Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 3(10) October 2013, pp. 1-8
Available online @ www.academians.org

ISSN1925-931X

Gender-Oriented Perusal of EFL Learners' Performance on Contextualized and De-Contextualized Tests

Kamal Heidari Soureshjani[1], Hamid Reza Hashemi[2]

[1]*Department of Linguistic, Shahrekord Branch, Islamic Azad University,Shahrekord, Iran*

[2]*Department of Linguistic, Sabzevar University, Iran*

Abstract
The current paper served as an attempt to shed light on two points. First, is there any difference in the performance of language testees on contextualized and de-contextualized vocabulary tests? Second, the study tried to examine gender differences in performing contextualized and de-contextualized vocabulary tests. To achieve the intended purposes, two versions of a vocabulary test, a contextualized and a de-contextualized one, were distributed among 72 Persian EFL language learners. Having run descriptive statistics and also independent t-test, the results revealed that first; language testees outperformed the contextualized vocabulary test than de-contextualized version. Second, though not significantly, females' performance was better than males' ones. Consideration of studies like the present one can be helpful in selecting the most effective strategies for assessing vocabulary aspect of language and as a result, causing a higher degree of performance in this area of language.

Keywords: Contextualized vocabulary test; De-contextualized vocabulary test; Gender Differences

I. Introduction

It is axiomatic and has also been several times mentioned by assorted researchers that assessing vocabulary is a must. They espouse this claim by arguing that words are the basic building blocks of language and it is the units of meaning from which larger structures such as sentences, paragraphs and whole texts are formed (Read, 2000; Schmitt, 2000). Considering the crucial role of vocabulary learning in second or foreign language learning contexts, one can makes aware of the importance of vocabulary teaching as well. In the past, vocabulary teaching and learning were often given little priority in second language programs, but recently there has been a renewed interest in the nature of vocabulary and its role in learning and teaching (Richards & Renandya, 2002). Furthermore, some studies have shown lexical problems of language learners indicating that lexical problems frequently interfere with communication. In fact, communication breaks down when people do not use the right words (Allen, 1983). Therefore, there is an increased interest in vocabulary as a component of every language.

Besides, as an important part of language testing, vocabulary testing has also been influenced by the development of linguistic theories. Structuralists insist that language is composed of a set

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 3(10) October 2013, pp. 1-8
Available online @ www.academians.org

Journal of Academic and Applied Studies

ISSN1925-931X

of structures and all languages can be analyzed from structures and can be divided into millions of small and independent linguistic elements. Learning of language is the acquisition of all these elements. Therefore, discrete-point test, in which every language skill is tested separately in different items, was adopted to test language competence and it was regarded as the most valid and effective way of testing.

However, since the late 1960s, with the development of transformative and generative grammar theory in the field of linguistics, discrete-point test was criticized by many linguists, for they hold that language is not the total of every small knowledge and skill, but an indivisible entity. As a result, integrative tests such as dictation and interview were adopted.

Nowadays, with the further research and development on linguistics, linguists have discovered that language competence is a multi-element entity, which are different from the traditional linguistic elements and skill theory and the mono-element theory. Both the integrative tests and discrete-point tests have advantages respectively. Integrative test is fit for the testing of language using and it is high in validity, while the discrete-point test is advantageous at reliability, plausibility and setting norm. Therefore, the two kinds of tests are both adopted in current language tests.

Usually vocabulary can be tested through the following ways: multiple-choice items, associated words, matching items, word formation test items, items involving synonyms, rearrangement items, and completion items and so on. Each form of testing has its own advantages and disadvantages. In the most public testing, multiple-choice items are preferred, for the convenience of scoring and its objectivity. No matter what way is adopted; vocabulary-testing items can be divided into context-independent items and context-independent items, when context is concerned. A context-independent vocabulary test presents words to the test-takers in isolation and requires them to select meanings for the words without reference to any linguistic context. While context-dependent vocabulary test is designed to assess the test-taker's ability to take account of contextual information in order to produce the expected response. In the context-dependent vocabulary tests, context can be sentences or passages.

Generally speaking, more and more tests begin to test vocabulary in contexts. For instance, TOFEL test vocabulary by giving only a word to be tested and four choices of single word at the beginning. When more and more criticism appeared against such form, sentence context was adopted in TOFEL vocabulary text. From 1995, TOFEL began to test vocabulary in reading materials. Other standard test, such as GMAT and IELTS, either adopt sentence context in vocabulary testing or paragraph context.

Understanding and accounting for possible different performances regarding context and gender has become a particular concern for educational researchers to ensure the fairness of all examinees. A study conducted by May (2007) found that context plays an important role in the performance of testees. He further concluded that the adoption of sentence context plays an important role in influencing the results of the tests. Therefore, context can affect test-takers' performance in a positive way. With the help of referring to sentence context, text-takers can comparatively easily get the right answer of the target words.

