



Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

A study of English reading ability based on multiple linear regression analysis

Lin Yu

Hubei University of Technology, Hubei Wuhan, China

ABSTRACT

As China's development continuously integrates with the whole world, English is more widely used as a world language. The understanding of technical papers which serve as the main carrier of technical exchanges has become the most concerned issue by people, and English reading ability is one of the crucial abilities influencing the reasonable uses of technical documentations. This paper aims to make some contributions to the exploration of ways of improving English reading ability by analyzing the influencing factors upon English reading ability. This paper starts with questionnaires on the influencing factors upon English reading ability and analyses of exam questions evaluating reading ability, and discovers that the subject factor, detail factor, inference factor, attitude factor and semantic factor are the major influencing factors upon English reading ability. Then, it elaborates the mathematical theories and application range of multiple linear regression models to provide theoretical basis for the exploration of the relationship between English reading ability and these five influencing factors. Next, an English reading ability test on a hundred students in the School of Electric Automation is conducted, sample data are obtained by using the test results and data processing method, a multiple linear regression model of English reading ability based on the sample data is built and the influencing factors of English reading ability have been attained, which provides a foundation for the probe of ways of improving English reading ability. Finally, based on the multiple linear regression models, a way of improving English reading ability by opportunely using intensive and extensive reading skills is presented.

Key words: English reading ability, influencing factors, multiple linear regression, SPSS statistical software.

INTRODUCTION

Any of human being's acquired ability has its own influences, if the influencing factors of an ability can be scientifically grasped and be fulfilled through practices, then it meets the theorem of "efforts+ methods=success". Likewise, the scientific exploration of the ways of improving English reading ability should start with its influencing factors. This paper seeks ways of improving reading ability for English learners, conducts a model study of them based on multiple linear regression methods and aims to contribute to the discovery of major influencing factors and the improving of reading ability.

Many scholars have made great efforts on the researches of improving approaches of English reading ability, and it is their efforts that enable contemporary China joins world's advanced fields. For example, Li Yanglong and his counterparts analyze the question types and contents in Comprehension part of 117 test papers chosen from the National Matriculation Exam papers between 2007 and 2012 and also from English test papers in provinces and cities which are granted independent rights for designing exam papers, point out the deficiencies in testing methods and the contents of English reading ability in current College Entrance Examinations and offer their advice on reform direction. Bao Gui thinks the core of language threshold hypothesis test is whether there is a interaction effect between the second language proficiency as the continuous variable and the native language reading ability when predicting the second language reading ability. He also puts forward the idea that the most appropriate

statistical analytical method should be the multiple regression analysis with interaction effect. Wang Zongying and others probe the impacts of native language reading ability and foreign language proficiency on high school students' reading ability and manage to verify the threshold hypothesis.

Based on previous studies, this paper analyzes the English test results of 100 students from the School of Electric Automation in a university, builds a multiple linear regression model of English reading ability and its influencing factors and aims to offer some more scientific ways of improving English reading ability.

2 AN ANALYSIS OF THE INFLUENCING FACTORS OF ENGLISH READING ABILITY

Reading ability is one of the yardsticks measuring a person's English proficiency and accounts for a large proportion of scores in Reading Comprehension part in various English exams, thus, examinees pay great efforts in reading ability training. Because of the CET 4 and CET 6 examinations during the college years, English learning has obtained a strongest learning atmosphere in campuses. English reading occupies an enormously important status in CET 4 and CET 6 examinations, as a result, there is such a saying: "if you can get good scores in Reading Comprehension part, then you can get good scores in the English exams". Therefore, it is very worthwhile to analyze the influencing factors of English reading ability.

English reading comprehension is process in which the students comprehend the authors' ideas and thoughts based on the language information they have acquired. This process can be interpreted as students' English cognitive process and thus the reading comprehending process can be treated as a kind of psychological process to analyze. The evaluation of the advantages and disadvantages of this psychological process needs to be graded through designing the examination questions: a higher mark means a superior reading ability, while a lower mark means an inferior one. To some extent, the marks can reflect students' reading ability, though the reading ability measured by marks is not wholly fair.

