CHAPTER FOUR : ANALYSIS OF DATA

This study seeks to explore the vocabulary sizes of urban and rural students and the possibility of it becoming a significant component in the teaching and learning of English in Malaysia. As there is currently no data on students' current vocabulary size, the first research question seeks to find out an estimate. This study would also explore the relationship between receptive and productive vocabulary of students, which would be addressed in the second research question. This is to see whether there is a pattern to which vocabulary size can be predicted. To extend vocabulary's significance further, the third research question seeks to find out to what extent vocabulary size correlates with students' ELA. The results of the analyses will be presented according to the research questions of the study.

4.1 The Receptive and Productive Vocabulary Size of Urban and Rural Students in the district of Kuantan (RQ1)

The first research question, which is *"What is the receptive and productive vocabulary size of urban and rural students in the district of Kuantan?"* is put forth to investigate the receptive and productive vocabulary sizes of urban and rural students. It will be calculated through the means of descriptive statistics (means and standard deviations). The overall mean scores will first be analysed to visualize the general pattern of the vocabulary sizes of the students.

The data analysis will be divided into two parts. In the first part, the mean scores are gathered from the translation tests, which test words based on the provided word list. An analysis of these mean scores will also be presented at different word levels. The

second part would be the analysis of the Lex30 test scores, which tests the production of low frequency words. Both the Lex30 scores (low frequency words, Level A+B) and Lex30 Level C scores (high frequency words) will be analysed. Lex30 Level C score will only be used for analysis here to provide a basis for comparison to the Lex30 score. It will not be analysed henceforth.

4.1.1 Receptive and Productive Vocabulary Size based on Translation Tests (Word List)

As the translation tests used target words from the word list provided by the Ministry of Education, the results reflect the vocabulary size based on the word list. The overall receptive and productive vocabulary sizes of urban and rural students will be analysed first, followed by a more detailed analysis at different word levels.

The descriptive statistics (means and standard deviations) of the overall receptive and productive translation test scores are reported in Table 4.1 and Figure 4.1.

. **Table 4.1**
Means and standard deviations for the overall receptive and productive scores of urban and rural students

	Mean		**Standard Deviation**	
	Urban	**Rural**	**Urban**	**Rural**
Receptive	222.68	174.52	44.584	35.257
Productive	224.60	175.78	36.095	37.510

Note. Maximum score = 300.

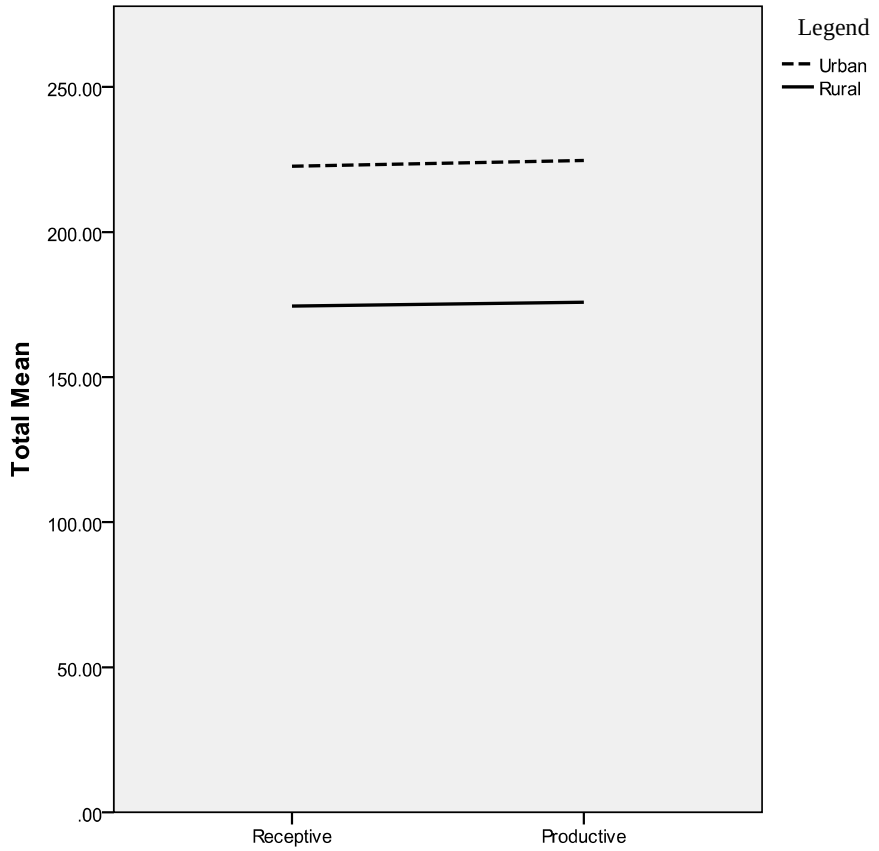

Figure 4.1
Mean for the overall receptive and productive scores of urban and rural students

The analysis revealed that the sizes of both receptive and productive vocabulary for urban students are larger than those of the rural students. This reflects past studies' results, where urban students outperformed rural students in terms of vocabulary (Rosli & Edwin, 1990, Lehr et.al., 2004, Graves et.al., 1982, cited in Berne & Blachowicz, 2008, Hart & Risley, 1995, cited in Lehr et.al., 2004) (see CHAPTER 2 Section 2.7.1). One-way repeated measures ANOVAs were conducted to prove its statistical significance. Urban students were found to have significantly larger receptive vocabulary size, Wilks' Lambda = .57, F (1, 49) = 36.34, $p < .0005$ and significantly larger productive vocabulary size, Wilks' Lambda = .50, F (1, 49) = 49.13, $p < .0005$, compared to rural students.