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 3(10) October 2013, pp. 1-8
Available online @ www.academians.org

Journal of Academic and Applied Studies

ISSN1925-931X

A number of studies conducted in various contexts have confirmed the presence of gender-related differences. In the context of second or foreign language testing, however, gender differences have only been explored to a limited degree. Ryan and Bachman (1992) studied the differential performance on two well-known international tests, the TOEFL (the Test of English as a Foreign Language) and the FCE (the First Certificate of English). Little evidence was found that male and female reacted differently at the item level to either test. Similar results were also reported when the reading comprehension test of the TOEFL was studied (Wainer & Lukhele, 1997). However, as Wainer and Lukhele (1997) suggested "it's not sufficient to merely examine each item for differential item functioning, but the test itself must be examined in its totality".

Hyde and Linn (1988) conducted a comprehensive mete-analytical study investigating gender differences in verbal ability. Among the 56 vocabulary studies included, six reported a significant difference in favor of males, while eight reported a significant difference in favor of females. Generally the meta-analysis demonstrated no significant gender difference in vocabulary, although there was significant heterogeneity in the effect size. In terms of reading comprehension, five out of twenty one studies reported a significant difference in favor of males, while ten found significant differences in favor of females. Generally, females were found to have slight advantages in reading, speaking, writing, and general verbal ability, but the differences were so small that Hyde and Linne argued the gender differences in verbal ability no longer existed.

Statistics from ACT of 2001 also showed no significant sex differences in English or reading, although the mean of females were slightly higher than those of males (Zwick, 2002). In contrast, a gender study carried out by Cole (1997) yielded completely different results. This study involved four hundred tests and millions of students. It was reported that a language advantage for females had remained unchanged compared to thirty years ago. Female superiority in verbal ability ranged from noticeable differences in writing and language use to very small differences in reading and vocabulary reasoning. At the same time, evidence also suggested that males were superior in listening vocabulary, that is, comprehension of heard vocabulary in both first and second language context. In general, despite the female advantage in general verbal ability, there seems to be no argument as to whether and to what extent gender differences exist in different types of verbal ability.

II. Methods

Participants

Subjects were 53 adult Persian-speaking students studying English at Jahad Language Institute, Iran. All of them were at upper-intermediate level and were studying Interchange 3. Among them, 30 students were males and 23 were females. Although these subjects were chosen through a survey using convenient sampling, their mean differences were compared in order to prevent any bias. Among males, 15 of them answered the contextualized vocabulary test and the

ISSN1925-931X

rest de-contextualized test. Among females, 8 of them answered contextualized and 13 answered isolated version.

Materials

Two versions of a twenty-item multiple-choice vocabulary test were designed. In the first on, items were contextualized at the sentence level and followed by four options. On the other hand, in the second test, vocabularies were presented in isolation and followed by four alternatives. It's worth noting that all the alternatives were the same.

Data collection Procedure

All the participants were asked to answer the questions within a time limit which was 20 minutes. Also, those who received the contextualized were told to use the context in order to guess any unknown words.

Data analysis

The author adopted SPSS 16.0 statistical package to compute collected data. Independent sample t-test was conducted to determine if there is a significant difference between those who received the contextualized version of the test and those who received the isolated one. Also, the same procedure was carried out in order to find out any variation between male and female performances with regard to the two versions of the test. The probability level of significance was set at 0.05.

III. Results and Analysis

As it is shown in table 1, we can see that the mean of the contextualized test is 4.6, whereas in the de-contextualized-version test, the mean is3.0. It follows that, generally, the participants who received the contextualized test achieved higher scores than those who answered the other version.

More specifically, Table 2 shows that there is a significant difference between the outcome of the first test and the second one ($p < 0.05$), it is indicated that the adoption of sentence context plays an important role in influencing the results of the tests. Therefore, we can draw the conclusion that context can affect test-takers' performance in a positive way. With the help of referring to sentence context, text-takers can comparatively easily get the right answer of the words.

Table 3 provides the information regarding the differences between males and females in terms of vocabulary knowledge. With regard to means, it is shown that females, overall, performed better than males. However, as it is illustrated in table 4 there's no significant

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 3(10) October 2013, pp. 1-8
Available online @ www.academians.org

ISSN1925-931X

difference between them (p >0.05). It seems that difference between the two means doesn't account for the differences among the testees' performance.

With regards to females, and as it is presented in table 5, we can observe a noticeable difference between the performance of those who received the contextualized version with the mean of 4.5 and those who received the isolated one with the mean of 3.9. But as the previous results for gender differences, by referring to table 6, the result is not significant. Therefore no conclusion can be drawn with respect to the result presented in table 5.