The factors influencing the reading ability are rooted in the factors of learners themselves which refer to their cognitive ability and the initial knowledge and ability, learning motives and attitudes towards the target knowledge, and learning methods and habits, psychological factors and characteristics. Due to the difficulties in analyzing the psychological factors and characteristics, this article starts with analyses of the above-listed three major factors:

1)Cognitive ability and the initial knowledge and ability refer to the learner's ability of grasping and using language knowledge such as vocabulary, phrases, sentences and linguistic rules. The exploration of the cognitive ability and the initial knowledge and ability can be analyzed from the following three aspects: the number of new words, the complexity of sentence structure and familiarity with the reading content.

2)The learning motive means the learner's internal impetus which directly motivates him to learn the language and is the force which initiates the learning action, maintains it and directs it to a certain goal. Learning attitude refers to the learner's cognitive, emotional and behavioral tendencies in learning activity. The exploration of the learning motives and learning attitudes can be analyzed from the following five aspects: the instrumental motivation, integrative motivation, attention condition, emotion condition and willpower condition.

3)According to different reading ways, the reading methods can be categorized into skimming, scanning and careful reading. Having grasped good learning methods allows the learners to shorten the reading time-span and improve reading efficiency. Learning habits refer to a kind of stable learning way that the learner has manifested during the learning process.

This paper has done a questionnaire survey based on the above-mentioned three influencing factors of the English reading ability. The questionnaire respondents are 150 students from the School of Electric Automation in a university, the research content includes A-the number of new words, B-the complexities of sentence structures, C-familiarities with the reading content, D-the degrees of learning interests, E-the positive degree of learning attitudes, F-word by word reading, G-consulting reference books whenever encountering new words, H-liking repeating the sentences that have been read, I-the uses of skimming and scanning. The respondents are required to choose among these nine influencing factors ranging from A to I, and then an evaluation of the influencing degrees of these factors on the reading ability based on the number of each factor being chosen by the respondents is made, the survey result is as shown in Table 1:

There are usually five questions following the reading comprehension content, the first question is summarizing the article's main idea or the gist, the second question is about the details of something mentioned in the article, the third question is a inference question, namely, inferring the right answer based on the title and the content of the original article, the fourth question is one about the author's attitude based on the judgment of the author's attitudes towards

the question, and the last one is semantic one which deals with English translations of the given sentences. The fifth question is aiming to test the reader's reading ability. Based on the factors of influencing the correctness of answers for these five questions, this paper summarizes eight factors such as English vocabulary, background information, basic grammatical knowledge, cultural differences, discourse comprehension, reading strategies and reading habits. These factors can be further classified into the English language factors and non-English language factors when being measured by language factors. English language factors include vocabulary and grammar, and non-English factors include background information, discourse comprehension, reading strategies and reading habits.

Tab. 1: The survey results list of importance degrees of influencing factors on the English reading ability

Factor symbols	Factor content	Number of people having chose this factor	percentage	Importance degrees ranking
A	The number of new words	115	76.67%	2
B	Complexities of sentence structures	106	70.67%	4
C	Familiarities with reading contents	75	50.00%	5
D	Degrees of learning interests	47	31.33%	7
E	Positive degrees of learning attitudes	41	27.33%	8
F	Word by word reading	117	78.00%	1
G	consulting reference books whenever encountering new words	107	71.33%	3
H	liking repeating the sentences that have been read	69	46.00%	6
I	the uses of skimming and scanning	31	20.67%	9

The ranking of importance degrees is like this: F>A>G>B>C>H>D>E>I, when analyzing these figures on the basis of cumulative importance degrees after the percentages being standardized, Table 2 can be obtained:

Tab. 2: The results list of importance degrees ranking of influencing factors on English reading ability and the cumulative degrees

Factor categories	F	A	G	B	C	H	D	E	I
Importance degrees	16.5%	16.2%	15.1%	15.0%	10.6%	9.8%	6.6%	5.8%	4.4%
Cumulative degrees	16.5%	32.8%	47.9%	62.9%	73.5%	83.2%	89.8%	95.6%	100%

In conclusion, the influencing factors of reading ability can be measured by the comprehension ability of these five questions in the Reading Comprehension part, namely, the gist understanding ability, detail understanding ability, inference ability, attitude judging ability and semantic understanding ability. The ways of improving reading ability can be obtained, since each kind of understanding ability has the corresponding influencing factor in Table 1. This paper seeks the correlation between total points of the reading comprehension part and the average points of each question, explores the relationship between the English reading ability and five different comprehension abilities and aims to provide proper ways for students to improve their English reading ability.