Another observation is that urban and rural students' productive vocabulary is slightly larger than their receptive vocabulary. However, the difference between their receptive and productive scores is marginal. The urban students' receptive vocabulary mean score was 222.68. It is only slightly less than the productive mean score, 224.60. A one-way repeated measures ANOVA was performed to determine whether there is any statistical significant difference between receptive and productive vocabulary sizes. The results showed no statistical significance between receptive and productive knowledge of urban students, Wilks' Lambda = .99, F (1, 49) = 0.325, $p > .0005$. For rural students, their receptive vocabulary mean score was 174.52. This is also slightly less than the productive mean score of 175.78. Statistical analysis also showed no significance in the difference of mean scores, Wilks' Lambda = .998, F (1, 49) = 0.112, $p > .0005$. Hence, there may not exist any difference in students' receptive and productive vocabulary. This is contrary to the popular belief that students generally have larger receptive vocabulary size, or that students at this age and level have a more pronounced vocabulary size gap (Laufer, 1998, cited in Meara and Fitzpatrick, 2000).

To study this general picture in more detail, descriptive statistic analysis was also employed to explore individual word level scores (Levels A, B and C) of the receptive and productive translation tests. Table 4.2 reports the data.

Table 4.2

Means and standard deviations for receptive and productive scores at different word levels

	Mean			Standard Deviation	
	Urban	**Rural**	**Ratio**	**Urban**	**Rural**
Receptive level A	60.57	36.46	60%	19.547	15.925
Receptive level B	72.79	58.76	81%	17.397	14.410
Receptive level C	89.33	79.30	89%	11.503	13.100
Productive level A	68.80	48.26	70%	16.183	15.721
Productive level B	74.60	57.84	78%	15.043	16.215
Productive level C	81.20	69.68	86%	8.894	11.493

Note : Maximum score = 100. Ratio = rural to urban mean scores.

The following line graphs illustrate the mean scores for the receptive (Figure 4.2) and productive (Figure 4.3) translation tests at each word level.

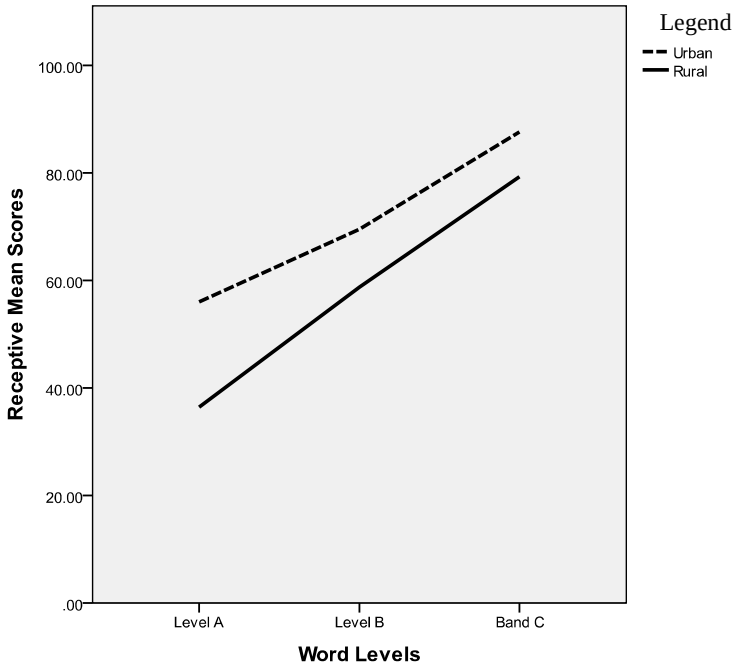

Figure 4.2

Mean scores for the receptive translation test at different word levels

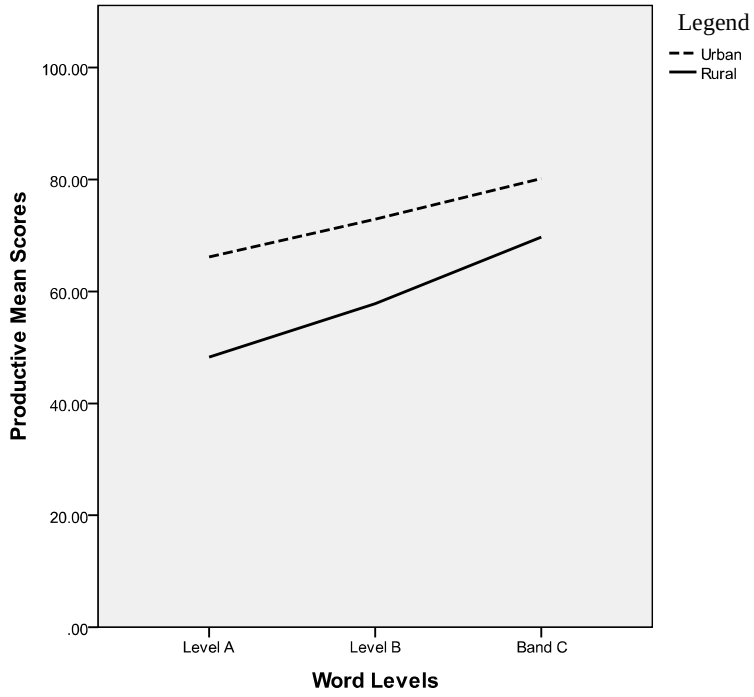

Figure 4.3
Mean scores for the productive translation test at different word levels

From the analysis, it can be concluded that the higher the word frequency, the higher the mean score. This is true for both rural and urban students, as well as for both the receptive and productive translation tests. Hence, both urban and rural students have larger receptive and productive high frequency vocabulary size compared to lower frequency vocabulary size.

