Validation of Lexical Frequency Profiles As a Measure of Lexical Richness in Written Discourse

Gholam Reza Abbasian[1] and Mehdi Shiri parizad[2]

Abstract: Technological developments and their utilities in various areas including education have offered great advantages for man. One of the greatest achievements in this trend has been the innovation in computer software like Lexical Frequency Profiles (LFP) and its pedagogical implications either in teaching or measurement. To take the maximum advantages, this study seeks to validate the LFP as a measure of lexical richness in written discourse of Iranian EFL Learners. 50 students majoring in English Translation participated in this study; each of them was encouraged to develop two compositions on general topics in order to establish VocabProfile indexes. To estimate the reliability of the LFP, the VocabProfile indexes of two writings were correlated, but for the validity purpose, first, a productive version of Vocabulary Levels Test (VLT) was administered and second, the students' compositions were fed into P_Lex software to elicit P_Lex index. After that, VocabProfile indexes were correlated with VLT scores and P_Lex index separately. The findings of the study revealed that students' VocabProfile indexes written on two different topics correlated significantly with each other. Because of such a significant correlation coefficients, and the LFP indexes are related to VLT active test and P_Lex index, it is conservatively safe to claim that VocabProfile indexes are to some extent reliable and valid measurement instruments but not strong enough to be used as a stand-alone measure for the assessment of lexical richness. Pedagogically speaking, the LPF is suggested as a relatively reliable and valid measure to be used along with other dependable devices in measuring lexical richness in discourses of various types.

Keywords: Lexical Frequency Profile, VocabProfile Indexes, Lexical Richness, Written Discourse, Reliability, Validity

1- Introduction

One of the most well-known, commonly-resorted, newly developed and unavoidably-interwoven phenomenon is the applications of technological developments in general and those related to computer science in particular in the fields of education ranging from teaching to testing and assessment. Then, employment of technological aids, especially those related to computers, have increasingly become a common feature of educational programs to the extent that computer–based instruction occupies a more central role in the contemporary world. Out of many opportunities offered by totality of computer technologies, feasibility of this industry for designing educational programs, virtual education, distance learning, learning exercises and more specifically assessment soft wares have opened fertile grounds for interested bodies from all flocks of life to take its advantages.

One of the main areas of employment of computer technologies is in language education, i.e., teaching and testing of language skills and components. To this end, specific devices and software have been developed and employed. For example, vocabulary assessment has been known one of the most amenable areas for involvement of computer software to the extent that it has been looked from different angles and various measures have been devised for this purpose so far [1-5]. In this line, developing various computerized measurement devices such as Vocabulary Level Test (VLT), which according to Beglar, enjoys reliability index of .95 based on Cronbach's alpha and .97 based on Rasch reliability estimate and P_Lex index, Academic Words List (AWL), Off list (OL), and more specifically the Lexical Frequency Profiles (LFP) by Laufer and Nation were initiatives and, truly, innovative developments to measure the learner's vocabulary knowledge in context [6,7].

Among these developments, the LFP is online software which enjoys highest educational utilities and advantages. It analyzes and then categorizes different samples of writings in terms of their richness of vocabulary. Amongst multiple applications, its output tells us what percentage of the words in a composition belong to the 1st thousand most frequent words (K1), 2nd thousand most frequent words (K2), AWL, and OL. Laufer and Nation (ibid) reported in their study that LFP provided similar stable results for two pieces of writing by the same individual and correlated well with other independent measures of vocabulary knowledge.

Furthermore, most language specialists have consensus that vocabulary is one of the components of language which plays a crucial role in language learning. Vocabulary as "the building block of

Received 2010/19/12 , Accepted 2011/14/02
[1] Assistant Professor, TEFL Department, Imam Ali University & Islamic Azad University, South Tehran Branch
E-mail: gabbasian@gmail.com.
[2] MA, TEFL Department, Islamic Azad University, Takestan branch, Iran

language" is regarded by learners themselves as one of the most important and difficult aspects of language learning [8]. Laufer asserts that "vocabulary correlates with holistic assessments of writing and general proficiency, and is the best single predictor of reading comprehension". Similarly, many researchers have reminded us of the role of vocabulary knowledge in reading and writing [5,9,10]. So, the research on vocabulary acquisition and assessment has been encouraged by specialists of this field [11-13]. Nevertheless, vocabulary assessment has, methodologically, faced crucial ups and downs during the history of language education. Among different aspects of vocabulary knowledge, vocabulary use in context has attracted the most attention. Read asserts that "in normal language use, words do not occur by themselves or in isolated sentences but as integrated elements of whole texts and discourse [5]. The way that we interpret vocabulary ability is significantly influenced by the context in which it occurs".

