THE INVESTIGATION OF READING COMPREHENSION TEST AND CLOZE TEST AMONG LEARNERS WHO LEARN ENGLISH AS A FOREIGN LANGUAGE WITH RESPECT TO THEIR LANGUAGE PROFICIENCY AND GENDER

Maryam Habibian

ABSTRACT

The present study investigates the relationship between Cloze tests and performance on traditional Reading Comprehension tests among EFL learners. There were 315 subjects majoring in English including both males and females taking part in the study. They were junior and senior students from Iran and they have been chosen from the fourth semesters and above. To carry out the study, first, the subjects were randomly selected and their scores based on skills and knowledge components in university achievements tests (listening, structure, reading and writing) from the fourth semesters and above, were employed as the criterion measure to place them in three levels, namely, elementary, intermediate and advanced. Second, the sample of standard tests of Reading Comprehension consist of 54 multiple choice items and Cloze tests consist of 75 multiple choice items with every 7th word deleted were administered. The analysis of the data indicated that there is a vital relationship between EFL learners' performance on Reading Comprehension and Cloze tests and they had a better performance on Cloze tests than reading Comprehension tests. Moreover, it was found that, the gender of the subjects has significant effect on the overall performance of the subjects on the Reading Comprehension and Cloze test and females student are outperformed males. Furthermore, significant relationships between language proficiency of the subjects' and their performance on Reading Comprehension and Cloze tests have been reported.

Keywords: *Language ability, test modules (Reading Comprehension and Cloze tests), level of language proficiency, gender*

INTRODUCTION

Language aspect which is one of the human capacities in linguistic and language testing domains, becomes the central attention of the scholars and the vast number of studies tried to investigate the underlying the linguistic factors of this particular quality. They have been studying the development of language proficiency for a long period of time. The close relation has been found between two terms of teaching and testing, therefore we cannot work in either field without paying attention to the other part. Another point which has been discovered by a number of researchers is learner's language ability which is not one- dimensional but rather multi-dimensional factor.

Multi- dimensional model is recognized with other terms such as multi-factorial model. Based on this model knowledge of language components is identical to language. Individual's language ability, according to the structuralism, consisted of separate skills and components. Each part of these aspects belonged to the two linguistic categories. One category is related to four skills of reading, writing, listening, and speaking, and another category consists of lexical, syntax and phonological features (Oller, 1977). In addition, knowledge is identical with the language components and these components are quite districted from each other. Therefore, it's possible to teach them and to test them separately. In this respect, Vollmer and Sang (1983: 6)

claimed that "knowing a language meant knowing its elements and having mastered different modes of putting them to use". Oller defined the *one-dimensional model* or UCH (unitary competence Hypothesis) as "second language ability may be a more unitary factor such that once the common variance on a variety of language tasks is explained essentially no meaningful unique variance attributable to separate component will remain" (1977: 13). Generally, the assumption behind one-dimensional model is that, it is the center of the fact that have things in common between the skills.

According to Farhady, et al. (1994), testing the proficiency of language has the same changes. For instance, discrete point tests, based on the multi-dimensional model, by using different subtests are trying to achieve the ability of language. On the other hand, one-dimensional model focused on assessing language ability through integrative approach. This approach is integrating various discrete components, therefore it can be considered as a "pragmatic" use of language. Moreover, Cloze test is associated to Reading Comprehension test and a number of different ability or competence, such as strategic competence in this area will take place.

On the other hand, the process of reading is considered as an important skill. This notion is in line with Dreyer (1998) consideration which suggests that, one of the vital skills for individuals who learns a second language is reading. Moreover, the importance of the process of Reading Comprehension is mentioned by Auerbach and Paxton (1997: 238) as follows: "Reading research suggests those learners' conceptions of reading correlate with their strategies and that when they view reading as a sound - or word - centered process, they often rely on processing strategies that impede Comprehension".

Another point that needs to be mentioned here is the concept of Cloze procedure. Since decades the term Cloze procedure has been used in language teaching. Brown (2002) noted that for the first time Cloze procedure has been used for measuring reliability of written materials of the students in United States, and later on is used for different purpose in language teaching. Taylor (1953) was the first person who determined the reliability of the materials which has been used in native language of reader's. Later on, in the 1960s measuring readability of first and second language becomes a central consideration and in order to measure Reading Comprehension ability, Cloze test were used. Since 1970s, Cloze test is considered a tool in order to measure language proficiency in second language and nowadays, it stands in language testing, for instance IELTS and TOEFL.