Another observation which can be reported is the difference or gap, between urban students' scores and rural students' scores. The gap gets bigger when word frequency reduces. The gap between rural to urban students' scores for receptive vocabulary increased, from a ratio of 89% for Level C (high frequency words), to 81% for Level B (mid frequency words) and to 60% for Level A (low frequency words). The same pattern is recorded for productive vocabulary, where the gap increases from a ratio of 86% (Level

C), to 78% (Level B) and to 70% (Level C). Hence, this shows that urban and rural students' vocabulary gap between high frequency words is smaller, compared to their gap between low frequency words. Urban students have only slightly more high frequency words than rural students, but distinctly more low frequency words.

4.1.2 Lex30

To look into the productive capabilities of students without the limitation of the word list, the scores for Lex30 (Level A+B) and Lex30 Level C was analysed using descriptive statistics. Lex30 Level C score is included in this analysis as a basis for comparison between high frequency words (Level C) and low frequency words (Lex30). Henceforth, it will not be included in further analysis. The results are shown in Table 4.3 and Figure 4.4.

Table 4.3
Means and standard deviations for Lex30 score and Lex30 Level C score

Lex30	Mean			Standard Deviation	
	Urban	Rural	Ratio	Urban	Rural
Lex30	42.24	42.03	99.5%	9.279	12.326
Level C	57.29	58.71	102.5%	9.139	10.931

Note : Maximum score = 100. Ratio = Rural to urban scores.

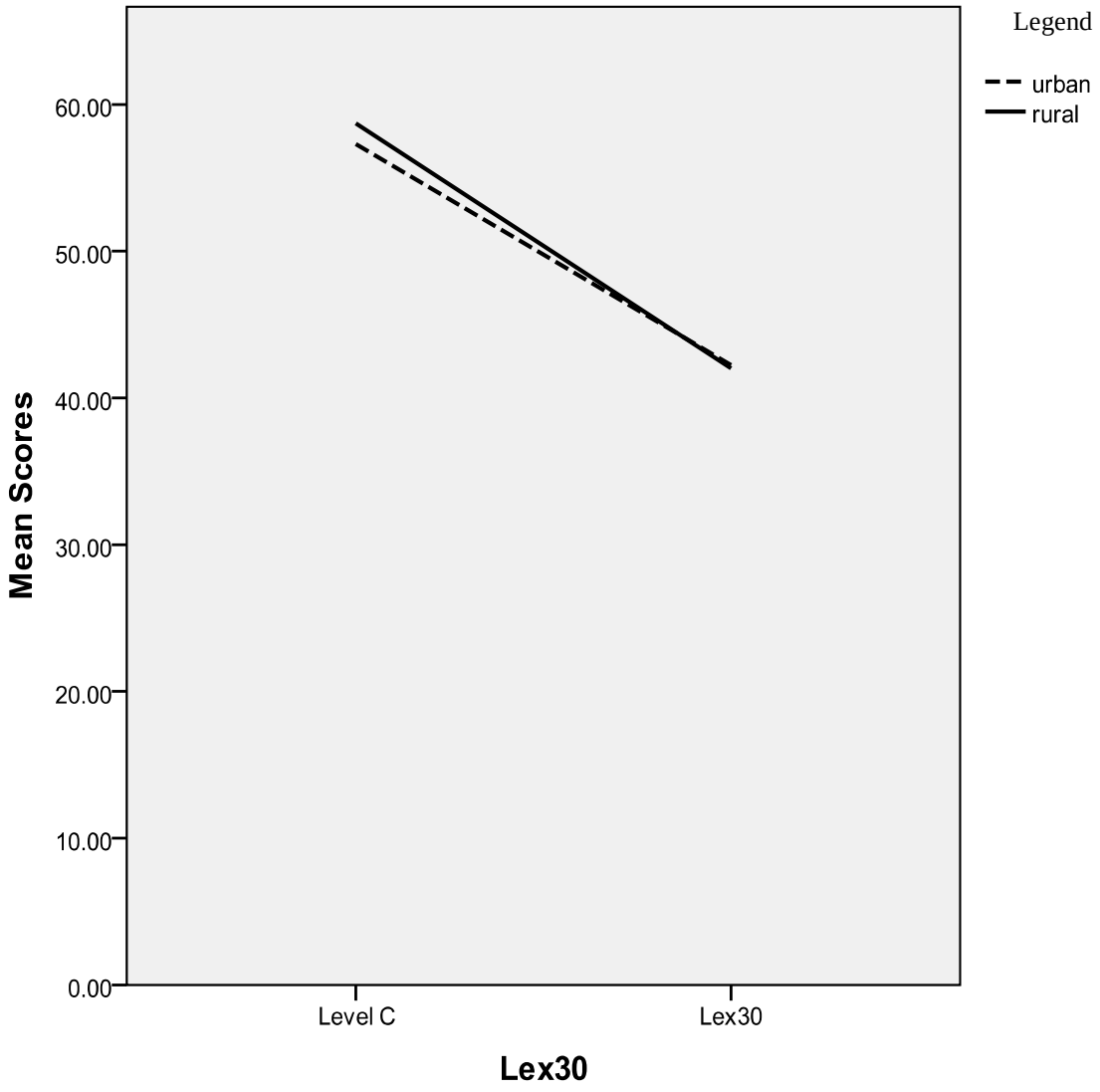

Figure 4.4
Mean scores for Lex30 and Level C

From Figure 4.4, it can be seen that both urban and rural students showed similar pattern in their word production. This figure is enlarged to note the different readings of urban and rural students, as their mean scores are very close. Both urban and rural students produced more high frequency words (Level C) compared to low frequency words (Lex30). The difference is statistically significant after a one-way repeated

measures ANOVA was conducted. The result showed Wilks' Lambda = .63, F (1, 99) = 59.05, p <.0005.