One of the effective ways for vocabulary assessment is through evaluation of the language learners' free writings, but free writing evaluation has always been a thorny task for language educators for its heavy reliance on the subjective judgment of the human raters. Earlier attempts to allay these shortcomings were focused on the development of assessment criteria that meant to increase the objectivity of free writing evaluation process. These criteria were an improvement over the traditional writing assessment but still relied on subjective judgment of raters especially when it came to the assessment of lexical sophistication. These criteria only asked the raters to devote a proportion of the total score to lexical diversity without a clear and objective definition of what lexical diversity meant, leaving the students' fate at the mercy of subjective judgment of raters. Two groups of measures for testing vocabulary knowledge in the context of use are some formulas and LFP measure.

1-1 Formulas

The formula group suggests using statistical approach for assessing lexical richness contains several formulas as presented by Linnarud**.** The most dominant formulas which Linnarud in suggests are, 1) Lexical Originality (LO) "is calculated by multiplying the number of tokens unique to one writer at 100 and then dividing it by the total number of tokens" [7].

$$LO = \frac{\text{Number of tokens unique to one writer X 100}}{\text{Total number tokens}}$$

2) Lexical Density (LD) is defined as "the number of lexical tokens multiplied at 100 divided by the total number of tokens" Linnarud, cited [7].

$$LD = \frac{\text{Number of lexical tokens X 100}}{\text{Total number tokens}}$$

3) Lexical Variation (LV) "is the type/token ratio, i.e. the ratio in per cent between the different words in the text and the total number of running words" (Linnarud, cited in [7]. In other words, the number of types multiplied at 100 is divided by the number of tokens to yield an index of LV.

$$LV = \frac{\text{Number of types X 100}}{\text{Number of tokens}}$$

4) Linnarud (ibid) multiplied the number of advanced tokens at 100 and divided the sum by the total number of lexical tokens to arrive at an index of Lexical Sophistication (LS).

$$LS = \frac{\text{Number of advanced tokens X 100}}{\text{Total number of lexical tokens}}$$

Mendelsohn cited in described his Semantic Variation (SV) measure as the number of types per topic [7]. Lexical Quality (LQ) formula is "the sum of the number of types and rare words minus twice the number of lexical errors" (Arnaud, cited in [7]. Cohen's cited in "T-unit length and error free T-unit length takes the length of a main clause together with its subordinate clauses as an indicator of lexical enrichment" [7]. However, according to Laufer and Nation each of the above mentioned formulas suffers from certain limitations which forced the scholars to search for alternative measures of testing vocabulary knowledge in the context of use, i.e., testing lexical richness of a composition through vocabulary profiler [7].

1-2 Lexical Frequency Profile (LFP)

Lexical Frequency Profile (LFP) shows the percentage of words a learner uses at different vocabulary frequency levels in his/her writing, put differently, the relative proportion of words from different frequency levels [7]. The LFP utilizes four word frequency levels: the first 1,000 most frequent words (K1), the second 1,000 most frequent words (K2), the academic words list (AWL), and the less frequent words that are not located in any of the other three lists (Off List). Laufer and Nation claim that this new measure overcomes various shortcomings of the conventional lexical statistics and it is capable to test the learners' lexical richness [7]. The LFP is based on the relative frequency of words in the language and involves simply calculating the percentage of word families in the learners' composition that belong to each of four frequency bands. Laufer and Nation rely on GSL word list for their K1 and K2 levels. As for their academic vocabulary level, they draw on Nation's and Xue and Nation's UWL (University Word List)

lists [3,14]. These lists contain 836 word families including vocabularies that is not in the first 2000 words of English, but which is frequent and wide range across a variety of written academic texts from different disciplines. But later on another AWL list which contains 570 most frequent word families of English proposed by Coxhead [15]. Less frequent words that are not included in West's GSL list and the AWL list are treated as Off-list (OL). For example, if a composition of an advanced learner consists of 300 word families and among these 300, 200 belong to the first 1,000 most frequent words, 40 to the second 1,000, 40 to the AWL, and 20 are not in any list. To calculate the LFP, we should convert these numbers into percentages out of the total of 300 word families. The LFP of the composition is approximately as follows:

Table 1 an example of LFP calculation

First 1000 words (K1)	Second 1000 words (K2)	Academic words (AWL)	Less frequent words (Off list)
67%	13%	13%	7%

The entire calculation will be done by a computer program which compares vocabulary lists against a text that we type in to see what words in the text are and are not in the lists [7]. According to Cobb the VocabProfile package consists of the program itself and three accompanying word lists [16]**.** The words in the lists that accompany the program are arranged under head words with derived forms listed below them indented by a TAB, for example [7]:

wash, washed, washing, washes

A word is defined in the program as a base form with its inflected and derived forms, i.e. a word family. Once a text has been typed into the program window, VocabProfile determines what percentage of the words in the text is covered by each of K1, K2, and the AWL lists. Cobb adapted this LFP program to the web for free online access under the name Web-VocabProfile (Web-VP) [16]. So, the LFP and VP can be used interchangeably. Laufer and Nation showed that the LFP measure of learners' texts can be compared with scores that the same learners achieve on standard vocabulary tests [7]. They (ibid) found that there is a correlation between performance on vocabulary tests and the proportions of low and high-frequency words in the free-written texts. They conclude that use of low frequency words is an indicator of richness in a learner's vocabulary, and recommend this procedure as a stable and reliable measure of lexical use in writing. Their study also shows that it is possible to obtain a reliable measure of lexical richness which can remain stable across two pieces of writing by the same learners [7].

In this regard, the LFP is supposed to give a reliable and objective approach for evaluation of lexical range of learners' writing and to save us lots of time and energy thereby. However, the reliability and validity of the instruments should be guaranteed. It is to this end that this study tries to investigate the reliability and validity of vocabulary profiling as a measure of lexical richness in written discourse in the Iranian context. To do so, one conventionally resorted method is to correlate its results with other established measures of lexical knowledge, so it is expected that elements of the system output correlate significantly not only with output's of P_Lex measure (a well established measure), but also with scores which they gain in Vocabulary Levels Test (VLT), another valid measure. To meet this purpose, this study tries to prove that the LFP outputs would remain stable in different samples of writing produced by the same subject on different topics and the instrument would measure actually what it is supposed to measure.

2- Method

2- 1 Participants

50 graduate Iranian EFL learners majoring in English Translation attended in this study in the form of two intact classes of English Translation junior students (3rd year) comprised the sample of this study. Of these, 28 were female and 22 were male with age range of 20 to 24.

2- 2 Instrumentations

2-2-1 VocabProfile indexes

The Lexical Frequency Profile (LFP) was used. It is based on the relative frequency of words in the language and involves simply calculating the percentage of word families in the learner's compositions that belong to each of four frequency bands.

2-2-2 Vocabulary Levels Test (Productive Version)

The productive version of Vocabulary Levels Test, proved .95 and .97 reliability estimates based on Gronbach's alpha and Rasch model, proposed by Laufer and Nation was also the second instrument employed [6,7]. The test has 18 items in each level and a total of 90 items. Each item in this productive version consists of a sentence with a missing word whose initial letters are provided. The letters are given to prevent the learners from producing an alternative form which might fit the context and to restrict them to producing the desired item. The participants were required to provide the missing word in each sentence.

2-2-3 P_Lex indexes

P_Lex proposed by Meara and Bell and as an exploratory tool that enables researchers to assess the lexical difficulty of texts was the third instrument [17]. It "divides the text into segments of 10 words each. Then it provides a profile showing the proportion of 10-word segments containing 0 difficult words, the proportion containing 1 difficult word, so on and so forth, up to 10" [18]. In other words, P_Lex divides the text into segments of 10 words each, and then counts the number of 'difficult' words in each segment. By difficult words it is meant "any word which is not found in a short list of high frequency words", which in practice means any word not included in the 1,000 most frequent English content words [17]. Thus, it looks at the distribution of 'difficult words' in texts, and provides a simple index i.e. 'lambda value' which indicates how likely the occurrence of these words is. When the students' essays are typed into it, it analyzes them and produces pertinent indexes.