3. MULTIPLE LINEAR REGRESSION THEORY FOUNDATION

The reading ability reflects the total point of test paper questions, the influencing factors of reading ability can be measured by using the gist understanding ability, detail understanding ability, inference ability, attitude judging ability and semantic understanding ability. In order to explore the correlation between these five influencing factors and the English reading ability, this paper applies the multiple linear regression method, studies the importance degrees of five influencing factors and the relationship between them and provides scientific ways for English learners to improve their reading abilities.

The theoretical models of the multiple linear regression analysis can be categorized into two types, namely, the research models between independent variables and the variables, and the research models between independent variables and independent variables. In order to more scientifically and rationally apply the multiple linear regression model to study the English reading ability, this chapter elaborates the theoretical models from the following two aspects:

3.1 Data table basic parameters

If m variables such as x_1, x_2, \dots, x_m are sampled n times, the obtained sample points n is shown by the formula(1):

$$(x_{i1} \quad x_{i2} \quad \dots \quad x_{im}), i = 1, 2, \dots, n \quad (1)$$

Formula(1) constitutes a data table which is a dimensional matrix \mathbf{X} of $n \times m$, as formula (2) shows:

$$\mathbf{X} = (x_{ij})_{n \times m} = \begin{bmatrix} e_1^T \\ \vdots \\ e_n^T \end{bmatrix} \quad (2)$$

In Formula(2), e_i^T is the i sample point in Formula (2), the solving of the mean sample data \bar{x} is as shown by the Formula(3)

$$\begin{cases} \bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m) \\ \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}, j = 1, 2, \dots, m \end{cases} \quad (3)$$

The sample covariance matrix \mathbf{S} and the sample correlation coefficient matrix \mathbf{R} is as shown in Formula(4)

$$\begin{cases} \mathbf{S} = (s_{ij})_{m \times m} = \frac{1}{n} \sum_{k=1}^n (e_k - \bar{x})(e_k - \bar{x})^T \\ \mathbf{R} = (r_{ij})_{m \times m} = \left(\frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}} \right); s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \end{cases} \quad (4)$$

3.2 Multivariable linear regression model

In real problems, there are generally more than one influential factors and there is no linear relationship between influential factors and the trend. Therefore, multivariable linear regression is an ideal way to find out the function relationship between these factors and real issues. This paper studies the relationship between reading ability and five influential factors. These factors are regarded as five variables, which constitute the five-variable regression model. If the random variable y is relevant to p ($p \geq 2$) regular variables $x_1, x_2, x_3, \dots, x_p$ and meets the equation(5), then this equation is called the mathematic descriptive method of multivariable linear regression.

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_p x_p + \varepsilon \\ E(\varepsilon) = 0, \text{Var}(\varepsilon) = \delta^2 < +\infty \end{cases} \quad (5)$$

Call equation (5) the linear regression model of p -variable theory, $\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_p$ the regression coefficients, and $x_1, x_2, x_3, \dots, x_p$ the regression factors or design factors, or shortly the factors. Reflecting the influence of factor x_i ($i = 1, 2, 3, \dots, p$) on the observation value y , β_i ($i = 1, 2, 3, \dots, p$) is also called the effect of factor x_i ($i = 1, 2, 3, \dots, p$).

Suppose that there are n groups of sample observations $(x_{i1}, x_{i2}, x_{i3}, \dots, x_{ip}; y_i)$ ($i = 1, 2, 3, \dots, n$), equation (6) can be deduced from equation (5).

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_p x_{ip} + \varepsilon_i \\ E(\varepsilon_i) = 0, \text{Var}(\varepsilon_i) = \delta^2 < +\infty \end{cases} \quad (6)$$

in which $i = 1, 2, 3, \dots, n$ and $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n$ are mutually independent. The matrix expression of equation (6) is equation (7).

