LET'S NOT IGNORE C-TEST AS A VALID AND RELIABLE PROFICIENCY TEST

Somaiyeh Shahri

Mrs., Adib Mazandaran Institute of Higher Education, Iran, somaiyeh_shahri1980@yahoo.com

Abstract

Integrative tests, due to some reasonable demerits of discrete-point tests, have drawn the attention of language testers to examine language as a whole and holistic phenomenon. Cloze test and C-test as the representatives of integrative tests are two appropriate and efficient tools for measuring language proficiency. However, C-test has been ignored in several cases in the process of developing language proficiency tests. The present study attempts to investigate the probable difference between C-test and cloze test in terms of their prediction of language learners' proficiency as measured by other standardized tests like the TOEFL. To this end, three tests (TOEFL, Cloze test, and C-test) were presented to 50 junior EFL learners. To investigate the probable difference between C-test and cloze test in terms of their prediction of language learners' proficiency, a regression analysis through the stepwise method was run. The C-test was entered as the first and most powerful predictor of the TOEFL. Thus, in this research, the C-test predicts the TOEFL much more accurately than cloze test.

Keywords: Integrative test, Cloze test, C-test, holistic approach, language proficiency test.

1. INTRODUCTION

Many of the most debates in the area of language testing and psychometrics in the last five decades have been about the advantages and disadvantages of discrete-point testing versus integrative testing. Discrete-point tests measure knowledge of individual language items at a time. Therefore each item tests only one element of particular component of the language. With the emergence of two alternative views of language proficiency assessment, integrative testing and research in the area of communicative competence, the discrete-point tests "have been pushed to periphery, at least theoretically". An integrative test requires a test taker to use several language skills at the same time (Duran, 1984; Oller, 1979; Farhady, 1979b Tajeddin, 2004).

Due to prominent position of cloze test among integrative tests, it has been positioned in the foreground of research. Research on item difficulty (Abraham & Chapelle, 1992, Kobayashi, 2002), construct validity (Chavez-Oller et al., 1985, Jonz,1990), text topic (Alderson & urquhart, 1985a, 1985b), deletion ratio (Alderson, 1979b,1980), gender (Farhady, 1982, Tajeddin, 1997), nationality (Farhady, 1982), learning style of field independence or dependence (Stansfield & Hansen, 1983, Chapelle & Roberts, 1986, Tajeddin,

1997), and native language (Spurling & Ilyin, 1985), background knowledge (Chacevych et al., 1982), trait structure (Bachman, 1982, Tajeddin, 1996), methods of student response (Bensoussan & Ramraz, 1984), language learning strategies and cloze (Tajeddin, 2004) are just some researches which have been conducted on cloze test. In consequence of these studies, much insight has been provided into the nature of the cloze.

In the 1960s, the use of cloze procedure has been the subject of intense debate in the profession; on the one side are those who see cloze as measuring the ability to make localized connections in the text (Alderson, 1978, 1979a, 1979b, 1983, Bachman, 1982, 1985). On the other side are the advocates of cloze who see the cloze task as involving the discourse processing ability which can provide a measure of language proficiency (Oller&Conrad, 1971, Chihara et al., 1977, Fotos, 1991).

As it is quite clear there were many researches in the cloze test domain. The researches encompass both linguistic and psychological aspects. Although some shortcomings have been ascribed to cloze test procedure, it is still one of the reliable language proficiency tests. To remedy the pitfalls of cloze test, c-test was first developed by Raatz and Klein-Braley (1981) as an alternative to cloze test. C-test has been investigated from different perspectives (e.g., Klein-Braley and Raatz, 1984, Cohen et al., 1984, Grotjahn, 1986, Grotjahn, 1987, 2002, Feldman and Sttemmer, 1987, Cleary, 1988, Chappelle and Abraham, 1990, Hood, 1990 Bradshaw, 1990, Kamimoto, 1992, 1993,Jafarpur, 1995, Jafarpur, 1999, Sigott and Korbel, 1996, Klein-Braley, 1997, Ikeguchi, 1998, Babaii and Ansari, 2001, Babaii and Jalali Moghaddam, 2006, Eckes and Grotjahn, 2006). These studies have shed light on merits and demerits of the c-test.

