Cloze Test and C-test Revisited: Appraising Collocational Competence on Second Language Learners' Performance

Mustapha Hajebi[1], Hosna Zarei[2]

[1]Department of Education, Bandar Abbas, Iran. E-mail: Hajebi2020@gmail.com

[2]Islamic Azad University, Central Branch, Tehran, Iran. E-mail: H.Zareii@yahoo.com

Correspondence: Mustapha Hajebi, Department of Education, Bandar Abbas, Iran. E-mail: Hajebi2020@gmail.com

Abstract

The current study tried to investigate the particular role of the text in EFL learners' performance on three types of tests, i.e. cloze test, C-test and open-ended test. This study aimed at comparing three test types of cloze test, C-test and open-ended test in measuring collocational knowledge of Iranian EFL learners. This was a quantitative research. This type of research placed more emphases on collecting data in the form of numbers. To this end, 84 Persian EFL learners were selected. They were both male and female with intermediate and advanced proficiency groups. The results showed that advanced participants in all of these three tests performed much more efficiently compared to their intermediate peers and indicated more collocational competence. The findings of this study had some implications for language learners, EFL instructors and material developers.

Keywords: cloze test, c-test, open ended test, language proficiency, collocation

1. Introduction

In a labor market that is increasingly globalized, knowledge of at least one foreign language is more relevant than ever before. So as to provide adequate language learning support, it is important to frequently evaluate learner progress on the basis of language proficiency tests that enable a fair comparison between learners (Beinborn, 2018). For language testing, redundancy can be reduced by eliminating words from a text and asking the learner to fill in the gap, also known as the cloze test. The C-test is a variant of the cloze test which contains more gaps but provides part of the solution as a hint and has been found to be a good estimate for language proficiency (Eckes and Grotjahn, 2006). Grotjahn et al. (2002) defined the C-test as a form of reduced redundancy testing and has been established as a standard entrance exam for many language centers. It usually consists of five coherent paragraphs or short texts. After an unaltered introductory sentence, every second word is transformed into a gap. When the intended number of gaps is reached (usually 20), the rest of the text is left intact. For each gap, the smaller half of the word is provided and the missing part has to be completed by the learner.

Reading is probably the most common and easiest skill of the four skills to be tested; however, testing reading has difficulties, and there are issues that anyone testing reading should be aware of. How reading ability might be assessed in a best way has interested language testing researchers for a long time. In English as a foreign/second language reading comprehension tests, include a series of related items that are based on the same reading passage (Lee, 2004). These items can be placed after a passage, as in traditional comprehension questions, multiple-choice or short-answer or embedded in the passage itself as in cloze or c-test (Klein- Braley, 1985).

Two areas of applied linguistic theories – reading and testing- come together when testers design a test of reading ability. In such cases, the test designer decides what s/he wants to test i.e. what s/he means by reading ability and finds a means of testing it. Alderson (2000), points out that there is no 'best method' for testing reading comprehension and no single test method fulfills all the purposes of tests. Discrete-point (multiple-choice) and integrative (cloze) tests are significant methods of testing comprehension.

As second language learners have great difficulty in learning and using collocations appropriately, this study is focused on measuring Iranian EFL learners' collocational competence through utilizing three test types of cloze test, C-test and open- ended test.

2. Review of Related Literature

The researcher explains more about the literature review of three test types and collocational knowledge of second language below.

2.1 Collocations and Second Language Teaching

The term *collocation* has only been given more attention after the development of electronic corpora in the 1960s and then the opportunity to observe combinational patterns in the English language increased. Since then, research on collocations has increased substantially (Ellis, 2001).

Even though the widespread use and importance of collocations in the language is recognized by many (Kjellmer 1984; Nation, 2001; Stubbs 1995), few attempts have been made to integrate the teaching of collocations in the English learning curriculum. Some exceptions are Michael Lewis, who has developed the Lexical Approach (1993), and McCarthy et al. (2006), authors of the *Touchstone* material that uses corpus information and includes collocations in the vocabulary work. Another important contribution is from Nesselhauf (2005), who studies collocations in a learner corpus and based on the results provides suggestions on how to select collocations for teaching**.**

The role collocations play in second language teaching is integrally related with a concept which was introduced in the 1970s. Based on this principle language is learned in a series of pre-fabricated blocks or chunks defined by Lewis (1994) as unanalyzed wholes. These chunks are said to be the basic data which enable learners to identify patterns in a language. Previously, this function was given to grammar which was considered as a necessary condition for successful communication.

