# On the Effect of Text Authenticity & Genre on EFL Learners' Performance  in C-Tests

**Mahmood Reza Atai***

Assistant Professor of TEFL, Tarbiat Moallem University, I.R. Iran

**Masoud Soleimany****

MA in TEFL, Tarbiat Moallem University, I.R. Iran

**Abstract**

The present study was conducted to investigate the effect of genre and text authenticity on the performance of EFL students in C-tests. Bachman's model of language ability and test method facets provided the theoretical framework for this research. The focus of the study was on the nature of input and expected response by considering genre and text authenticity in C-tests. The participants were 120 senior students majoring in English language translation who were screened from a larger pool of subjects based on their performances on a retired TOEFL test (Sharpe, 2001). The instrumentation included four tailored C-tests based on two different genres (political vs. literary) and two kinds of texts, i.e. authentic and inauthentic texts. The C-tests were administered to four homogenized groups each comprising 30 subjects. Data were analyzed through two-way analysis of variance and post hoc Scheffe tests. The results showed a significant effect of genre and text authenticity on the C-test performances of the participants. However, the interaction between genre and text authenticity did not prove to be significant .The findings provided data to support two aspects of Bachman's model ; the effect of the nature of input and the nature of the expected response on C-tests. The findings make it clear that genre and text authenticity can be considered as test method facets or 'specific aspects of test method' (Bachman, 1995 : 115) in the C-tests.

**Key Words:** C-Test, Method facet, Text authenticity, Political genre, Literary genre.

* Tel: 021-88304896, Fax: 021-88304896, E-mail: mreatai@yahoo.com

** Tel: 021- 88304896, Fax: 021-88304896, E-mail: __

**Introduction & Background**

According to Klein-Braley (1997), C-test consists of four to six short authentic texts, each complete as a sense unit in itself. Since C-tests consist of different short texts, the selection of these short passages or input is of the great importance and has been the subject of many investigations in order to make these tests more appropriate and least biased (McBeath, 1990). When C-test was first developed, it was meant to account for the specific weaknesses that had been discovered in the cloze format. It was claimed to be a justified replacement for the cloze test from both theoretical and psychometric perspectives (Klein-Braley and Raatz, 1984, p.138). It is supposed that the inclusion of several short passages will lead to the superiority of a C-test over cloze involving only one text (Klein-Braley, 1984). But as Siong (2004:5) argues, for administering an English test, the starting point should be an examination of its 'text types' and 'genre'.

Considering Bachman's (1990, 1995) model of language ability, *'the nature of input'* which can affect test performance in language testing is associated with *'test organization'* and *'method characteristics'*. As Bachman (1995) asserts, if the language of the input in a given test is characteristic of a 'genre' that is unfamiliar to the test takers, it can be hypothesized that tasks that depend on the interpretation of the input would be relatively difficult and the genre may affect the test takers' performance because "unfamiliarity with the characteristics of a given genre may make the expected response more difficult" (Bachman, 1995:139). Along the same line, Bachman (1995) argues that one essential factor in the framework of the test method facets is 'the nature of input' which can potentially affect test performance.

Following Bachman's model (1990; 1995), Kobayashi (2002, 2004) carried out a series of studies on *'test organization'* and *'response format'* and found supporting evidence for method facets in reading comprehension tests and cloze. Kobayashi (2004) argues that there is a systematic relationship between the students' test performance and the effects of '*the nature of the input*' and the '*nature of the expected response*', while the nature of input in testing refers to '*test organization*' and '*test method characteristics*'. Similarly, Koberl and Sigott (1994) reiterated the

importance of *'text type'* in the format of reading comprehension tests. Also, an array of studies on text characteristics and text organization of reading comprehension tests conducted by Paltridge (1996) fountas and Pinell (2001) reconfirm that the students' performance in reading comprehension tests could be sensitive to the different *'genres'* and *'text types'* used in the tests. Moreover, different *' text types'* not only have significant effects on reading comprehension, but they also interact with response format (Kobayashi ,2002 ) .