The mean scores between urban and rural students at Level C and Lex30, however, are very close. From the raw data, urban students produced only marginally more low frequency words than rural students and rural students produced slightly more high frequency words compared to urban students. The ratios show that the gap between rural and urban students is very small (99.5% and 102.5%). Statistically, the difference is almost perfectly insignificant for low frequency words (Lex30), Wilks' Lambda = 1.00, F (1, 49) = .01, p = .921 and also statistically insignificant for high frequency words (Level C), Wilks' Lambda = .99, F (1, 49) = .554, p = .460. From this analysis, it can be concluded that urban and rural students produce an almost similar ratio of high and low frequency words when not confined to a set of predetermined words (word list).

4.1.3 Summary of the Results of Research Question 1

In summary to the first research question, there seems to be a difference in the vocabulary sizes of students, according to the type of tests. Based on the translation tests, the study shows that urban students have larger receptive and productive vocabulary sizes at every word level compared to rural students. Both urban and rural students also showed higher scores for higher frequency words compared to lower frequency words. In terms of the gap between urban and rural students' vocabulary size at different word levels, the gap increases as word frequency decreases. This indicates that urban students have only slightly more high frequency words than rural students, but distinctly more low frequency words. However, there is almost no difference between the sizes of the students' receptive and productive vocabulary, hence, there is a balanced receptive and productive vocabulary size for both urban and rural students.

In contrast, from the analysis of scores of Lex30, urban and rural students showed no significant difference in the size of their low-frequency productive vocabulary. However, the result that remains constant is that students produced significantly more high frequency words compared to low frequency words.

Table 4.4 summarizes the findings.

Table 4.4
Summary of the findings of the first research question

Translation Tests	Urban Receptive and Productive vocabulary sizes > Rural Receptive and Productive vocabulary sizes
	Insignificant difference between Receptive and Productive vocabulary sizes of Urban and Rural students
	High frequency words > Low frequency words
	Gap between high frequency words of urban and rural students < Gap between low frequency words of urban and rural students
Lex30	Urban Productive (low-frequency) vocabulary size = Rural Productive (low-frequency) vocabulary size

Note : > more than
< smaller than
= is the same as

4.2 The Relationship Between Receptive and Productive Vocabulary (RQ2)

From the broad analysis on the receptive and productive vocabulary sizes of students earlier, the second research question, seeks to investigate in greater depth, its relationship with each other. The second research question was:

What is the relationship between:

a) the overall receptive and productive vocabulary of students?

b) receptive and productive vocabulary at different word levels of urban and rural students?

It seeks to find out whether the size of one influences the other and whether there exists a predictive element.

Two types of relationships were identified to be investigated. First of all, the relationship between the overall receptive and productive vocabulary of rural and urban students was explored. Secondly, the relationship between receptive and productive vocabulary at different word levels was investigated. The Pearson Product Moment Correlational analysis was used to explore the relationships.

(Note: Lex30 results will only be compared to the overall vocabulary and not at different word levels, as it represents students' overall low frequency productive vocabulary and not at different word levels.)

4.2.1 The Relationship Between The Overall Receptive and Productive Vocabulary

The relationship between the overall receptive and productive vocabulary of urban and rural students was analysed. The scores for urban and rural students were combined for both tests. This was done to investigate the relationship between receptive and productive vocabulary, regardless of students' ecological background. Lex30 was also

included in the analysis as it represents students' overall low frequency productive vocabulary beyond the confines of the word list. Table 4.5 reports the results of the analysis.

Table 4.5

Correlations between students' overall receptive, productive and Lex30 vocabulary

Correlations

		Overall Receptive	Overall Productive	Lex30
Overall Receptive	Pearson Correlation	1	.848**	.081
	Sig. (2-tailed)		.000	.423
	N	100	100	100
Overall Productive	Pearson Correlation	.848**	1	.078
	Sig. (2-tailed)	.000		.443
	N	100	100	100
Lex30	Pearson Correlation	.081	.078	1
	Sig. (2-tailed)	.423	.443	
	N	100	100	100

**. Correlation is significant at the 0.01 level (2-tailed).

Students' overall receptive and productive vocabulary correlate significantly, r = .85, P < .01. It shows a strong positive relationship. Based on these results, it is possible to make predictions. It can be said that those with larger receptive vocabulary sizes would therefore have larger productive vocabulary sizes; conversely, those with smaller receptive vocabulary sizes would also have smaller productive vocabulary sizes and vice versa. However, Lex30 is not in any way affecting or being influenced by the receptive and productive vocabulary, hence, cannot predict the overall receptive and productive vocabulary of students.

4.2.2 The Relationship Between Receptive and Productive Vocabulary at Different Word Levels

For the second part of the research question, the relationship between receptive and productive scores at different word levels was analysed. The analysis was done separately for urban and rural students.

Urban students' results will be reported first, followed by rural students' results.

4.2.2.1 Urban Students

The Pearson Product Moment Correlational analysis was used to explore the relationship between the variables in order to determine their direction, whether they share a significant relationship and if they have the ability to influence and predict one another. The analysis will correlate the relationships between the different word levels of firstly, the receptive vocabulary, secondly, the productive vocabulary, thirdly, the receptive and productive vocabulary and finally, the overall receptive, productive and Lex30 vocabulary. The following tables report the results of the analyses.