According to Nattinger (1980), language production is based on piecing together ready-made units which are appropriate for a particular situation. Comprehension of such units depends on knowing the patterns to predict in different contexts. That is why a learner should be instructed to know in which cases those units can be combined. Szulc (1984) claimed that acquisition of collocations in a second language is fundamental when a learner tries to gain language competence. Collocational errors occur even for proficient learners.

Carter (1987) perceives collocations as crucial factors of lexical coherence and emphasizes the need for teaching collocation at all levels of language proficiency. In Celce-Murcia's view, familiarity with the way words combine is a basic, native-like aspect of learning and using vocabulary. This knowledge helps learners to encode and decode the language, it follows the rule that it is easier to understand a message if its elements are highly predictable. Lewis (1997) in his 'lexical approach' theory emphasized that fluency in a foreign language is conditioned by the acquisition of a number of pre-fabricated chunks. He also regards collocation as a central feature of a language production. Therefore, students' attention should be fully directed to it. Lewis (2000) also claims that the number of collocations which are understood as word combinations is greater than the number of all words because the same words can occur in various collocations. That is why collocations create enormous problems even for those who are the most proficient in English. In addition, collocational competence enables students to produce texts which are grammatically correct and authentic, it means that it is this collocation which a native speaker uses in this specific situation. Lewis comes to a conclusion that it is possible to achieve proficiency level in mastering the syntax of a second language through expanding a range of memorized whole word combinations. Collocations also strengthen the generation of a learner's lexicon which is especially true for nouns. Every time nouns are presented in class they should be accompanied by a range of adjectives and verbs which are their collocates. A learner will not use a noun in a proper context unless he knows which words co-occur with it. As a result, knowing a word cannot be limited to simply knowing its meaning; it is also crucial to know its collocational range. Hill (1999) even raises the term 'collocational competence'. Learners have considerable difficulty developing collocational competence unless they are able to collocate words successfully. If there are no ready-made chunks in their lexical corpus, they have to generate novel ones on the basis of grammar rules. This leads to numerous mistakes. A wide range of meaningful chunks and collocations in the learner's mental lexicon makes it possible to find the right word quickly. It also facilitates and accelerates the communication process.

2.2 C-Tests vs Cloze Tests

The main problem with cloze tests is the ambiguity of the solution. Unless function words are deleted, the gap allows many alternative solutions such as synonyms and hypernyms, but also entirely different words that change the meaning of the text but also fit the context. Language teachers have proposed two ways of dealing with this ambiguity: the application of relaxed scoring schemes and the use of distractors. In relaxed scoring, teachers accept all tolerable candidates for a gap and not only the intended solution as in exact scoring. Unfortunately, this scoring method turned out to be quite subjective and time-consuming as it is not possible to anticipate all tolerable solutions (Raatz and Klein-Braley, 2002). The use of distractors circumvents this open solution space by providing a closed set of candidates from which the solution needs to be picked. Several approaches have been proposed for automatic distractor selection

(Sakaguchi et al., 2013; Zesch and Melamud, 2014) to make sure that the distractors are not too hard nor too easy and are not a valid solution themselves. However, the presence of the correct solution in the distractor set enables the option of random guessing leading to biased results. In order to overcome this and other weaknesses of the cloze test, Klein-Braley and Raatz (1984) propose the C-test as a more stable alternative. Thorough analyses following the principles of test theory indicate advantages of the C-test over the cloze test regarding empirical validity, reliability, and correlation with other language tests (Babaii and Ansary, 2001; Klein-Braley, 1997; Jafarpur, 1995). However, Jakschik et al. (2010) transform the C-test into a true recognition test by providing multiple choice options and find that this variant is significantly easier than open C-test gaps. This indicates that C-test solving requires both, receptive and productive skills, and we reflect this in our feature choice.

2.3 Purpose of the Study

This study aimed at comparing three test types of cloze test C-test and open-ended test in measuring collocational knowledge of Iranian EFL learners. There are many problems for EFL learners when they intend to learn collocations of the second language because native speakers know these items through their extensive exposure in their first language as they have many opportunities to hear and make use of them while second language learners do not enjoy such opportunities to be exposed to and know how to use them in different communicative situations. Therefore, second language learners may face serious problems in using collocations in their correct order as collocations have their specific co-occurrence of words that need to be learned together. Generally, this study is an attempt at shedding more light on the nature of lexical and grammatical collocations and tries to find Persian EFL learners' weak points and difficulties in acquiring this very important part of a second language and may propose new ways in learning them.