On the other hand,  some researchers ( e.g. Klein-Braley 1994 ;  Sigott, 1995 ; Nakamura, 1998 ;   Woolls & King, 1999  ; Brantmeier,2005) indicated the significant differences between authentic, inauthentic and quasiauthentic texts. As Woolls and King (1999:28) argued, "there are different inter-relationships between the source language and the parallel-translated texts of it". Additionally, different paragraph and sentence boundaries, different alignment in features of layout and paragraphing and considerable differences in text encoding were also observed by Woolls and King (1999) between the authentic and inauthentic texts. Likewise, some empirical studies on text authenticity and translation conducted by Sigott (1995),  and Laviosa-Braithwaite (1997) indicate that there are differences between authentic texts and translated texts in terms of relative frequencies of certain features compared with their counterpart corpora in terms of lexical, grammatical and even pragmatic elements. As Sigott (1995:47) maintains, "translation creates different texts". These differences may indicate the concept of 'authentic stimulus' in language testing that is emphasized by Widdowson (1983:30) as "attested instances of language".

Regarding the role of genre of the texts in the language tests, Paltridge (2001:57) argues that in language testing the genre plays an undeniable role since "students need an understanding of the sociocultural context and setting of a genre, as well as the relationship between themselves and the audience(s) of the texts". As Nunan (1991:73) asserts, "one of the claims of genre theory is that language fulfills certain functions and that these functions will largely determine the structure of the text and the language it contains." Therefore, text organization is related to the

'genre' since "genre describes text's structure" (Paltridge, 1996: 239); besides, according to Paltridge (2001), the study of genre is important in English for Specific Purposes (ESP) and in the entire field of English language teaching and testing.

Following Bachman's model for test method facets (1990, 1995) and considering the current empirical findings (e.g. Paltridge, 2001; Kobayashi, 2002) on the effect of genre and text type on test performance, it could be assumed that C-test performance might be affected by genre and text type too. Accordingly, the focus of present study was on the nature of input and the expected response by manipulating the short passages of C-tests as the main variable of the study. More specifically, the study was designed to find out whether 'genre' and 'text authenticity' were facets or "the specific aspects of test method" (Bachman, 1995:115) in the C-tests. In this regard, the effects of authentic and inauthentic texts with different genres (political vs. literary genre) on the performance of EFL students in the C-tests were investigated. In addition to this primary focus, the study also looked into interaction between genres (political versus literary genre) and text types (authentic versus inauthentic texts) on the performance of EFL students in C-tests. To this end, the following questions were raised:

1. Is there any significant relationship between the performance of EFL students on C-tests and different genres (political vs. literary), while the level of language proficiency is controlled?

2. Is there any significant relationship between the performance of EFL students on C-tests and text authenticity, while the level of language proficiency is controlled?

3. Is there any interaction between genres (political vs. literary) and text types (authentic vs. inauthentic) in the performance of EFL students on C-tests, while the level of language proficiency is controlled?

To explore the above questions, the following nullhypotheses were formulated and probed :

H01. There is no significant relationship between the performance of EFL

students on C-tests and different genres (political vs. literary), while the level of language proficiency is controlled?

H02. There is no significant relationship between the performance of EFL students on C-tests and text authenticity, while the level of language proficiency is controlled?

H03. There is no interaction between genres (political vs. literary) and text types (authentic vs. inauthentic) in the performance of EFL students on C-tests, while the level of language proficiency is controlled?

## Method

### Subjects

The subjects of the study were 120 EFL male and female senior students (21-30 years old) in Islamic Azad University (IAU) who were majoring in English language translation. The rationale behind selecting senior students was that these students were familiar with a wide range of genres including political and literary genres. In order to ascertain homogeneity of subjects in terms of their general English proficiency at the outset of the study, a retired TOFEL test (Sharpe, 2001) was administered to 153 participants. Considering the normal curve and based on the histogram of the participants' TOEFL scores, those having abnormally too high or too low scores were excluded. Finally, 130 participants remained. Through random elimination of 10 participants, 120 students were chosen out of 130 participants and assigned randomly to four groups, i.e. groups 1, 2, 3 and 4 (each of them consisted of 30 participants). In order to test the homogeneity of the aforementioned four groups, a one-way ANOVA was run the results of which attested that there was no significant difference among the groups in terms of language proficiency.

### Instrumentation

The instrumentation included a retired Test of English as a Foreign Language (TOEFL) and four C-tests constructed by the researchers for the purpose of this

study. The TOEFL (Sharpe, 2001) was used as a criterion to determine the examinees' level of language proficiency and to estimate the reliability and validity of the newly developed C-tests as well.