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \quad (7)$$

$$\begin{cases} Y = X\beta + \varepsilon \\ E(\varepsilon) = 0, \text{Var}(\varepsilon) = \delta^2 I_n \end{cases}$$

in which I_n is n -order unit matrix, 1_n is the n -dimensional column vector of elements whose value is 1. Call Y the observation vector of random variable, β the unknown parameter vector, X the design matrix and ε the n -dimensional error vector. Suppose in the regression analysis that $\text{rank}(X) = p + 1$, which means that X is required to be full column rank. Therefore, $E(Y) = X\beta$, $\text{Var}(\varepsilon) = \delta^2 I_n$, in which $\varepsilon \sim N(0, \delta^2)$. Then the expression of equation (7) is shown as equation (8).

$$\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \delta^2 I_n) \end{cases} \quad (8)$$

3.3 The calculation method of regression coefficient

The least square method is adopted to figure out the estimated column vector $\hat{\beta}$ of regression coefficient in model (7), and the sum of error squares of regression coefficient is as equation (9) shows:

$$Q(\beta_0, \beta_1, \beta_2, \dots, \beta_p) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \dots - \beta_p x_{ip})^2 \quad (9)$$

According to the idea of the least square method, the group with the minimum sum of error squares is chosen to estimate the parameter of regression coefficient. That is to say, if $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$ exists and can meet the equation (10),

$$Q(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p) = \min\{Q(\beta_0, \beta_1, \beta_2, \dots, \beta_p)\} \quad (10)$$

3.4 Regression effect and the significance test of regression coefficient

The closeness between y and the linear relationship of $x_1, x_2, x_3, \dots, x_p$ is relevant to the proportion of the sum of regression squares to the total sum of squares. And the sum of regression squares U and the total sum of squares L_{yy} are expressed as the following equation (11):

$$\begin{cases} L_{yy} = Q + U = \sum_{i=1}^n (y_i - \bar{y})^2 \\ Q = \sum_{i=1}^n (y_i - \hat{y})^2; U = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \end{cases} \quad (11)$$

Call $R = \sqrt{\frac{U}{L_{yy}}}$ the negative correlation coefficient (goodness of fit) of samples of y and $x_1, x_2, x_3, \dots, x_p$.

In the application of the multi-variable linear regression, the negative correlation coefficient R represents the goodness of fit of regression equations to sample data and has a range of $0 \leq R^2 \leq 1$. The closer the value approaches 1, the greater the fitness is. But when the number of independent variables is close to that of samples, R will easily approach 1. Then second thoughts are required whether the model's goodness can be decided by R .

When the regression effect is significant, it means that $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ are not totally 0 whereas 1 or $m(m < p)$ β_i are 0, which further reflects that y is irrelevant to x_i and the role of x_i is replaced by $x_j (j \neq i)$, and then x_i should be eradicated from the regression equation.

4. EMPIRICAL ANALYSIS

4.1 Statistics of influential factors of English reading

Through testing 100 sophomores who major in electric automation, test content, accumulation of words and sentence patterns, grading standard, data collection and treatment are illustrated as follows.

1) Test content: choose the first reading from the English exam in post-graduate entrance examinations in 2008, 2009, 2010, 2012 and 2013. The test is designed by the difficulty standard of post-graduate entrance examination, in which the first article is easy. Difficulty coefficients in choosing articles and designing questions are the same in these five years of examinations papers.

2) Vocabulary and sentence pattern: Students who take the test have finished the English readings for the above five years of postgraduate entrance examination, but they have not received any trainings for this exam. With different reading abilities, students are asked to prepare the vocabulary and sentence pattern one week in advance.

3) Grading standard: There are 5 articles and 25 questions in the English paper. The questions in the test are marked by year, namely 1-5 for year 2008, 6-10 for year 2009...and 21-25 for year 2013, and each question has 4 scores, totaling 100 scores.

4) Data collection: the total score and each question's score of 100 students are collected.

5) Data processing: scores are ranked from high to low and every five neighboring students are assembled in a group to get the moving average value. Average total score from question 1 to 5 represents the first sample value of reading ability, and average total score from questions 6 to 10 represents the second sample value. Likewise, 96 sample values of reading ability can be deduced in total. For the first student of each group, he or she will choose questions 1, 6, 11, 16, 21, and their mean score is the sample value of principal independent variable; for the second student of each group, he or she may choose questions 2, 7, 12, 17, 22, and their mean score is the sample value of detailed independent variable; for the third student of each group, he or she may choose questions 3, 8, 13, 18, 23, and their mean score is the ample value of deduced variable; for the fourth student, he or she may choose questions 4, 9, 14, 19, 24, and their mean score is the sample value of attitude variable; for the fifth student, he or she may choose questions 5, 10, 15, 20, 25, and their mean score is the sample value of sematic variable. This paper will explore the relationship between independent variable and dependent variable sample values.