Table 4.6

Correlations between urban students' receptive vocabulary at different word levels

Correlations

		Receptive level A	Receptive level B	Receptive level C
Receptive level A	Pearson Correlation	1	.844**	.637**
	Sig. (2-tailed)		.000	.000
	N	50	50	50
Receptive level B	Pearson Correlation	.844**	1	.776**
	Sig. (2-tailed)	.000		.000
	N	50	50	50
Receptive level C	Pearson Correlation	.637**	.776**	1
	Sig. (2-tailed)	.000	.000	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.6 shows the correlation between the receptive vocabulary of urban students at different word levels. The results revealed that statistically significant relationship existed between all of the word levels. Level A is significantly related to Level B at a positively strong $r = 0.84$, $P < .01$ and vice versa, and Level C at a positively strong $r = 0.64$, $P < .01$ and vice versa. Level B is significantly related to Level C at a positively strong $r = 0.78$, $P < .01$ and vice versa. As all three word levels quite strongly and positively correlate with each other, it can safely be said that those who have a larger receptive vocabulary size at any word level will have relatively larger sizes for the other receptive vocabulary word levels; conversely, those who have a smaller receptive vocabulary size at any word level will have relatively smaller sizes for the other word levels. For example, those with larger Level C receptive vocabulary size, tend to have larger Levels B and A sizes; conversely, those who have smaller Level C receptive vocabulary size, tend to have smaller Levels B and C sizes as well. Hence, it has the ability to influence and predict each other.

Table 4.7

Correlations between urban students' productive vocabulary at different word levels

Correlations

		Productive level A	Productive level B	Productive level C
Productive level A	Pearson Correlation	1	.797**	.605**
	Sig. (2-tailed)		.000	.000
	N	50	50	50
Productive level B	Pearson Correlation	.797**	1	.649**
	Sig. (2-tailed)	.000		.000
	N	50	50	50
Productive level C	Pearson Correlation	.605**	.649**	1
	Sig. (2-tailed)	.000	.000	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.7 shows the correlation of productive vocabulary of urban students at different word levels. The results revealed that statistically significant relationship existed between all of the word levels, similar to receptive vocabulary. Level A is significantly related to Level B at a positively strong $r = 0.80$, $P < .01$ and vice versa, and Level C at a positively strong $r = 0.61$, $P < .01$. Level B is significantly related to Level C at a positively strong $r = 0.65$, $P < .01$ and vice versa. Hence, it can safely be said that those who have a larger productive vocabulary size at any word level will have relatively larger sizes for the other productive vocabulary word levels; conversely, those who have a smaller productive vocabulary size at any word level, will have relatively smaller sizes for the other word levels. For example, those with larger Level C productive vocabulary size, tend to have larger Levels B and A sizes; conversely, those who have smaller Level C productive vocabulary size, tend to have smaller Levels B and C sizes as well. The productive vocabulary of urban students, hence, is predictable.

Table 4.8

Correlations between urban students' receptive and productive vocabulary at different word levels

Correlations

		Productive level A	Productive level B	Productive level C
Receptive level A	Pearson Correlation	.701**	.724**	.682**
	Sig. (2-tailed)	.000	.000	.000
	N	50	50	50
Receptive level B	Pearson Correlation	.763**	.769**	.612**
	Sig. (2-tailed)	.000	.000	.000
	N	50	50	50
Receptive level C	Pearson Correlation	.677**	.683**	.521**
	Sig. (2-tailed)	.000	.000	.000
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.8 shows the correlation between receptive and productive vocabulary of urban students at different word levels. The results reveal that statistically significant relationships existed between both receptive and productive vocabulary at all of the word levels. Receptive vocabulary at Level A is significantly related to productive Level A at a positively strong r = 0.70, P < .01, productive Level B at a positively strong r = 0.72, P < .01 and productive Level C at a positively strong r = 0.68, P < .01. Receptive Level B is significantly related to productive Level A at a positively strong r = 0.76, P < .01, productive Level B at a positively strong r = 0.77, P < .01 and productive Level C at a positively strong r = 0.61, P < .01. Receptive Level C is significantly related to productive Level A at a positively strong r = 0.68, P < .01, productive Level B at also a positively strong r = 0.68, P < .01 and productive Level C at a positively moderate r = 0.52, P < .01. Productive levels A, B and C, hence, are significantly related to receptive levels A, B and C. As all three word levels between receptive and productive vocabulary quite strongly

and positively correlate with each other, it can safely be said that those who have a larger receptive vocabulary size at any word level will have relatively larger sizes for productive vocabulary at every word level; conversely, those who have a smaller receptive vocabulary size at any word level will have relatively smaller productive vocabulary sizes at every word level, and vice versa. For example, those with larger Level C receptive vocabulary size, tend to have larger productive vocabulary at Levels A, B and C; conversely, those who have smaller Level C receptive vocabulary size, tend to have smaller productive vocabulary at Levels A, B and C as well.

Table 4.9

Correlations between urban students' overall receptive, productive and Lex30 vocabulary

Correlations

		Overall Receptive	Overall Productive	Lex30
Overall Receptive	Pearson Correlation	1	.846**	-.105
	Sig. (2-tailed)		.000	.467
	N	50	50	50
Overall Productive	Pearson Correlation	.846**	1	.032
	Sig. (2-tailed)	.000		.824
	N	50	50	50
Lex30	Pearson Correlation	-.105	.032	1
	Sig. (2-tailed)	.467	.824	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

With a strong statistically significant relationship between receptive and productive vocabulary at every word level, receptive and productive vocabulary revealed an overall positively strong relationship, $r = .85$, $P < .01$. A large receptive vocabulary size would indicate a large productive vocabulary size and vice versa. A large productive vocabulary size would also indicate a large receptive vocabulary size and vice versa.