**Construction of C-Tests**

In order to probe the research questions of the study, four tailored C-Tests were constructed:

a) Tailored C-Tests 1 and 2: Based on Political Genre (PG), Authentic Text (AT) and Inauthentic Text (IT), PGAT (1) and PGIT (2)

b) Tailored C-Tests 3 and 4: Based on Literary Genre (LG), Authentic Text (AT) and Inauthentic Text (IT), LGAT (3) and LGIT (4)

Each of the above C-Tests comprised four short passages.

**Selection of genres and subgenres**

From all the genres mentioned in the British component of the International Corpus of English (ICE) and the British National Corpus (BNC), in this study the option was narrowed down to the political and literary genres. These two genres were among the most frequent ones in the participants' curriculum. Commentary was chosen as the subgenre since it is available in both literary and political genre and is typically common in translation courses.

**Selection of Authentic and Inauthentic texts**

In order to explore the probable effect of text authenticity on EFL learners' C-test performance, we took text type as one of the variables in this study. Inauthentic texts were considered as those texts translated from Persian into English. To avoid the undesired cultural impact of authentic texts, sometimes the test developers prefer to use translated texts for their C-test development. Therefore, we selected a variety of political and literary commentaries of authentic and inauthentic texts

**Selection of passages for C-Tests**

Each of the four C-tests included four passages that were randomly selected from a large pool of political and literary sources. The readabilities of the texts were estimated through Flesch Reading Ease Score (FRES) index and Kincaid Formula. The participants were at upper-intermediate of language proficiency, so the average readability of the passages according to Greenfield (2004) should be around 53.24 on the Flesch Reading Ease Score (FRES) index of readability. This average readability with one standard deviation below and above it, ranging from 43.24 to 63.24 was applied to randomly chosen passages to determine their readability. The estimated average Flesch reading ease and Kincaid indexes for the texts were 47.3 and 10.7, respectively.

Given the guidelines of Klein-Braley (1997), four C-tests, each of them comprising four subtests and each subtest with 25 mutilations were constructed. Each subtest comprised a complete sentence as a contextual clue, followed by twenty-five deletions of every other second half of alternate words beginning with the second word of the second sentence.

**Pilot study**

A pilot study was designed to check the characteristics of the items and the psychometric features of the C-tests. The pilot group consisted of 30 participants of similar characteristics to the target group. Item discrimination and item facility indices were also estimated and after improving or changing defective items, the scores of the newly developed tests were correlated with those of the proficiency test for estimating their criterion validity. The Pearson product-moment correlation coefficients between the C-tests fell within the high level (i.e. all beyond .50). Estimated internal consistency measures (KR-21) of individual subtests revealed that all of them enjoyed acceptable reliability values The final versions of tailored C-tests consisted of only those subtests which enjoyed an IF index above 0.32. Furthermore, by using the individual texts as superitems in Cronbach alpha formula (Klein-Braley, 1997), C-tests reliability was again estimated. The estimated internal

consistency measures (KR-21) and the results of Cronbach alpha formula revealed that all of the C-tests enjoyed acceptable reliability values, i.e. $0.65 < alpha < 0.9$ and $0.68 < KR-21 < 0.92$.

**Procedure**

The four C-tests , i.e. PGAT, PGIT, LGAT, and LGIT, were administered to groups 1,2,3, and 4 , respectively. Similar test instructions were given to all groups and the participants were asked to complete the tests in 40 minutes. Having administered the C-tests to the corresponding groups of testees, the researchers corrected the papers. To probe the research null hypotheses, two-way ANOVA and post-Hoc Scheffe tests were carried out.

**Results & Discussion**

Initially, descriptive statistics for the performances of the subjects on the four C-tests involving different genres and text types (i.e. PGAT, PGIT, LGAT and LGIT) were computed. A two-way ANOVA was run in order to probe the null hypotheses of the research and a Post-Hoc Scheffe test was carried out in order to locate the exact differences in the performances of the target groups. Table 1 summarizes basic descriptive statistics including mean, standard deviation, etc. for all the C-tests used in the study.