2.4 Research Questions

The present study attempted to address the following questions regarding the efficiency of C-test and cloze test in measuring Persian EFL learners' collocational competence:

1. Is C-test more excessive than cloze test and open-ended test in measuring EFL learners' lexical and grammatical collocations knowledge?

2. Is cloze test capable of measuring EFL learners' lexical and grammatical collocations knowledge?

3. Is C-test capable of measuring EFL learners' lexical and grammatical collocations knowledge?

4. Is open-ended test capable of measuring EFL learners' lexical and grammatical collocations knowledge?

5. Is there any difference between intermediate and advanced learners in terms of their performance on cloze test, C-test and open ended test items?

3. Methodology of the Research

First, the design and participants discussed. In the next step, the applied and measurement instruments mentioned. Third, it focuses on the steps and procedures taking from the very beginning towards the end of this work. And finally, it provides the view about how the data analyzed.

3.1 Research Design

This study used quantitative methods in utilizing cloze test, C-test and open ended test in measuring collocational competence of Iranian EFL learners.

3.2 Participants

The participants of the study were 84 Persian EFL learners of English. They were male and female students and their ages were between18 to 42 for intermediate and advanced group. The criteria to select participants of the study included: (a) previous academic L2 learning background. (b) An Oxford Placement Test was conducted to ensure the least difference among reading proficiency level of students in intermediate and advanced levels.

3.3 Procedure

The researcher explained to the students the details of the study for fifteen minutes. After explaining to the students, the researcher asked them whether they want to participate in the study or not.

When the tests distributed among the learners, some instructions passed them to respond test items more effectively .For example, when C-tests are distributed among the learners, they faced difficulty in dealing with so many blanks which existed in this test and in making meaningful sentences out of these many blanks. So, the students were asked to concentrate only on filling those fifty collocational items which were written in bold prints and they were made aware that the dashes after the first half of the word showed the number of letters complete the words.

First, to indicate the participants' proficiency level, the standardized Oxford Placement Test was administered to all

participants. The original test includes 200 test items from which we chose the second part including 50 items for two main reasons: 1. the second part looked more like a cloze test. 2. With only 50 items, we would not have problems of time and participants would not get bored and tired.

Second, in order to make a comparison between performances of students on cloze test, 14 intermediate and 14 advanced students took this test to investigate which proficiency level performed higher on the test and enjoyed more collocational competence.

Third, 14 intermediate and 14 advanced learners participated on C-test. Advanced learners were used to provide a fair comparison between collocational competence of intermediate and advanced learners and their performance on this test.

Fourth, 14 intermediate and another 14 advanced learners participated in the open-ended test to provide comparison among collocational competence of both proficiency groups.

The scoring method for cloze test, C-test and open ended test was exact word scoring method. This method was objective so that obtaining scores were reliable. The multiple-cloze test was scored like usual multiple-choice items and each item had one point. The participants were expected to write and guess the exact word using in the original passage. This type of scoring method was used because it is both easier and more reliable than the other methods. Being objective, this method of scoring provided equal opportunities for participants' performances to be compared with each other. In the C-test, each item had one point, too. The participants asked to read the text carefully and provided the needed letters which were originally deleted from the text. In this study, it was decided to tolerate minor spelling problems in C-test which had not changed the words' meaning and parts of speech. If these two problems happened, the written word would not gain mark.

4. Results

In the current study, learning collocation was chosen to study three types of test e.g, cloze test, C-test and open-ended test that are effective in measuring the collocational competence of EFL learners, and which of these three tests are more effective in measuring the collocational competence of learners.

Two-way ANOVA Regarding the Effect of Test type and Proficiency Level on Collocational Scores

Table 1. Two Way ANOVA Results for Test Types and Proficiency Level

Descriptive Statistics				
Dependent Variable: Scores				
Test type	Proficiency level	Mean	Std. Deviation	N
Cloze	Intermediate	17.84	03.033	42
	Advanced	24.49	04.095	42
	Total	21.16	03.564	84
C-test	Intermediate	16.61	04.617	42
	Advanced	24.52	03.786	42
	Total	20.56	04.201	84
open ended test	Intermediate	09.94	01.097	42
	Advanced	12.99	01.482	42
	Total	11.46	01.289	84
Total	Intermediate	14.72	04.757	42
	Advanced	20.63	06.348	42
	Total	17.67	05.552	84

Tests of between Subjects Effect showing the Results of Two way ANOVA Regarding the Effect of Test type and Proficiency level on Collocational Scores

Table 2. Two Way ANOVA Results between-Subjects Effects.