Numerous studies (Helfeldt, Henk and Fotos, 1986) revealed that there is a close relation between individual's performance on Cloze test and Reading Comprehension test. Rye (1982), for instance, confirmed the results of Taylor's study in which Cloze test and Reading Comprehension test are 85% overlapped. High correlation between vocabulary test which was self-made and Cloze test and Reading Comprehension test has been found in study done by Koda's (1989) among 24 student at university who were learning Japanese language as a second language. In comparing Cloze test and Reading Comprehension test among fifth-grader native speakers, McKenna and Layton (1990) reported that with using Cloze test we can measure inter-sentential Comprehension. Hicks (1979) investigation among sixth grades learners revealed that there is no differences in mean scores of Cloze tests and informal reading inventory. Another study done by Davidson (1980) reported that there is no strong relationship between Cloze test and informal reading inventory during studies of science and social passages. Using Cloze test as

a valid measurement of Reading Comprehension is reported in Greene (2001) studies by using a Cloze test in order to measure macro level of Comprehension and 'global coherence' in the domain of economic. Shanahan et al. (1982) with comparing 'natural', 'scramble', and 'intruded' Cloze tests come up with the conclusion that, it is unreasonable to use a Cloze tests for global comprehension.

In sum, although various researchers have been tried to examine the relationship between Reading Comprehension and Cloze tests, the question still remains unanswered as to whether there is a close relation between the two or not because the results of each study extensively differs from another. The main interest of this article is to find out the relationship between Reading Comprehension and Cloze test among EFL learners with emphasis on their level of proficiency and gender.

STATEMENT OF THE PROBLEM

According to Alderson (1986), Vollmer and Sang (1983), the degree of Integration or Discreteness is related to the level of proficiency. This means that when learners are more advanced, they are more integrated their language proficiency and vice-versa. On the other hand, researchers in the area of testing second language suggest that performance on Cloze test is considered an indication of the readers' Comprehension ability in general and their language ability in particular (Farhady, Jafarpour, and Birjandi, 1994). Based on these ideas the present experiment is trying to investigate learner's level of proficiency in Cloze test and Reading Comprehension test with respect to the gender and level of language proficiency of those who take these tests. In sum, working on testing procedure is strongly justifiable to establish certain techniques from the results of the research, in order to improve language teaching. Further implications of the present study will be delineated in another section.

RESEARCH QUESTIONS

1. How and to what degree Reading Comprehension and Cloze test of language ability are related?
2. Is there any relationship between the learner's performance on Reading Comprehension and Cloze test and their gender?
3. What is the relationship between Reading Comprehension and Cloze tests and level of language proficiency?

RESEARCH HYPOTHESES

HO1: There is no significant relationship ($p < .05$) between EFL learners performance on Reading Comprehension and Cloze tests.

HO2: There is no significant difference ($p < .05$) between EFL learner's performance on Reading Comprehension and Cloze tests and their gender.

HO3: There is no significant difference ($p < .05$) between EFL learner's performance on Reading Comprehension and Cloze test and their level of language proficiency.

METHODOLIGY

There were 315 subjects out of almost 2000 undergraduate students form Iran majoring in English language were selected randomly. They were chosen from different university levels namely junior and senior. Based on the Mean score and the standard deviation, students who scored in half a standard deviation below the mean are considered as an elementary level and those who scored in half a standard deviation above the mean are considered as an advanced level and thoes between the two are selected as an intermediate level.

The instrumentation used in the present study consists of a number of language proficiency and achievement measures as described below.

First, student's scores based on four skills in university achievement tests (listening, grammar, reading and writing) from the fourth semesters and above, were employed as the criterion measure. After calculating different statistical procedures on the subject's scores, they were classified into three proficiency levels, namely, elementary, intermediate and advanced. **Second,** samples of standard tests of Reading Comprehension including 54 multiple choice items and Cloze test consisting of 75 multiple choice items with every 7th word deleted , taken from Oller and Perkins (1980) were given to 390 students. After removing the extreme cases 315 subject were chosen. The Reliability indices are computed to show that the instruments employed in this study is reliable. The KR-21 formula is employed to measure the reliability indices of the components of the achievement test. The components of the achievement test enjoy moderate to low reliability indices that might be due to the small number of items in each section.