**Table 1: Descriptive Statistics for C-Tests 1, 2, 3 and 4**

| C-Tests | N | Range | Min. | Max. | Mean | Std. Dev. | Variance |
|---------|----|-------|------|------|-------|-----------|----------|
| PGAT 1 | 30 | 48 | 41 | 89 | 61.26 | 15.21 | 231.44 |
| PG IT 2 | 30 | 31 | 55 | 86 | 72.26 | 9.45 | 89.37 |
| LGAT 3 | 30 | 30 | 27 | 57 | 39.83 | 8.53 | 72.90 |
| LGIT 4 | 30 | 44 | 34 | 78 | 51 | 10.85 | 117.72 |

The pattern of C-tests means shows that the participants outperformed on C-test

2 (PGIT) and C-test 4 (LGIT), i.e. the C-tests that were made based on the translated (Persian into English) texts, compared with their authentic counterparts. In other words, the participants` performances on C-test 1 (PGAT) and C-test 3 (LGAT), made from authentic English texts, were lower than the other two C-tests (PGIT & LGIT).

In order to probe the nullhypothesis formulated earlier in this study, a two-way ANOVA was run so that we could compare the means of the four groups' performances on the C-tests. The results are depicted in Table 2.

**Table 2: Tests of Between-Subjects Effects, Univariate Analysis of Variance**

| Source | Sum of Squares | DF | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| GENRE | 13696.03 | 1 | 13696.03 | 106.71 | .000 |
| TEXT | 3674.13 | 1 | 3674.13 | 28.62 | .000 |
| GENRE * TEXT | 0.30 | 1 | 0.300 | .002 | .962 |
| Total | 409924.00 | 120 | | | |

Concerning the first null hypothesis, predicting no significant relationship between the performance of EFL students on the C-tests and different genres (political vs. literary), the results of the two way ANOVA show that the p-value for the F ratio of 106.71 is .000, which is much smaller than the level of significance (0.05) set for this study. Therefore, the first null hypothesis of the study is rejected and there is obviously a significant relationship between genre and the performance of EFL students on C-tests.

Regarding the second null hypothesis, stating no significant relationship between the performance of EFL students on C-tests and text authenticity, again, the results of the two-way ANOVA reveal that the p-value for the F ratio of 28.62 is .000, which is much smaller than the level of significance (0.05) set for this study. Accordingly, the second null hypothesis of the study is rejected and, the results confirm a significant relationship between the performance of EFL students on C-

tests and authentic and inauthentic texts.

As for the third null hypothesis, predicting no significant interaction between genres (political vs. literary) and text types (authentic vs. inauthentic) in the performance of EFL students on C-tests, as the results of the two-way ANOVA demonstrate, the p-value for the F ratio of .002 is .962, which is much larger than the level of significance (0.05) set for this study. In other words, the third null hypothesis is not rejected or there is not any significant interaction effect of genres (political vs. literary) and text types (authentic vs. inauthentic) on the performance of EFL students in the C-Tests.

In order to find out the location of the differences, post hoc comparisons of means were completed. The results of multiple comparisons and the homogeneous subsets for the C-Tests scores are illustrated in Tables 3 and 4.

**Table 3: Post hoc Scheffe Test, The Multiple comparisons**

| C-Tests (I) CELLCODE | C-Tests (J) CELLCODE | Mean Difference (I - J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1(PGIT) | 2(PGIT) | -10.96(*) | 2.92 | .004 | -19.26 | -2.66 |
| | 3(LGAT) | 21.46(*) | 2.92 | .000 | 13.16 | 29.76 |
| | 4(LGIT) | 10.30(*) | 2.92 | .008 | 2.00 | 18.59 |
| 2(PGIT) | 1(PGAT) | 10.96(*) | 2.92 | .004 | 2.66 | 19.26 |
| | 3(LGAT) | 32.43(*) | 2.92 | .000 | 24.13 | 40.73 |
| | 4(LGIT) | 21.26(*) | 2.92 | .000 | 12.96 | 29.56 |
| 3(LGAT) | 1(PGAT) | -21.46(*) | 2.92 | .000 | -29.76 | -13.16 |
| | 2(PGIT) | -32.43(*) | 2.92 | .000 | -40.73 | -24.13 |
| | 4(LGIT) | -11.16(*) | 2.92 | .003 | -19.46 | -2.86 |
| 4(LGIT) | 1(PGAT) | -10.30(*) | 2.92 | .008 | -18.59 | -2.00 |
| | 2(PGIT) | -21.26(*) | 2.92 | .000 | -29.56 | -12.96 |
| | 3(LGAT) | 11.16(*) | 2.92 | .003 | 2.86 | 19.46 |

* The mean difference is significant at the .05 level

Table 3 shows Multiple Comparisons of means, listing the differences between means (Mean Difference (I-J)), their standard errors, *p*-values and 95% confidence

interval for each pair. Inspection of the *p*-values shows that the scores obtained from C-Test 1 (PGAT) differ significantly (p<0.05) from those of C-tests 2, 3, 4 and vice versa. This is more clearly shown in table 4 in which the groups are divided into homogeneous subsets, thus showing which means differ significantly from one another (i.e. the members within each subsets).