Tests between Subjects Effects

Dependent Variable: Scores

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	76413.714[a]	5	15282.743	220.597	.000	.727
Intercept	820528.800	1	820528.800	1.184E4	.000	.966
Test type	50767.600	2	25383.800	366.400	.000	.639
Proficiency level	22880.952	1	22880.952	330.273	.000	.444
Test type * Proficiency level	2765.162	2	1382.581	19.957	.000	.088
Error	28681.486	78	69.279			
Total	925624.000	84				
Corrected Total	105095.200	83				

First, the main effects of the independent variables i.e. test type and proficiency level is checked. Test type row indicates a significance value of 0.001, which shows that the test types can affect collocation scores. The proficiency level has also a significance value of 0.001 which shows that proficiency level affects collocation scores. The effect size of the test type variable, as shown under partial Eta Squared column, is .639 indicating a large effect size.

Given the significance of the test type variable, it should become clear which test type is significantly different from the other tests.

Multiple Comparison Table Providing a Comparison between the Results of Different Test types

Table 3. Multiple Comparison Results of Different Test types

Multiple Comparisons

Scores Schefe

(I) Test type	(J) Test type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Cloze	C-test	1.36	.995	.395	-1.09	3.80
	open ended test	23.94[*]	.995	.000	21.50	26.39
C-test	Cloze	-1.36	.995	.395	-3.80	1.09
	open ended test	22.59[*]	.995	.000	20.14	25.03
open ended test	Cloze	-23.94[*]	.995	.000	-26.39	-21.50
	C-test	-22.59[*]	.995	.000	-25.03	-20.14

The results show that cloze test and open ended test, as two test types are significantly different from each other with a p value of .001. But cloze test and C-test are not significantly different from each other with a significance value of .395. C-test and open ended test show a significance value of .001 which shows that they are significantly different from each other. The open-ended test shows a significance value of .001 in comparison with the cloze test and C-test which indicates that it is significantly different from both of these tests.

In order to investigate the relationship between two independent variables, a means plot is conducted.

Descriptive Statistics for Test type and Collocational Categories

Table 4. MANOVA Results for Test Type and Collocational Categories

	Test type	Mean	Std. Deviation	N
Noun collocation	cloze test	10.51	04.116	28
	c-test	10.42	03.604	28
	open-ended test	07.88	02.711	28
	Total	09.60	03.477	84
Verb collocation	cloze test	09.54	03.397	28
	c-test	09.31	03.903	28
	open-ended test	05.57	02.690	28
	Total	08.14	03.330	84
Adjective collocation	cloze test	09.17	03.218	28
	c-test	08.48	03.852	28
	open-ended test	03.44	01.089	28
	Total	07.03	02.986	84
Adverb collocation	cloze test	11.87	04.001	28
	c-test	10.15	03.861	28
	open-ended test	04.57	02.011	28
	Total	08.86	03.291	84
Preposition collocation	cloze test	11.55	03.174	28
	c-test	12.94	04.119	28
	open-ended test	11.85	02.940	28
	Total	12.11	03.411	84

Descriptive Statistics for Proficiency level and Collocational Categories

Table 5. MANOVA Results for Proficiency level and Collocational Categories

Descriptive Statistics				
	Proficiency level	Mean	Std. Deviation	N
Noun collocation	intermediate	16.05	06.387	42
	advanced	22.38	07.009	42
	Total	19.44	06.738	84
Verb collocation	intermediate	13.02	06.466	42
	advanced	16.02	07.361	42
	Total	14.07	06.913	84
Adjective collocation	intermediate	11.10	06.152	42
	advanced	17.02	08.573	42
	Total	14.06	07.362	84
Adverb collocation	intermediate	14.85	07.335	42
	advanced	20.60	10.034	42
	Total	17.72	08.684	84
Preposition collocation	intermediate	18.26	07.725	42
	advanced	23.71	08.272	42
	Total	20.98	07.998	84

The importance of the impact of proficiency level on collocational categories' scores can be evaluated using the effect size statistic provided in the final column. Partial Eta Squared represents the proportion of the variance in the dependent variables that can be explained by the independent variable which is proficiency level.

Preliminary assumption testing was conducted to check for linearity, homogeneity of variance covariance matrices, and multi collinearity, with no serious violations. There was a statistically significant difference between intermediate and advanced learners on the combined dependent variables. When the results for the dependent variables were considered separately, all of the dependent variables reached statistical significance. An inspection of the mean scores indicated that

advanced learners reported higher collocational score in noun collocation than intermediate learners, advanced learners reported a better score in verb collocation than intermediate learners, advanced learners in adjective collocation showed these scores and advanced learners in adverb collocation had such scores.