However, Lex30 scores did not show any significant correlation between the overall receptive and productive scores. This shows that Lex30 is not in any way affecting or influenced by the receptive and productive scores and vice versa.

4.2.2.2 Rural Students

Table 4.10

Correlations between rural students' receptive vocabulary at different word levels

Correlations

		Receptive level A	Receptive level B	Receptive level C
Receptive level A	Pearson Correlation	1	.579**	.348*
	Sig. (2-tailed)		.000	.013
	N	50	50	50
Receptive level B	Pearson Correlation	.579**	1	.527**
	Sig. (2-tailed)	.000		.000
	N	50	50	50
Receptive level C	Pearson Correlation	.348*	.527**	1
	Sig. (2-tailed)	.013	.000	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Where rural students are concerned, Table 4.10 shows the correlation within the receptive vocabulary at different word levels. The results revealed that statistically significant relationship existed between all of the word levels, although slightly lower than urban students' results. For rural students, receptive Level A is significantly related to Level B at a positively moderate $r = 0.58$, $P < .01$ and Level C at a positively weak $r = 0.35$, $P < .05$. Level B is significantly related to Level A at a positively moderate $r = 0.58$, $P < .01$ (as have been reported) and Level C at a positively moderate $r = 0.53$, $P < .01$. Level C is significantly related to Levels A and B as have been reported. As all three word levels quite moderately and positively correlate with each other, it can safely be said that

to some extent, those who have a larger receptive vocabulary size at any word level will have relatively larger sizes for the other receptive vocabulary word levels; conversely, those who have a smaller receptive vocabulary size at any word level will have relatively smaller sizes for the other word levels. For example, to some extent those with larger Level C receptive vocabulary size, tend to have larger Levels B and A sizes; conversely, those who have smaller Level C receptive vocabulary size, tend to have smaller Levels B and C sizes as well. Hence, the ability to 'predict' and 'influence' one another for rural students' receptive vocabulary exist, but is weaker than that of urban students'.

Table 4.11

Correlations between rural students' productive vocabulary at different word levels

Correlations

		Productive level A	Productive level B	Productive level C
Productive level A	Pearson Correlation	1	.670**	.620**
	Sig. (2-tailed)		.000	.000
	N	50	50	50
Productive level B	Pearson Correlation	.670**	1	.534**
	Sig. (2-tailed)	.000		.000
	N	50	50	50
Productive level C	Pearson Correlation	.620**	.534**	1
	Sig. (2-tailed)	.000	.000	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.11 shows the correlation between the productive vocabulary of rural students at different word levels. The results revealed that statistically significant relationship existed between all of the word levels, similar to the receptive vocabulary test. Again, the value of significance here is slightly lower than that of urban students. However, it is slightly stronger than the significance level of rural students' receptive

scores. Productive Level A here is significantly related to Level B at a positively strong r = 0.67, P < .01 and vice versa, and Level C at a positively strong r = 0.62, P < .01 and vice versa. Level B is significantly related to Level C at a positively moderate r = 0.53, P < .01 and vice versa. Hence, it can safely be said that those who have a larger productive vocabulary size at any word level will have relatively larger sizes for the other productive vocabulary word levels; conversely, those who have a smaller productive vocabulary size at any word level, will have relatively smaller sizes for the other word levels. For example, those with larger Level C productive vocabulary size, tend to have larger Levels B and A sizes; conversely, those who have smaller Level C productive vocabulary size, tend to have smaller Levels B and C sizes as well.

Table 4.12

Correlations between rural students' receptive and productive vocabulary at different word levels

Correlations

		Productive level A	Productive level B	Productive level C
Receptive level A	Pearson Correlation	.536**	.421**	.421**
	Sig. (2-tailed)	.000	.002	.002
	N	50	50	50
Receptive level B	Pearson Correlation	.638**	.400**	.432**
	Sig. (2-tailed)	.000	.004	.002
	N	50	50	50
Receptive level C	Pearson Correlation	.656**	.595**	.521**
	Sig. (2-tailed)	.000	.000	.000
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.12 shows the correlation between receptive and productive vocabulary test scores of rural students at different word levels. The results revealed that statistically

significant relationship existed between both receptive and productive vocabulary at all of the word levels. When compared to urban students' correlational results, the analysis for rural students, in general, showed slightly weaker relationships. For rural students, receptive vocabulary at Level A is significantly related to productive Level A at a positively moderate $r = 0.54$, $P < .01$, productive Level B at a positively moderate $r = 0.42$, $P < .01$ and productive Level C at also a positively moderate $r = 0.42$, $P < .01$. Receptive Level B is significantly related to productive Level A at a positively strong $r = 0.64$, $P < .01$, productive Level B at a positively moderate $r = 0.40$, $P < .01$ and productive Level C at a positively strong $r = 0.43$, $P < .01$. Receptive Level C is significantly related to productive Level A at a positively strong $r = 0.66$, $P < .01$, productive Level B at also a positively strong $r = 0.60$, $P < .01$ and productive Level C at a positively moderate $r = 0.52$, $P < .01$. Hence, productive levels A, B and C are significantly related to receptive levels A, B and C as well. As all three word levels between receptive and productive vocabulary positively correlate with each other moderately, it can be said that to some extent, those who have a larger receptive vocabulary size at any word level will have relatively larger sizes for productive vocabulary at every word level; conversely, to some extent, those who have a smaller receptive vocabulary size at any word level will have relatively smaller productive vocabulary sizes at every word level, and vice versa. For example, those with larger Level C receptive vocabulary size, to some extent, tend to have larger productive vocabulary at Levels A, B and C; conversely, those who have smaller Level C receptive vocabulary size, to some extent tend to have smaller productive vocabulary at Levels A, B and C as well. Hence, the ability to predict and influence one another is reinforced.

