measurement and Evaluation of Essay Writing Ability. - ProQuest. *Unpublished Doctoral Dissertation, The University of Chicago.* https://www.proquest.com/openview/07d5ec16ba34c06feb9099bb8d47b92b/1.pdf?pq-

origsite=gscholar&cbl=18750&diss=y

Applebee, A. N. (1986). *Writing: Trends Across the Decade, 1974-84.* National Assessment of Educational Progress, Educational Testing Service, Rosedale Road, Princeton, NJ 08541 ($12.50 for single copies, half price for more than 3 copies).

Attali, Y., & Burstein, J. (2006). *With e-rater V . 2*. *4*(3).

Bachman, L. F. (1990). *Fundamental considerations in language testing*. 408.

Beck, S. W., & Jeffery, J. V. (2007). Genres of high-stakes writing assessments and the construct of writing competence. *Assessing Writing*, *12*(1), 60–79. https://doi.org/10.1016/j.asw.2007.05.001

Bereiter, C., & Scardamalia, M. (1987). An Attainable Version of High Literacy: Approaches to Teaching Higher-Order Skills in Reading and Writing. *Curriculum Inquiry*, *17*(1), 9–30. https://doi.org/10.1080/03626784.1987.11075275

Berninger, V. W., & Abbott, R. D. (2010). Listening Comprehension, Oral Expression, Reading Comprehension, and Written Expression: Related Yet Unique Language Systems in Grades 1. 3. 5, and 7. *Journal of Educational Psychology*, *102*(3), 635–651. https://doi.org/10.1037/a0019319

Burdick, H., Swartz, C. W., Stenner, A. J., Fitzgerald, J., Burdick, D., & Hanlon, S. T. (2013). Measuring Students' Writing Ability on a Computer-Analytic Developmental Scale: An Exploratory Validity Study. *Literacy Research and Instruction*, *52*(4), 255–280. https://doi.org/10.1080/19388071.2013.812165

Chenoweth, N. A., & Hayes, J. R. (2003). The Inner Voice in Writing. *Written Communication*, *20*(1), 99–118. https://doi.org/10.1177/0741088303253572

Chopin, B. (1982). A Fully Conditional Estimation Procedure for Rasch Model Parameters. *Center for The Study of Evaluation. Graduate School of Education University of California, Los Angeles*, *Cm*.

Cooper, C. R. (1977). Holistic evaluation of writing. In C. R. Cooper & L. Odell (Eds.). *Evaluating Writing: Describing, Measuring, Judging. Buffalo: State University of New York.*, *67*(9), 71. https://doi.org/10.2307/815139

Copeland, B. M., Specialist, L. C., Liben, D., & Specialist, S. C. (2013). *Quick and Easy Access to Measures of Text Complexity Using the Lexile Framework for Reading*.

De Gruijter, D. N. M. (1984). *Two Simple Models for Rater Effects*. *8*(2), 213–218. http://www.copyright.com/

Duin, A. H., & Graves, M. F. (1986). Intensive Vocabulary Instruction as a Prewriting Technique. *Reading Research Quarterly*, *22*(3), 311. https://doi.org/10.2307/747971

Engelhard, G. (2009). Applied Measurement in Education The Measurement of Writing Ability With a Many-Faceted Rasch Model. *Applied Measurement in Education*, *3*(January 2012), 37–41. https://doi.org/10.1207/s15324818ame0503

Engelhard, George, Gordan, B., & Gabrielson, S. (1991). Writing Tasks 1 Writing Tasks and The Quality of Student Writing: Evidence Mlim a Statewide Assess/Wit of Miting Rest Copy Available. *Annual Meeting of the American Educational Research Association*.

Fitzgerald, J., Stenner, A. J., Sanford-Moore, E. E., Koons, H., Bowen, K., & Kim, K. H. (2015). The Relationship of Korean Students' Age and Years of English-as-a-Foreign-Language Exposure with English-Reading Ability: A Cross-Age Study. *Reading Psychology*, *36*(2), 173–202. https://doi.org/10.1080/02702711.2013.843063

Flower, L., & Hayes, J. R. (1980). *The Cognition of Discovery: Defining a Rhetorical Problem*. *31*(1), 21–32.

Goodman, K. S. (2014). Reading: A psycholinguistic guessing game. *Making Sense of Learners Making Sense of Written Language: The Selected Works of Kenneth S. Goodman and Yetta M. Goodman*, 103–112. https://doi.org/10.4324/9780203366929

Gorman, T. P., Purves, A. C., & Degenhart, R. E. (1988). The IEA study for written composition I : the international writing tasks and scoring scales. *Oxford. 1988. 1st Ed.. 214 P..*

Graham, S., Harris, K., & Hebert, M. (2011). Informing Writing. *The Benefits of Formative Assessment.*, 40. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Informing+Writing#3

Jahin, J. H. (2012). The effect of peer reviewing on writing apprehension and essay writing ability of prospective EFL teachers. *Australian Journal of Teacher Education*, *37*(11), 60–84. https://doi.org/10.14221/ajte.2012v37n11.3

Kellogg, R. (2008). *Training writing skills: A cognitive developmental perspective Related papers The Effect of Psycholinguist ic Research on the Teaching of Writing*.

Kim, Y. S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences*, *21*(5), 517–525. https://doi.org/10.1016/j.lindif.2011.06.004

Klobucar, A., Deane, P., Elliot, N., Ramineni, C., Deess, P., & Rudniy, A. (2012). Automated essay scoring and the search for valid writing assessment. *International Advances in Writing Research: Cultures, Places, Measures, 103-119*, 5–24.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2–3), 259–284. https://doi.org/10.1080/01638539809545028

Linacre, J. M. (1989). *Simple and Effective Algorithms: Computer-Adaptive Testing*.

Merjen, A., Yogyakarta, U. N., & Kassymova, G. K. (2019). *Impact of reading on students ' writing ability*. *August*. https://doi.org/10.31643/2019.001

Nunan, D. (1999). *Second language teaching & learning*. 330.

Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing*, *22*(5), 545–565. https://doi.org/10.1007/s11145-008-9124-z

Raimes, A. (1983). *Teaching technique in English as a second language*. 163.

Rea, P. W., Applebee, A. N., Langer, J. A., Jenkins, L. B., Millis, I. V. S., & Foertsch, M. A. (1991). Learning to Write in Our Nation's Schools: Instruction and Achievement in 1988 at Grades 4. 8, and 12. In *College Composition and Communication* (Vol. 42. Issue 1). https://doi.org/10.2307/357554

Silverman, R. D., Coker, D., Patrick Proctor, C., Harring, J., Piantedosi, K. W., & Hartranft, A. M. (2015). The relationship between language skills and writing outcomes for linguistically diverse students in upper elementary school. *Elementary School Journal*, *116*(1), 103–125. https://doi.org/10.1086/683135

Traxler, M. J., & Gernsbacher, M. A. (1993). Improving written communication through perspective-taking. *Language and Cognitive Processes*, *8*(3), 311–334. https://doi.org/10.1080/01690969308406958

Tribble, C. (2009). Writing academic English—a survey review of current published resources. *ELT Journal*, *63*(4), 400–417. https://doi.org/10.1093/ELT/CCP073

Wright, B. D., & Masters, G. N. (1982). Rating Scale Analysis: Rasch Measurement. In *Journal of the American Statistical Association* (Vol. 78, Issue 382). https://doi.org/10.2307/2288670

Wright, B. D., & Stone, M. H. (1979). Best Test Design. Rasch Measurement. *MESA Press*, 245.