Table 1: Reliability Indices

KR-21		Mean	Variance
Listening	.29	13.5954	6.007
Structure	.41	13.8945	6.932
Reading	.36	13.3591	6.709
Writing	.48	13.2872	8.173

RESULTS AND ANALYSIS

How and to what degree Reading Comprehension and Cloze tests of language ability are related?

Table 2: Pearson Correlation Reading Comprehension and Cloze

		Cloze
Reading Comprehension	Pearson Correlation	.538 (**)
	Sig. (2-tailed)	.000
	N	315

The R-observed value is .53 (Table 2). At 315 degrees of freedom, this value of R is much higher than the critical value of R, i.e. .18. Based on these results, it can be claimed that there is a statistically significant relationship between the Iranian EFL learners' performance on the Reading Comprehension and Cloze tests, thus the first null hypothesis is rejected. The finding of the first hypothesis supported Farhady (1983) ideas that, there is a high correlation between

Discrete Point and Integrative tests and they are measuring the same things and they are of the same nature.

Is there any relationship between the learner's performance on Reading Comprehensions and Cloze tests and their gender?

Table 3: Multivariate ANOVA Within-Subjects Contrast

Source	Tests	Type III Sum of Squares	df	Mean Square	F	Sig.
Tests	Linear	469.379	1	469.379	54.808	.000
tests* GENDER	Linear	245.722	1	245.722	28.692	.000
Error (tests)	Linear	2680.564	313	8.564		

As displayed in Table 3, the F-observed value for comparing the overall mean scores of the Reading Comprehension and Cloze tests is 54.80. At 1 and 313 degrees of freedom, this amount of F-value is greater than the critical value of F, i.e. 3.87. Based on these results, significant difference between the overall mean scores of the Iranian EFL learners on Reading Comprehension and Cloze tests disregarding their gender can be concluded.

Table 4: Descriptive Statistics Reading Comprehension and Cloze tests

Test2	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Reading Comprehension	7.981	.163	7.660	8.302
Cloze	9.774	.247	9.287	10.260

The subjects performed better on the Cloze Test. (Table 4)

Table 5: Multivariate ANOVA Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	46032.797	1	46032.797	2696.563	.000
GENDER	3110.803	1	3110.803	182.228	.000
Error	5343.197	313	17.071		

These results indicate that the gender of the subjects has a significant effect on the overall performance of the subjects on the Reading Comprehension and Cloze tests. (Table 5)

Table 6: Descriptive Statistics Overall Mean Scores by Gender

Gender	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Female	11.185	.207	10.779	11.591
Male	6.570	.272	6.034	7.106

These results indicate that there is a significant interaction between the students' gender and their performance on the Reading Comprehension and Cloze tests. (Table 6)

Table 7: Descriptive Statistics Interaction between Gender and Tests

Gender	Test	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Female	Reading Comprehension	9.640	.197	9.252	10.028
	Cloze	12.730	.299	12.142	13.318
Male	Reading Comprehension	6.322	.260	5.810	6.833
	Cloze	6.817	.394	6.042	7.593

As displayed in Table 7, the female students outperformed the male students on both the Reading Comprehension and Cloze tests, thus the second null-hypothesis is rejected.

What is the relationship between Reading Comprehension, Cloze tests and level of language proficiency?

Table 8: Multivariate ANOVA Within-Subjects Contrast

Source	Tests	Type III Sum of Squares	df	Mean Square	F	Sig.
Tests	Linear	723.214	1	723.214	81.854	.000
Tests * PROFICIENCY	Linear	169.638	2	84.819	9.600	.000
Error (Tests)	Linear	2756.648	312	8.835		

As displayed in Table 8, the F-observed value for comparing the overall mean scores of the Reading Comprehension and Cloze tests is 81.85. At 1 and 312 degrees of freedom, this amount of F-value is greater than the critical value of F, i.e. 3.87. These results indicated that there is a significant difference between the overall mean scores of the Iranian EFL learners on Reading Comprehension and Cloze tests disregarding their language proficiency.