Sample data is collected according to the above five principles and the result is as Table 3 shows:

Tab. 3: Sample data

sample group	variable sample value	sample value of independent variable 1	sample value of independent variable 2	sample value of independent variable 3	sample value of independent variable 4	sample value of independent variable 5
1	90.26	3.91	3.92	2.35	3.59	3.35
2	89.60	3.54	3.83	2.31	3.61	2.46
3	89.43	2.99	3.75	2.54	2.76	3.14
4	88.97	3.76	3.26	2.35	3.22	2.51
5	88.46	3.02	3.57	3.15	2.87	2.73
6	88.28	3.15	3.38	3.00	3.48	3.67
7	86.19	2.17	3.73	2.57	2.99	2.63
8	85.31	3.37	3.18	2.31	2.68	2.63
9	84.97	2.68	3.09	3.22	3.64	3.31
10	84.95	2.02	2.55	3.48	3.63	3.38
11	84.76	3.41	2.56	3.49	2.28	3.81
12	84.64	3.98	2.58	2.16	2.19	2.64
13	84.51	3.55	2.32	2.15	2.13	2.96
14	84.22	3.50	2.29	2.12	2.12	2.59
15	83.46	2.40	2.19	2.21	2.39	3.53
⋮	⋮	⋮	⋮	⋮	⋮	⋮
96	42.34	1.19	0.22	1.34	1.76	0.94

4.2 The establishment and analysis of multi-variable linear regression model

A SPSS multi-variable linear regression analysis of the first 15 groups of data in Table (3) is made. In the SPSS analysis, VAR00006 represents the variable-reading ability, VAR00001 refers to independent variable 1 – theme factor, VAR00002 refers to independent variable 2 – detail factor, VAR00003 refers to independent variable 3 – deduction factor, VAR00004 refers to independent variable 4 – attitude factor, and VAR00005 refers to independent variable 5 – sematic factor. The result of SPSS coefficient analysis is illustrated in table 4.

Tab. 4: Analysis result of multi-variable linear regression coefficients

model	non-standardized coefficient		standard coefficient (trial version)	t	Sig.	correlation			linear statistics	
	B	standard error				zero-order	one-sided	partial	tolerance	VIF
constant	70.782	3.717	/	19.041	.000	/	/	/	/	/
VAR00001	1.140	.563	.299	2.026	.073	.302	.560	.256	.736	1.359
VAR00002	2.934	.687	.772	4.268	.002	.879	.818	.540	.488	2.049
VAR00003	-.209	.834	-.045	-.251	.808	-.023	-.083	-.032	.506	1.977
VAR00004	.692	.742	.175	.932	.376	.586	.297	.118	.451	2.220
VAR00005	.546	.828	.107	.659	.526	-.157	.215	.083	.607	1.647

As can be deduced from Table 4, the multi-variable linear regression equation of the theme factor x_1 , detail factor x_2 , deduction factor x_3 , attitude factor x_4 and semantic factor x_5 is as equation (12) illustrates.

$$y = 1.140x_1 + 2.934x_2 - 0.209x_3 + 0.692x_4 + 0.546x_5 + 70.782 \quad (12)$$

Above results of SPSS is shown as Table 5.

Tab. 5: Multi-variable regression test result

model	sum of squares	df	mean square	F	Sig.
regression	65.559	5	13.112	10.718	.001
residual	11.010	9	1.223	R	R square
total	76.569	14		.925	0.856

Histogram of standardized regression residual is as Fig.1 illustrates.

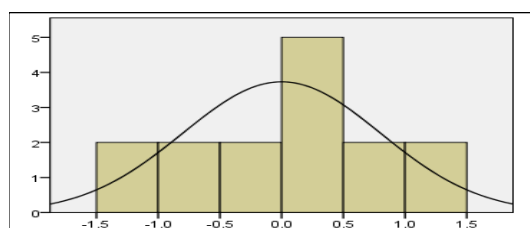


Fig.1: Histogram of standardized residuals

In the figure, horizontal ordinates stands for standardized regression residuals and vertical ordinates refer to frequencies. The mean value is $-1.44E-14$, the standard deviation is 0.802 and the number of samples is 15.