2-3 Procedure

In order to carry out this study a set of distinct steps were taken. First, the participants were exposed to the productive version of VLT in order to measure their vocabulary knowledge based on the notion of the LFP index. Second, the participants were asked to write two argumentative essays on two distinct topics. Then, the texts were entered into Web-VocabProfile for lexical richness analysis [16]. Third, the compositions were analyzed with the Lexical Frequency Profile (LFP) measure employing the VocabProfile as an online computer program used to calculate the percentage of the lexical items in the writing samples that come from different word frequency levels (1st and 2nd 1000 most frequent words, academic word list and off list words). For each text the profiler calculated the percentage of words of the text that fall into (K1), (K2), (AWL), and off-list (OL) words group. In the last step all compositions were entered into the P_Lex software. Having carried out these stages, the researchers went through lexical item cleansing process thereby a set of deletions and corrections were applied. The last steps were devoted to the focus of the study; estimation of the reliability and validity indexes. To do the former one, the indexes of two writings were correlated with each other pair by pair (K1-K1, K2-K2, AWL1-AWL2, and OL1-OL2). As far as the latter one was concerned, the mean of the LFP indexes of each subject were correlated with P_Lex indexes and VLT scores of the same subject.

3- Results and Discussions

Each participant's score were triangulated from three perspectives; triple-calculation through: 1) Vocabulary Levels Test (VLT), 2) LFP (VP) indexes, and 3) P_Lex indexes of the compositions. All the essays were typed into 'Web-VocabProfile' for lexical richness analysis. During this process some manipulation on the essays was done. Proper nouns were deleted to prevent Web-VocabProfile identify them as Off-List words because they were not considered part of learners' vocabulary knowledge. There were also some syntactic and spelling errors in the essays. Spelling errors were corrected since they could show the semantic knowledge and active vocabulary knowledge of the learner by themselves. Descriptive statistics for LFP indexes, P_Lex indexes, and VLT scores are presented in table 2.

Moreover, to answer the research questions and mostly test the reliability and validity estimates of VocabProfile indexes a set of statistical analyses was run. To this end, t-test and correlational analyses of the various indices or measures were first run. Each set of scores was compared with each other to determine the equivalence. For this purpose, samples t-test show that the variances of two LFP indexes are equal.

Since for the reliability purposes the LFP indexes should remain stable across two writings, as such the mean of indexes for writing 1 and writing 2 should not be significantly different from each other. Given the above statistics, the mean scores of indexes in writing 1 and writing 2 prove nearly equal, but the t-test analysis could not reveal a significant difference between them. The assumption behind using t-test is that the variances are not equal and there is a significant difference between them. As (Table 3) above shows, since in all cases t-values were greater than significance levels, the hypothesis can be rejected and it can be claimed that variances of two writings are equal.

However, as the focus of the study was to measure the reliability and validity indexes of the LPF, correlational analyses as a statistics commonly used to estimate both test characteristics; reliability and validity, received prime importance.

Table 2 Descriptive statistics for LFP Indexes, P_Lex indexes, and VLT scores

N	K1	K2	AWL	OL	VLT	P_Lex index
Mean	89.5676	5.1464	3.7800	1.6000	39.6200	1.7800
Std. Error of Mean	.44450	.24814	.27085	.14842	2.04478	.06131
Median	90.4350	4.5700	3.5650	1.2850	37.0000	1.7350
Mode	92.39	4.22(a)	3.64	.96	22.00(a)	1.52
Std. Deviation	3.14309	1.75463	1.91523	1.04949	14.45879	.43350
Variance	9.87899	3.07874	3.66812	1.10142	209.05673	.18793

Table 3 Paired samples t-test

		Paired Differences					t	d	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	K 1- K 1	-.8518	2.47524	.35005	-1.5553	-.1483	-2.433	49	.019
Pair 2	K 2- K 2	.4652	1.46160	.20670	.0498	.8806	2.251	49	.029
Pair 3	AWL 1- AWL	.3282	1.38227	.19548	-.0646	.7210	1.679	49	.100
Pair 4	OL 1- OL 2	.0534	1.13667	.16075	-.2696	.3764	.832	49	.741

3-1 Estimating reliability index of LPF

As far as the reliability was concerned, the indexes of two writings were then correlated with each other pair by pair (K1-K1, K2-K2, AWL1-AWL2, and OL1-OL2). The results are presented below (table 4).