5. Discussion

This study tried to measure the Iranian EFL learners' collocational competence. The five collocational categories of noun, verb, adjective, adverb and preposition are chosen to be further investigated in EFL learners' performances. Each of these three test types had fifty items and each of these tests had ten noun, verb, adjective, adverb and prepositional collocations.

The findings of this study led to the following conclusions. First, cloze test measured noun, verb, adjective and adverb collocations better than C-test and open ended tests which may be related to the point of the cloze test provides four choices for each item and the learners are supposed to choose among them. Second, the C-test has been more effective in measuring prepositional collocations of the participants as half of the prepositional words are provided and this serves as a big clue for language learners. Third, in case of the open-ended test, as participants had no clue and no choice was provided for them, they performed poorly on this test through all the five collocational categories in comparison to the other two tests. Finally, it was shown that participants' performance on all the three test types through all the collocational categories were significant. (p=.001).

These findings are in line with some of the research projects which have been conducted in this regard. These findings support this view point that different proficiency levels influence the learners' performances on lexical and collocational categories and higher proficiency levels learners can perform far more effectively on these categories. It should also be mentioned that the more students are in command of collocations, the more they show reading proficiency as different collocational categories enable language learners to learn more chunks of words and word clusters. Based on the results from the conducted studies, it has been shown that the language learners learn vocabularies and phrases in groups more effectively than when they are isolated from context.

This study was conducted to fill a gap in the existing literature in measuring the effectiveness of the cloze test, C-test and open ended test in assessing collocational competence of Iranian EFL learners. This study tried to utilize the lexical and grammatical collocations framework in assessing Iranian EFL learners' collocational competence. On the whole, the results of the conducted analyses suggested that C-test was not superior to cloze test and open ended test in assessing collocational competence of EFL learners. In addition, the analyses confirmed that proficiency level is an important and determining factor and influences participants' performances on different test types. Therefore, the learners from higher proficiency levels performed more effectively on different test types compared with their lower proficiency peers.

Reference

Esmat, B., & Hasan, A. (2001). The C-test: a valid operationalization of reduced redundancy principle? *System, 29*(2), 209–219. https://doi.org/10.1016/S0346-251X(01)00012-4

Jack, M., & Hyeju, J. (2012). Generating Diagnostic Multiple Choice Comprehension Cloze Questions. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pages 136–146. Association for Computational Linguistics. https://doi.org/10.3115/v1/P14-5011

Johannes, D., Oliver, F., Iryna, G., & Torsten, Z. (2014). DKPro TC: A Java-based Framework for Supervised Learning Experiments on Textual Data. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 61–66. Association for Computational Linguistics.

Jonathan, C. B., Gwen, A. F., & Maxine, E. (2005). Automatic Question Generation for Vocabulary Assessment. In *HLT '05: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 819–826, Morristown, NJ, USA. Association for Computational Linguistics.

Lisa, B., Torsten, Z., & Iryna, G. (2014). Readability for foreign language learning: The importance of cognates. *International Journal of Applied Linguistics*, pages 136–162. https://doi.org/10.1075/itl.165.2.02bei

Lisa, B., Torsten, Z., & Iryna, G. (2018). Predicting the difficulty of language proficiency tests Asian federation of natural language processing. http://www.ukp.tu-darmstadt.de

Manish, A. & Prashanth, M. (2011). Automatic Gap-fill Question Generation from Text Books. *Proceedings of the 6th Workshop on Innovative Use of NLP for Building, Educational Applications*, pages 56–64.

Mark, H., Eibe, F., Geoffrey, H., Bernhard, P., Peter, R., & Ian, H. W. (2009). The WEKA data mining software: an update, *11*(1).

Richard, E. de C., & Iryna, G. (2014). A broad-coverage collection of portable NLP components for building shareable

analysis pipelines. In *Proceedings of the Workshop on Open Infrastructures and Analysis Frameworks for HLT (OIAF4HLT) at COLING 2014*, pages 1–11, August. https://doi.org/10.3115/v1/W14-5201

Thomas, E. (2011). Item banking for C-tests: A polytomous Rasch modeling approach. *Psychological Test and Assessment Modeling*, *53*(4), 414–439.

Thorsten, B., & Alex, F. (2006). Web 1T 5-gram corpus version 1.1. *Linguistic Data Consortium*.