Table 4.13

Correlations between rural students' overall receptive, productive and Lex30 vocabulary

Correlations

		Overall Receptive	Overall Productive	Lex30
Overall Receptive	Pearson Correlation	1	.733**	.280*
	Sig. (2-tailed)		.000	.049
	N	50	50	50
Overall Productive	Pearson Correlation	.733**	1	.127
	Sig. (2-tailed)	.000		.378
	N	50	50	50
Lex30	Pearson Correlation	.280*	.127	1
	Sig. (2-tailed)	.049	.378	
	N	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

With a moderate statistically significant relationship between receptive and productive vocabulary at every word level, the overall receptive and productive vocabulary of rural students revealed an overall positively strong relationship, $r = .73$, $P < .01$. This relationship is naturally slightly lower than urban students' overall results, considering the prior analyses (see TABLES 4.10 to 4.12). From this result, it can be interpreted that a large receptive vocabulary size would indicate a large productive vocabulary size and vice versa. A large productive vocabulary size would also indicate a large receptive vocabulary size and vice versa. Hence, a strong 'predictability'.

Lex30 also shares a statistically significant relationship with the overall receptive vocabulary, $r = 0.28$, $P < .05$. This is in contrast with the result of the urban students' analysis, which did not show any significant correlation. As this is quite a weak relationship, it can be said that rural students' production of lower frequency words may be influenced by their receptive vocabulary size to some extent; conversely, their receptive vocabulary size may be influenced by the production of low frequency words.

Hence, those who produce more low frequency words may have larger receptive vocabulary size to some extent; conversely, those who produce less low frequency words may have smaller receptive vocabulary size to some extent. The opposite is also true. Hence, Lex30 has some ability to predict and influence rural students' receptive vocabulary.

4.2.3 Summary of the Results of Research Question 2

In summary, the response to the second research question shows that both urban and rural students' receptive and productive vocabulary have a statistically significant positive relationship between the overall scores of receptive and productive vocabulary. This indicates that receptive and productive vocabulary are influenced by and can predict one another. However, there was no significant relationship between receptive and productive vocabulary and Lex30.

When correlating the scores at different word levels, separately for urban and rural students, statistically significant positive relationships were found. Urban students' correlation results revealed generally stronger relationships compared to rural students' results in all areas for this part of the analysis, except when compared to Lex30. Urban students' receptive and productive vocabulary scores were not statistically significant to Lex30. However, where rural students are concerned, their receptive vocabulary score was statistically significant with Lex30, but it is only a weak relationship.

Hence, it can be said that the receptive and productive vocabulary of students are predictable. The size of one may be able to predict the size of the other. When receptive vocabulary is large, the productive vocabulary also tends to be large, to some extent, and vice versa. Urban students' scores also tend to be more predictable than rural students' scores. However, Lex30 scores, or the ability to produce low frequency words, cannot be

predicted as it does not share statistically significant relationships with the other variables, although it may influence or be influenced slightly by rural students' receptive vocabulary.

Table 4.14 summarizes the findings related to the second research question of the study.

Table 4.14
Summary of the findings of the second research question

Translation Tests	Overall receptive vocabulary correlates with overall productive vocabulary strongly, and also at every word level
	Strength of relationships for urban students > rural students
Lex30	Only a weak correlation between Lex30 and rural students' Receptive vocabulary

Note : >greater than

4.3 The Extent of the Correlation between Vocabulary Size and ELA (RQ3)

With past research and studies showing significant correlation between vocabulary size and ELA (Berne & Blachowicz, 2008, Chall & Snow, 1988, Morris & Cobb, 2004, Zareva et.al., 2005) (see CHAPTER 2 Section 2.7.2), this study seeks to explore this possibility with Malaysian students. With encouraging results from the previous findings from research questions one and two, the analyses that will be conducted here will shed more possibilities to the significance of vocabulary in the teaching and learning of English in Malaysia.

The third research question of this study is *'To what extent do the vocabulary sizes of urban and rural students correlate with their English language achievement (ELA)?'* To answer this research question, ELA was correlated with firstly, the overall vocabulary size (receptive + productive) using the Pearson Product Moment Correlational analysis. Separate analysis was also carried out to distinguish between urban and rural students' correlation. Lex30 results were not correlated as it only represents students' low frequency productive vocabulary.

4.3.1 Overall Correlation between Vocabulary Size and ELA

To gain an overview of the extent of the correlation between students' vocabulary sizes and ELA, regardless of the ecological factor and the different vocabulary types (receptive and productive), an overall correlational analysis was done. The overall vocabulary size is derived form the total receptive and productive vocabulary scores of urban and rural students. The following table reports the result of the analysis:

Table 4.15
Correlations between students' overall vocabulary size and ELA

Correlations

		Overall Vocabulary Size
ELA	Pearson Correlation	.503**
	Sig. (2-tailed)	.000
	N	100

**. Correlation is significant at the 0.01 level (2-tailed).

The result of the analysis revealed that a statistically significant relationship existed between the overall vocabulary size and ELA, $r = 0.50$, $P < .01$. This means a statistically significant positive relationship existed between the overall vocabulary size

and ELA of students. The relationship between the two variables is moderate. To some extent, those who have larger vocabulary sizes tend to score better in the English test, and conversely, those who score better in the English test tend to have larger vocabulary sizes as well.