Table 4 Correlation between measurement indices

(K1, K2, AWL, & OL with the Writing Pieces: 1 & 2)

Correlations Coefficients	Pearson corr.	Sig,t-tailed	No.
K1 W1-K1 W2	.751**	.000	50
K2 W1-K2 W2	.798**	.000	50
AWL W1-AWL W2	.757**	.000	50
OL W1-OL W2	.626**	.000	50

**Correlation is significant at the 0.01 level (2-tailed),
W: writing

As the table displays, there are a set of significant correlations between these indexes reported as .75, 80, 75, and 62, respectively. After that, the indexes of writing 1 were correlated with indexes of writing 2 one by one, i.e. (K1-K1, K2-K2, AWL 1-AWL 2, and OL 1-OL 2). As tables 4 shows, there are significant correlations between all the indexes.

Since higher correlation degrees are observed between the LFP indexes of the two writings, and also there is a close correspondence between the findings of this study and Laufer and Nation's findings, it can be claimed that the LFP is a reliable measure of lexical richness [7].

Then, the data convinces the researcher and reader/s to claim and believe that LFP is a reliable measure of lexical richness as it produces nearly same results across two writings on two different topics written by the same subjects. So, high correlation between these sets of scores will be interpreted as reliability of LFP scores.

3-2 Validation process

To calculate the index validity, the mean of the LFP indexes of each subject were correlated with P_Lex indexes and VLT scores of the same subject. First, the mean the LFP indexes and VLT scores were correlated with each other. Since VLT was an already established measure of vocabulary knowledge of learners, if high levels of correlation between them were observed, the LFP would have been implied to be a valid measure of lexical richness. Then the Pearson correlation between mean LFP indexes and mean P-Lex indexes of the group were established using SPSS. Since P_Lex was an already established measure, high correlation

between the LFP and P-Lex indexes would have been interpreted as an indication of the validity of the LFP.

Table 5 Correlation between VP indexes & VLT & VP indexes & P_Lex index

	Pearson Corr		Sig		N
	VLT Test	P-Lex	VLT	P_lex	
K1	-.218*	-.633**	.050	.000	50
K2	-.177	.222	219	.120	50
AWL	.271*	.638**	.050	.000	50
OL	.437**	.513**	.00	.000	50

As to the validity issue, the assumption was that use of more K1 and K2 words in writing by a student would co-occur with his/her weak performance on VLT test. That is, the negative relationships were expected between these two indexes (K1 and K2) and VLT scores. On the other hand, it was expected that increased use of less frequent words i.e. academic words (AWL) and Off List words (OL) in writing would coincide with high VLT test scores. So, it was expected to find a significant correlation between VLT scores and AWL and OL indexes. As it was expected and displayed in table 5, the correlation between K1 index and VLT scores is negative (-0.21) which means if a student uses more K1 words in his/her writing, he/she would perform weakly in his/her VLT test. There is no then significant correlation between K2 index and VLT scores. Lack of correlation between K2 and VLT scores may be due to the fact that, as part of most frequent words, K2 items are known and used equally well by all members of the sample, thus reducing the dispersion of K2 scores and excluding the possibility of significant correlation. As for AWL and OL indexes, positive correlations between AWL and VLT scores (0.27) on the one hand, and significant positive correlations between OL and VLT scores (0.43) on the other hand, are observed. Given the close correspondence between the findings of this study and Laufer and Nation's study, although in this study moderate correlations are observed between VocabProfile indexes and VLT scores, and are not as strong as Laufer and Nation's, it can be conservatively claimed that VocabProfile indexes are valid measures of lexical knowledge at least in the context of this study. For testing the validity of VocabProfile indexes once more they were correlated with P_Lex indexes [7]. On the other hand, there are significant correlations between P_Lex indexes and VocabProfile indexes K1 (-0.63), AWL (0.63), and OL (0.51). Correlation between K1 and P_Lex index is negative. This may be due to the point that K1 words are very frequent and used more extensively by learners whose lexical knowledge is very restricted. Negative correlation between K2 index and P_Lex index may be due to the fact that, K2 words, as part of most frequent words, are known and used equally well by all members of the sample, thus reducing the dispersion of K2 scores and excluding the possibility of significant correlation. As expected, there are positive significant relationships between AWL and OL indexes and P_Lex index. Meara and Bell cited in Espinosa reported findings which are very similar to the findings of this study [18]. Negative correlation between K1 index and P_Lex index and also positive correlations between AWL and OL indexes with P_Lex index are observed in both studies.