As can be concluded from above regression results and parameter characteristics, the most significant influential factor of English reading ability is detail, which is followed by theme, attitude and semantics. And deduction is a negative correlation factor, which shows that the more capable a learner makes a deduction, the more likely that the learner will misunderstand the connotation of original texts. Concerning that the most important factors are detail and theme, the improvement of students' reading ability can be made accordingly.

5. IMPROVEMENT APPROACHES OF ENGLISH READING ABILITY

As can be seen from the above analysis in Chapter 4, among five factors which influence English reading ability, detail is the most important one, followed by theme, semantics and attitude. From the characteristics of regression parameters, there is a high correlation between semantics and detail factors and between attitude and theme. In practical reading, it is familiar to us that if we are sure about the details, we will always be right about semantic questions; and if we are certain about the theme, we can always be right about attitude-related questions. When it comes to the negative correlation between deductions and English reading ability, there are two main assumptions: a strong deductive ability can lead to misunderstanding or the author's ideas; or data characteristics can contribute to the negative correlation in the regression equation. The author tends to agree with the first assumption. As long as English learners can accurately understand the details and the theme of an article, he can choose the right answer of deductive questions without too much consideration.

As can be concluded from the above analysis, attention should be paid to the theme and details of an article. There are many factors which can influence learner's understanding of the theme and details, like the knowledge of vocabulary, grammar and background knowledge, which is actually acquired through own efforts. For example, to

enlarge vocabulary, learners can recite English dictionaries or vocabulary outline. This paper mainly explores improvement methods of reading skills instead of hard indicators which can be changed through efforts. Specifically extensive and intensive reading skills are discussed, which are respectively associated with the theme and details accordingly. Next is the elaboration of skills in extensive reading and intensive reading to enhance readers' reading ability.

1) Intensive reading: it refers to the English texts learned in the class or incorporated in the teaching programme. Intensive reading aims at developing learners' language knowledge and training their basic language skills. It requires readers to read aloud and memorize some important sentence patterns or dialogues so as to develop their language sense and help them to accumulate knowledge.

2) Extensive reading: based on intensive reading, extensive reading involves a wide range of articles and is the best way to enhance learners' understanding level. New words are very common in extensive reading, but their meanings can be guessed from the context. But if learners find it hard to guess the meaning of some new words, they can deduce from the context and resort to dictionary to see if the deduction is accurate.

Extensive reading and intensive reading can be mutually transformed. In the preview of articles in an English class, extensive reading can first be adopted to extract the theme and then make deductions of new words' meanings from the text. Then intensive reading can be adopted to overcome language obstacles. In extracurricular extensive reading practices, learners can also do intensive reading so as to accumulate language skills.

To sum up, the approaches to improve English reading ability include enlarging vocabulary, firmly grasping the grammar knowledge and choosing to do extensive reading or intensive reading in an appropriate manner.

CONCLUSION AND DISCUSSION

This paper first makes a questionnaire of influential factors of English reading ability and then an analysis of exam questions of reading ability evaluation, and comes to a conclusion that theme, details, deductions and semantics are main influential factors of English reading ability.

Then the mathematical theory and application range of the multi-variable linear regression model are elaborated, laying the theoretical basis for exploring the relationship between English reading ability and five influential factors.

Next, after an English reading test is given to 100 students who major in electrical automation, sample data is collected according to the text result and data processing method. The sample data is later used to establish the multi-variable linear regression model of English reading ability and get main influential factors of reading ability, laying the basis of exploring the improvement approaches of reading ability.

Finally, on the basis of the regression model, the paper concludes the approaches to improve English reading ability through either the extensive or intensive reading in an appropriate way.

REFERENCES

- [1] Li Yanglong et al. *Journal of Shandong Foreign Languages Teaching*. 2013(2):56-51.
- [2] Bao Gui. *Foreign Languages*. 2012.28(4):63-68.
- [3] Wang Zongying et al. *Journal of Basic English Education*. 2011.13(5):7-15.
- [4] Ma Hongxiang. *English Square-Academic research*. 2013(31).
- [5] Qiu qiufen et al. *Journal of Hunan Institute of Humanities, Sciences and Technology*. 2013(2):102-105.