Table 9: Descriptive Statistics Reading Comprehension and Cloze tests

Tests	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Reading Comprehension	8.429	.169	8.095	8.762
Cloze	10. 571	.275	10.031	11.112

As displayed in Table 9 the students performed better on the Cloze test.

Table 10: Multivariate ANOVA Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	56857.500	1	56857.500	2374.263	.000
Error	7471.600	312	23.947		

Table 11: Descriptive Statistics Overall Mean Scores by Language Proficiency

Proficiency	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Advanced	10.776	.338	10.112	11.441
Intermediate	9.919	.338	9.255	10.583
Elementary	7. 805	.338	7.140	8.469

These results indicate that the language proficiency of the subjects has a significant effect on the overall performance on the Reading Comprehension and Cloze tests (Table10, 11). It supported Alderson (1986) notions that based on the level of proficiency, language proficiency may vary.

Table 12: Descriptive Statistics Interaction between Language Proficiency and tests

Proficiency	Test	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Advanced	Reading Comprehension	9.962	.293	9.385	10.539
	Cloze	11.590	.476	10.655	12.526
Intermediate	Reading Comprehension	8.124	.293	7.547	8.701
	Cloze	11.714	.476	10.779	12.650
Elementary	Reading Comprehension	7.200	.293	6.623	7.777
	Cloze	8.410	.476	7.474	9.345

These results indicate that there is significant interaction between the students' language proficiency and their performance on the Reading Comprehension and Cloze tests. As displayed in Table 12:

a. All three proficiency levels performed better on the Cloze test. The Intermediate group had the highest mean (11.71). Second comes the Advanced level with a mean of 11.59 and finally the Elementary level had a mean of 8.41.
b. However they showed a different pattern on the Reading Comprehension test, the Advance level had the highest mean (9.96), the Intermediate group had a mean of 8.12 and finally the Elementary group showed a mean of 7.20.

Thus the third null-hypothesis is rejected.

It has been proven that the students had a better performance on Integrative test than Discrete Point test. This difference may be due to the familiarity of the subjects with Integrative tests, or to the teaching methodology they were instructed with, in their English learning periods. The reason can be the fact that in Integrative tests different linguistic elements help each other. Students have different resources in solving a given problem. For instance, if the item is tapping vocabulary knowledge, the grammatical framework of the sentence or the previous sentences can be of great help.

CONCLUSION

The results of the present study can be relevant to language testing and language teaching improvement. Based on the findings of the present study there is strong relationship between performances of the subjects in the two tests of language proficiency. It supported Brown's (1994) notions that one of the ability that influences individual performance on Cloze test is Reading Comprehension skills. The point needs to be mentioned here is because of the significant correlation between Reading Comprehension parts of language proficiency test (TOEFL) with Cloze test, therefore, Cloze test may be regarded as test of Reading Comprehension. It seems fair and sensible to agree with Carroll's (1983) arguments that,

"although there is a general language ability, at the same time, the language skills have some tendency to be developed and specialized to different degrees or rates, so that different language skills can be separately recognized and measured ". In this respect Oller (1979: 453) claimed that "every teacher in every area of the curriculum should be teaching all of the traditionally recognized language skills". Another point needs to be mentioned is that Cloze tests can be performed successfully regardless of Reading Comprehension skills. This can be implied from the words of Bachman (1990: 121) that *"Some individuals may answer items in the sequence presented, while other may not, that is, some test takers may move systematically through the test, answering the items in sequence irrespective of how easy they find them or of how much time they spend on them, while others may adopt a strategy of going through the test several times, answering only the items they are absolutely certain of the first time … In observing hundreds of individuals take Cloze tests over the year, for instance, I have noticed that these two clearly stand out"*. Consequently, different performance of the subjects is results of their different language proficiency regardless of having ability in reading strategies.

In addition, strong correlation coefficient between Discrete Point and Integrative tests in general and also in different proficiency levels have been reported. Thus, we come up with a contribution that, on the basis of the level of proficiency the correlations were determines. The findings confirmed Alderson's (1991) notion in that, the nature of language proficiency and the level of proficiency are interdependence. Confirmation of second hypothesis was another finding of this experiment, which indicated that there is a significant difference between gender performance in different tests. It seems reasonable to agree that teachers must take these differences into account because it may have an effect on the results of subject's achievement. The last important factor that is highlighted by the present study is the nature of language proficiency which is level-dependent. The rationale behind this alternative would be the proficiency level of the learner. Based on this approach, language skills and components must be treated differently at different levels of language proficiency. This theoretical framework can be considered for the pedagogical purpose in which how we should deal with different level of language proficiency. In short, we have to bear in mind that different methods, strategies and techniques are required for language teaching.