To investigate this moderate relationship further, the analysis was also conducted separately for urban and rural students.

4.3.2 Correlation between Urban Students' Vocabulary Size and ELA

The analysis for the overall vocabulary size and ELA will be presented first, followed by separate analyses between receptive and productive vocabulary and ELA.

Table 4.16
Correlations between urban students' overall vocabulary size and ELA

Correlations

		Overall Vocabulary Size
ELA	Pearson Correlation	.291*
	Sig. (2-tailed)	.041
	N	50

*. Correlation is significant at the 0.05 level (2-tailed).

The overall vocabulary size of urban students showed a statistically significant relationship with their ELA, $r = 0.29$, $P < .05$. It concurs with the overall finding. However, this relationship is quite weak. Urban students' vocabulary size affect and predict their ELA and vice versa only to some extent.

The following separate analysis between receptive and productive vocabulary sizes and ELA may shed more light on this result.

Table 4.17

Correlations between urban students' overall receptive vocabulary size and ELA

Correlations

		Overall Receptive Vocabulary Size
ELA	Pearson Correlation	.281*
	Sig. (2-tailed)	.048
	N	50

*. Correlation is significant at the 0.05 level (2-tailed).

The results of the analysis (Table 4.17) revealed that a statistically significant relationship existed between the overall receptive vocabulary size and ELA, $r = 0.28$, $P < .05$. This also supports the previous results, where it reveals a positive relationship and the strength of it concurs with the overall results of the urban students. Hence, the ability to influence and predict between urban students' receptive vocabulary and ELA can only be reliable to some extent.

Table 4.18

Correlations between urban students' overall productive vocabulary size and ELA

Correlations

		Overall Productive Vocabulary Size
ELA	Pearson Correlation	.279
	Sig. (2-tailed)	.050
	N	50

However, the same cannot be said of the relationship between urban students' productive vocabulary size and ELA. The results, as shown in Table 4.18, reveal that there is no statistically significant relationship between them. In the output, the correlation is a low $r = 0.28$, which is not statistically significant, where the P value is 0.050. Hence, the

productive vocabulary size of urban students does not influence their ELA; conversely, their ELA does not predict or determine their productive vocabulary size.

This may be the reason why urban students' overall correlational strength is quite weak.

4.3.3 Correlation between Rural Students' Vocabulary Size and ELA

For rural students, the results of the analyses will be presented similar to the sequence of the urban students. However, the analyses here reports contrasting results compared to urban students' results.

Table 4.19
Correlations between rural students' overall vocabulary size and ELA

Correlations

		Overall Vocabulary Size
ELA	Pearson Correlation	.257
	Sig. (2-tailed)	.075
	N	49

The overall correlational analysis between rural students' overall vocabulary size and ELA revealed no significant relationship. This may have contributed to the overall moderate relationship between vocabulary size and ELA. Hence, rural students' overall vocabulary size cannot indicate their ELA and vice versa.

Table 4.20

Correlations between rural students' overall receptive vocabulary size and ELA

Correlations

		Overall Receptive Vocabulary Size
ELA	Pearson Correlation	.164
	Sig. (2-tailed)	.254
	N	50

Rural students' overall receptive vocabulary size also concurs with this result. The analysis revealed that there is no statistically significant relationship between the overall receptive vocabulary size and ELA, r = 0.16, P = .254. This means that the overall receptive vocabulary size of rural students does not affect or influence their ELA and vice versa.

Table 4.21

Correlations between rural students' overall productive vocabulary size and ELA

Correlations

		Overall Productive Vocabulary Size
ELA	Pearson Correlation	.281*
	Sig. (2-tailed)	.048
	N	50

*. Correlation is significant at the 0.05 level (2-tailed).

However, the overall productive vocabulary size of rural students painted a different picture. The results revealed that there is a statistically significant relationship between the overall productive vocabulary size and ELA (Table 4.21). In the output, the correlation recorded r = 0.28, P < .05. This is only a weak relationship which may not

have been significant enough for it to contribute to rural students' overall vocabulary size to indicate any significant relationship with ELA. Hence, this can only suggest that rural students who have larger productive vocabulary size may also do well in the English test; conversely, those who have smaller productive vocabulary size may not do as well in the English test.

4.3.4 Summary of the Results of Research Question 3

In summary, the study have shown that there exist a significant relationship between ELA and the receptive and productive tests, albeit a moderate one. The correlational results are indicative that students' vocabulary sizes have the ability to predict their ELA to some extent. However, urban and rural students' separate analyses showed different patterns in their correlation. First of all, urban students' overall vocabulary size, receptive vocabulary size in particular, revealed that it can only predict ELA marginally, due to a weak relationship. Secondly, rural students' overall vocabulary size did not reveal any ability to predict ELA; nevertheless, their productive vocabulary size may do so marginally.

Table 4.22 provides a summary of the findings related to the third research question.

Table 4.22
Findings of the third research question

Translation Tests	Overall vocabulary size correlates with ELA moderately : -Urban students' overall vocabulary size correlates with ELA -Rural students' overall vocabulary size does not correlate with ELA
	Urban students' receptive vocabulary size correlates with ELA, whereas rural students' productive vocabulary size correlates with ELA

All the analyses on the three research questions of this study have revealed encouraging results. These findings may indicate that vocabulary is an area which is worth exploring. The following chapter will discuss in greater detail the factors and possibilities pertinent to these findings.