4- Conclusions

Findings of this study can be clearly touched upon from three perspectives, given the data collected and the statistical analyses run: t-test, reliability correlation coefficients, and validity estimate. The t-test calculation which was based on the assumption that the variances are not equal and there is a significant difference between them, shows in all cases t-values are greater than significance levels, then the raised hypothesis can be rejected and it can be claimed that variances of two writings are equal. However, to prove the consistency of measurement scores, the indexes of writing 1 were correlated with indexes of writing 2 one by one, i.e. (K1-K1, K2-K2, AWL 1-AWL 2, and OL 1-OL 2). Existence of higher correlation degrees between the LFP indexes of the two writings, and also the close correspondence between the findings of this study and Laufer and Nation's study it can be claimed that the LFP is a reliable measure of lexical richness [7]. The third perspective refers to validity measure. In this regard, correlations of the LFP indexes with VLT scores and P_lex present three different pictures: Negative, Moderate, and Positive ones which can be attributed to a variety of variables, but sustain in many cases the related research findings by Laufer and Nation, Meara and Bell, and Espinosa among others. Thus, based on the findings of correlations between VocabProfile indexes of two writings and then between VocabProfile indexes and VLT scores on the one hand, and between VocabProfile indexes and P_Lex on the other hand, it could be concluded and summarized that VocabProfile indexes could be regarded as both reliable and valid measures of vocabulary knowledge [7,17,18]. To take a bit conservative stance, however, it is safe to say that VocabProfile indexes are reliable and valid to some extent but not so strongly as to be used as a stand-alone measure for the assessment of lexical richness.

References

[1] Madsen H., *Techniques in testing,* Oxford University Press, **1983**.

[2] Bachman L. and Palmer A., *Language testing in practice*, Oxford, Oxford University Press, **1997**.

[3] Nation P., *Teaching and learning vocabulary*, New York, Heinle & Heinle, **1990**.

[4] Laufer B. and Paribakht T.S., *The relationship between passive and active vocabularies: Effects of language learning context*, Language Learning, Vol.48, No.3, **1998**, pp.365-391.

[5] Read J., Assessing *vocabulary*, Cambridge, Cambridge University Press, **2000**.

[6] Beglar D., *Estimating vocabulary size*, JALT testing & evaluation, SIG Newsletter, Vol.4, No.1, April, **2000**, pp.2-4.

[7] Laufer B. and Nation P., *Vocabulary size and use: Lexical richness in L2 written production*, Applied Linguistics, Vol.16, No.3, **1995**, pp.307-322.

[8] Schmitt N., Schmitt D. and Clapham C., *Developing and exploring the behavior of two new versions of the Vocabulary Levels Test*, Language Testing,Vol.18,No.1, **2001**, pp.55-88.

[9] Saville-Troike M., *What really matters in second language learning for academic achievement?* TESOL Quarterly, Vol.18, No.2, **1984**, pp.199-219.

[10] Alderson J.C., *Assessing reading,* Cambridge, England, Cambridge University Press, **2000**.

[11] Hermann F., *Differential effects of reading and memorization of paired associates on vocabulary acquisition in adult learners of English as a second language*, TESL-EJ, Teaching English as a Second or Foreign Language, Vol.7,No.1, **2003**, pp.1-16.

[12] Read J., *Measuring the vocabulary knowledge of second language learners*, RELC Journal, Vol.19,No.2, **1988**, pp.12-25.

[13] Zareva A., *Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency*, System,Vol.33, No.4, **2005**, pp.547-562.

[14] Xue G. and Nation I.S.P., *A university word list,* Language learning and communication, Vol.3, **1984**, pp.215-229.

[15] Coxhead A., *A new academic word list*, TESOL Quarterly, Vol.34, No.2, **2000**, pp.213–238.

[16] Cobb T., *Web VP V2.5*, Available at: http://www.lextutor.ac/vp, **2009**.

[17] Meara P. and Bell., *P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts*, Prospect, A Journal of Australian TESOL, Vol.16, No.3, **2001**, pp. 5-19.

[18] Espinosa M.S., *Can P_Lex Accurately Measure Lexical Richness in the Written Production of Young Learners of EFL?* Porta Linguarum, Vol.4, **2005**, pp. 7-21.