REFERENCES

Alderson, J. C. 1991. Language testing in the 1990s: How far have we come? How much further have we to go? In J. C. Anivan (Ed.), *Current developments in language testing.* (pp.1-19). Singapore: SEAMEO Regional language Center.

Alderson, J. C. 1986. Innovations in language testing. In M. Portal (Ed.), *Innovations in language testing.* Windsor: NFRE-NELSON.

Auberbach, E. R & Paxton, D. 1997. It's not the English thing: Bringing Reading Research into the ESL classroom. *TESOL QUARTERLY,* 31-2.

Bachman, L. F. 1990. *Fundamental considerations in Language testing.* Oxford: OUP.
Brown, H. D. 1994. *Principles of language learning and teaching.* 3rd edition. Englewood Cliffs, NJ. Prentice Hall Regents.

Carroll, J. B. 1983. Issues in the testing of languag

for specific purposes. In A. Hughes & D. Porter (Eds.), *Current developments in language testing*. London: Academic Press.

Davidson, E. S. 1980. An analysis of fourth and sixth grade reader performance using Cloze tests, group reading inventory, and a standardize reading test. (ERIC Document Reproduction Service NO. ED 195 984).

Dreyer, C. 1998. Improving Students' Reading Comprehension by Means of Strategy Instruction. *Journal for Language Teaching of the South African Association for Language Teaching.* Vol.32.1: 18-27.

Farhady, H. 1983. New directions for ESL proficiency testing. In J.W. Oller (Ed.), *Issues in language testing research* (pp. 253-270). U. S. A. Newbury House.

Farhady, H. 1983. The disjunctive fallacy between discrete-point and integrative tests. In J. W. Oller (Ed.), *Issue in language testing research* (PP. 311-323). U. S. A.: Newbury house.

Farhady, H. & Jafarpur, A. & Birjandi, P. 1994. *Testing language skills from theory to practice.* Tehran: SAMT.

Greene, B. B. 2001. Testing reading comprehension of theoretical discourse with cloze, *Journal of Research in Reading*, 24 (1), 82-98.

Helfeldt, J. P., Henk, W. A. & Fotos, A. 1986. A Test of Alternative Cloze Test Formats at the Six-Grade Level. *Journal of Educational Research.* Vol.79.4: 216-221.

Hicks, R. D. & Monroe, E. A. 1979. A comparison of reading achievement, current reading placement, sex, age, intelligence, informal reading inventory, and the Cloze procedure. (ERIC Document Reproduction Service NO. ED 181 438).

Koda, K. 1989. The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals* 22: 529-540.

McKenna, M. C. & Layton, K. (1990). Concurrent validity of cloze as a measure of intersentential comprehension, *Journal of Educational Psychology*, 82 (2), 372-377.

Oller, J.W. 1977. *The psychology of language and contrastive linguistics: the research and the debate*. Foreign Language Annals.

Oller, J. W 1979. *Language tests at school.* Albuquerque, University of New Mexico. London: Longman.

Oller, J. W. & Perkins, K. (Eds.) 1980. *Research in language testing.* Rowley, Mass.: Newbury House.

Rye, J. 1982. *The Cloze Procedure and the Teaching of Reading*. Oxford: Heinemann Educational Books Ltd.

Shanahan, T., Kamil, M. L. & Tobin, A. W. 1982. Cloze as a measure of intersentential comprehension, *Reading Research Quarterly*, 17 (2), 229-255.

Taylor, W. L. 1953. Cloze procedure: A new tool for measuring readability, Journalism Quarterly, 30, 413-433.

Vollmer, H. J. & Sang, F. 1983. Competing hypotheses about second language ability: a plea for caution. In J. W. Oller (ED.), *Issues in language testing research* (pp. 29-75). Rowley, Mass.: Newbury House.

Dr. Maryam Habibian
Email: habibian_2001@yahoo.com