**Table 4: Post hoc Scheffe tests for the C-Tests, means for groups in homogeneous Subsets**

| C-Tests | N | Subset for alpha = .05 | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| 3 (LGAT) | 30 | 39.83 | | | |
| 4 (LGIT) | 30 | | 51.00 | | |
| 1 (PGAT) | 30 | | | 61.30 | |
| 2 (PGIT) | 30 | | | | 72.26 |
| Sig. | | 1.00 | 1.00 | 1.00 | 1.00 |

To sum up the findings, the results of the present study indicate that genre and text authenticity significantly affected the performances of the testees on the C-tests, but no significant interaction between genre and text authenticity was observed.

The findings of the present study support Paltridge's claim (2001) holding that learners' awareness of the purpose and structure of the genre can help them in dealing with the unfamiliar content and vocabulary in the testing materials. The research results are also in consistency with Bachman's (1990:138) assertion that "the unfamiliarity with the characteristics of a given genre may also make the expected response more difficult".

In the present research, the performance of the EFL learners on those C-tests composed of the literary texts was rather low in comparison with the C-tests based on political texts. This finding supports the claims of Mackay (1992, p. 205) holding that cloze tests should not be based on literary texts because the 'normal criterion of semantico-grammatical appropriacy ' does not apply to these texts, and these texts

may be "low in cohesiveness and thus negatively affect a cloze test's ability to measure constraints across sentences" (Chavez-Oller et al., 1985, p.185).

The results indicate the existence of a significant relationship between text authenticity and students' performances on the C-tests. The findings are compatible with Nakamura's (1998) idea who suggested that in dealing with text, the crucial feature of authenticity in the testing materials and in the performance of the students should be considered. Likewise, the research findings are congruent with the findings of Sigott's (1995) study which showed that translation creates different texts. They are also congruent with the idea of Elmore and Headrick (2002:131) maintaining that "the wording used in the translated version of an assessment does not undergo the same process of refinement as the wording used in the assessment written in the original language".

Similarly, the results of the present study support Klein-Braley's (1994) suggestion that it is permissible to work with quasi-authentic texts when the authentic texts are too difficult for learners of a foreign language. Ranalli's (2002) already claimed that in text selection for C-tests, source and text difficulty are important variables.

On the other hand, the results indicated that 'genre' and 'text type' can be test method facets ; so the findings are in line with Lee (2000) who maintained that knowledge about genre and text type can help test makers to make informed selections from the mass of texts available for cloze, cloze elide and C-tests. Aitken's (1977:60) argues that "content can be a source of bias in scores". Along the same line, Klein-Braley and Raatz (1984:135) also argued that "care must be taken not to place an undue burden on test takers' knowledge of the content matter".

Additionally, since no interaction between the two factors under study (i.e. genre and text authenticity) was observed, it can be concluded that according to Gravetter and Wallnau (1999), the effects of one factor (genre) do not depend on the different levels of another factor (text authenticity) and vice versa.

Finally, the overall findings of this study support Kobayashi's (2005) comment holding that for language testers, or any participants making judgments on the basis

of test results, it is extremely important to pay attention to the test methods in order to develop proper assessment instrumentation or interpret test scores.

## Conclusion

The results of the study showed significant effects of genre and text authenticity on the performance of the EFL learners on the C-tests, i.e. genre and text authenticity were found as test facets in the C-tests. In other words, test performance was significantly affected by text type and genre. However, the interaction between genre and text type (authentic vs. inauthentic) turned out to be a very weak link.

Based on the findings of the present study, genre and text authenticity can be considered as test method facets in the C-tests, or according to Bachman (1995. p.113) they are contextual features of language that affect performance on language tests. It is specifically important since the facets may affect test takers' scores resulting in bias against certain test takers (Bachman, 1995, p.114).

Considering the purpose of language testing that according to Bachman and Palmer (1996:341) is "to enable the tester to make inferences about test takers ability to use language to perform tasks in a particular domain", the results of the present study may promise some implications for assessment of English for Academic Purposes (EAP) and in English for Specific Purposes (ESP) especially in the domains of international politics, political news and literary translations.