Contrary to what we observed in table, there is a much greater difference between the males group with regard to the performance on the two versions of the test. As it is shown in table 7, the difference between the mean of performance on the contextualized test and that of the isolated test is about 2.6. Also, the difference can be confirmed by looking at table 8 below which shows a significant difference between the two performances ($p < 0.05$).

VI. Conclusion

The present study was in fact an attempt to investigate first; whether contextualization or de-contextualization of a vocabulary test makes any difference in performance of language testees on these tests; and second, if there is any gender differences in the performance on these vocabulary tests.

The first finding of the study was that language testees usually outperform the contextualized version of a test than a de-contextualized test. Furthermore, the results also indicated that there was no significant difference in overall vocabulary knowledge between two groups, which seemed consistent with some previous finding: gender differences in verbal ability no longer exist (Hyde and Linn, 1988). However, it was also found out that though not significantly, female language testees' performance was higher than males'.

Besides, although, generally speaking, we can demonstrate the positive influence of sentence context on the text-takers' performance of vocabulary testing, there are still further development to be achieved for the process of experiments and more elements should be taken into account in the design of experiment. For example, an interview after the tests would be beneficial, through which the designer can have a casual talk about their performance of the text-takers after the texts. Meanwhile, the difficulty of the target words should be controlled carefully in the further experiments. In addition, the choosing of the context of the target words should be paid more attention to.

Finally, some limitations of the current study must be noted. To begin with, the size of the sample wasn't that much large enough. Therefore, generalization should be made quite cautiously. Also, the participants were all from one institute and not randomly selected. This lack of representativeness, again, may affect the generalizability of the study to a certain degree.

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 3(10) October 2013, pp. 1-8
Available online @ www.academians.org

ISSN1925-931X

Tables

TABLE 1.
DESCRIPTIVE STATISTICS

	Context	N	Mean	Std. Deviation	Std. Error Mean
Mark	contextualized	23	4.6	1.7	.3
	zero-context	30	3.0	2.1	.3

TABLE 2
INDEPENDENT T-TEST

		df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Mark	Equal variances assumed	51	.00	1.62	.56
	Equal variances not assumed	50.705	.00	1.62	.54

TABLE 3
GROUP STATISTICS

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Mark	male	30	3.50	2.28	.41
	female	23	4.13	1.98	.41

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 3(10) October 2013, pp. 1-8
Available online @ www.academians.org

Journal of Academic and Applied Studies

ISSN1925-931X

TABLE 4
INDEPENDENT T-TEST

		df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Mark	Equal variances assumed	51	.29	-.63	.59
	Equal variances not assumed	50.1	.28	-.63	.58

TABLE 5
GROUP STATISTICS

	Girlscontext	N	Mean	Std. Deviation	Std. Error Mean
Girlsmark	contextualized	8	4.5	2.1	.75
	zero-context	15	3.9	1.9	.50

TABLE 6
INDEPENDENT T-TEST

	Girlscontext	N	Mean	Std. Deviation	Std. Error Mean
Girlsmark	contextualized	8	4.5	2.1	.75
	zero-context	15	3.9	1.9	.50

TABLE 7
GROUP STATISTICS

	Boyscontext	N	Mean	Std. Deviation	Std. Error Mean
Boysmark	contextualized	15	4.8	1.6	.42
	zero-context	15	2.2	2.1	.54

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 3(10) October 2013, pp. 1-8
Available online @ www.academians.org

ISSN1925-931X

TABLE 8
INDEPENDENT T-TEST

		df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Boysmark	Equal variances assumed	3.7	28	.00	2.6
	Equal variances not assumed	3.7	26.4	.00	2.6

References

Cole, N. S. (1997). *The ETS gender study: how females and males perform in educational setting.* Princeton, NJ: Educational Testing Service.
Hyde, J., & Linn, M. (1988). Gender differences in verbal ability: a meta-analysis. Pschological Bulletin, 104(1), 53-69.
Mei, G. (2007). The influence of contexts on vocabulary testing. *US-China Education Review, 4*(7) , 41- 56.
Read, J. (2000). Assessing vocabulary. Oxford: Oxford University Press.
Read J. & Chappell C. (2001). A framework for second language vocabulary assessment. *Language Testing*. *19*(2), 57-79.
Ryan, K., & Bachman, l. (1992). Differential item functioning on two tests of EFL proficiency. *Language Testing*, *9*(1), 12-29.
Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
Wainer, H., & Lukhele, R. (1997). How reliable are TOEFLE scores? Educational and Pschological Measurment, 57(5), 741-759.
Zwick, R. (2002). *Fare game? The use of standardized admission tests in higher education.* New York: Routledge Falmer.