## References

Aitken, K. G. (1977). Using cloze procedure as an overall language proficiency test. *TESOL Quarterly*,11(1): 59-67.

Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

---. (1995). *Fundamental Considerations in Language Testing*. Third impression. Oxford: Oxford University Press.

Bachman, F., and S. Palmer. (1996). *Language testing in practice.* Oxford: Oxford University Press.

Brantmeier, C. (2005). Effects of Reader's Knowledge, Text Type, and Test Type on L1 and L2 Reading Comprehension in Spanish. *The Modern Language Journal*, 89 (1), 37-53.

Chavez - Oller, M. A., T. Chihara, et al. (1985). "When are close items sensitive to constraints across sentences?" *Language Learning*, 35: 181-206.

Elmore, P.B. and T.C. Headrick. (2002). A Proposed Number Correct Scoring Procedure Based on Classical True-Score Theory and Multidimensional Item Response Theory. *International Journal of Testing*, 2002, Vol. 2, No. 2.

Fountas, I. and G.S. Pinell. (2001). *Guiding Readers and Writers Grades 3-6: Teaching Comprehension, Genre, and Content Literacy*. Portsmouth, NH: Heinemann.

Greenfield, G. (2004). Readability Formulas for EFL. *JALT Journal* , 26 (1), 5-24.

Gravetter, F. J. and L. B. Wallnau. (1999). *Essentials of Statistics for the Behavioral Sciences*. Third Edition. Pacific Grove, CA: Brooks/Cole Publishing Company.

Kobayashi, M. (2002). Method effects on reading comprehension test performance: text organization and response format. *Language Testing* 19(2): 193-220.

---. (2004). Investigation of test method effects: text organization and response format: a response to Chen, (2004). Language Testing, 21, 235-244.

---. (2005). Lifelong learning. Proceedings of JALT Pan-SIG Conference. Tokyo, Japan : Tokyo Keizai University.

Koberl, J. and G. Sigott. (1994). Adjusting the C-Test difficulty in German. In Rudiger Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagenund praktische Anwoendungen*. Bochum: Brockmeyer.

Klein-Braley, C. (1994). *Language testing with the C-Test: A linguistic and statistical investigation into the strategies used by C-Test takers, and the prediction of C-Test difficulty*. Duisburg: Habilitationsschrift.

---. (1997). C-Tests in the context of reduced redundancy testing: an appraisal. *Language Testing*, 14, pp. 47–84.

Klein-Braley, C. and U. Raatz. (1984). A survey of research on the C-Test. *Language Testing* ,1, 134-146.

Laviosa-Braithwaite, S. (1997). *The English Comparable Corpus (ECC): A resource and a methodology for the empirical study of translation*. Manchester: UMIST.

Lee, D. (2000). Defining core vocabulary and tracking its distribution across spoken and written genres: Evidence of a gradience of variation from the British National Corpus. *Journal of English Linguistics, 27,* 52-58.

McBeath, N. (1990). C-Tests - some words of caution. English Teaching. *Forum,* 28, 45-46.

Nakamura, Y. (1998). *The Construct of Reading for Testing.* The English Teacher, Vol. XXVIII. Tokyo: Keizai University.

Nunan, D. (1991). *Language Teaching Methodology.* London: Prentice Hall International.

Paltridge, B. (1996). *Genre and the language-learning classroom.* Ann Arbor: University of Michigan Press.

---. (2001). *Genre and the Language Learning Classroom* . Michigan : University of Michigan Press.

Ranalli, J. M. (2002). Comparing Scoring Procedures on a Cloze Test. University of Birmingham.

Sharpe, P.J. (2001). *How to prepare for the test of English as a foreign language TOEFL*. New York: Barron's Educational Series, Inc.

Sigott, G. (1995). The C-Test: some factors of difficulty. *AAA- Arbeiten aus Anglistik und Amerikanistik,* 20 (1), 43-53.

Siong, K. (2004). Exploring the connection between the testing of reading and Literacy. *Journal of Language Studies*, 4(1) .

Widdowson, H. G. (1983). *Learning Purpose and Language Use*. Oxford: Oxford University Press.

Woolls, D. and P. King. (1999). *Creating and using a multilingual parallel Concordancer.* Birmingham: The University of Birmingham.