C-test has been praised for its high reliability and validity indices, ease of construction, administration and scoring (e.g. Feldman and Sttemmer, 1987; Connelly, 1997). Beside these merits, lack of face validity (Jafarpur, 1995; Bradshaw, 1990), Poor item discrimination (Cleary, 1988; Kamimoto, 1993), and unclear construct validity (Kamimoto, 1992; Feldman and Sttemmer, 1987) are some flaws that leveled against c-test. Scholars propose some alternatives to remedy these flaws. For example, Cleary (1988) suggests left-hand deletion rather than customary right-hand deletion to enhance discrimination. Kamimoto (1992) in a study to find out what a C-test measures claims that it tends to assess vocabulary competence as well as grammatical competence. Dorney and Katona (1992) carried out a study to validate the C-test among Hungarian EFL learners. They come up to the conclusion that it is both reliable and valid instrument. It also has high correlation with other language tests. It is a highly integrative language test which measures global language proficiency but is less efficient in testing grammar. They also consider sentence length as a predictor of efficacy of the measurement. The shorter the sentence, the better the text measures general language proficiency. Average sentence length is a desirable index of text cognitive load and syntactical complexity. Kamimoto (1993) proposes tailoring the C-test in terms of the classical item analysis. She concludes that leaving out items with less than 0.19 for Item Discrimination (ID), and less than 0.30 and more than 0.70 for Item Facility (IF) will be a possible solution to some problems with item characteristics. However, Jafarpur (1999) constructed a tailored C-test in which appropriate items are selected. The findings show that, the tailored C-test however has no advantage over the natural one. He reports the uselessness of tailoring the items on the basis of classical item analysis. Sigott and Korbel (1996) maintain that the tester may increase the difficulty level of texts to avoid ceiling effects in the C-test. It is possible through (a) increasing redundancy reduction by deleting 2/3 rather than 1/2 of the words or, leaving only the first letter; and (b) changing deletion pattern by deleting the first half of the word — left-hand deletion in Cleary's (1988) terms.

The most controversial issue about C-testing, however, is what it really measures. Chappelle and Abraham (1990) maintain that C-test checks more grammatical and less textual competence. Dornyei and Katona, (1992), state that since deletions are not at the sentence but at the word level it is not efficient in testing grammar. Some researchers have found indications of micro-level processing in the C-test with almost no trace of macro-level processing. Cohen et al. (1984, p. 225) maintain that because test takers are presented with half of the word in each mutilation, the examinees who do not understand the macro-context could use their vocabulary skills to fill in the gaps without much involving in higher-level processing. Feldman and Sttemmer (1987), based on the analysis of test takers' verbal protocols, conclude that both bottom-up and top-down processes are involved in completing a C-test. Babaii and Ansari's investigation (2001) revealed the "the C-test triggers both macro and micro aspects of language. That is to say, to the extent that, it conforms well to the principle of reduced redundancy which fundamentally emphasizes that both a global and a local knowledge are required to supply the missing elements in a distorted linguistic message" (Babaii and Ansary,2001, p.216).

In consequence of these studies, much insight has been provided into the nature of the C-test. However, it

appears that there is a dearth of evidence in studies which have approached to examine C-test and its probable superiority over cloze test. Hence, the researcher intends to investigate the probable differences between C-test and cloze test in terms of their prediction of language learners' proficiency as measured by other standardized tests like the TOEFL. Since C-test has been neglected in Iran, the findings might provide much insight into the nature and use of language proficiency tests. It might also shed light on the importance and usefulness of C-test as a valid and reliable proficiency test.

2. METHODOLOGY

2.1 Participants

A sample of 70 junior male and female students studying Teaching English at Sari Azad University was selected. Their ages ranged between 20 and 30.Since in this study Three tests (TOEFL, Cloze test, C-test) were given to the participants during three weeks, some of the subjects were absent for one of the tests. For the final analysis, the researcher took into consideration just those participants who took part in all the three tests. Thus, the final analysis covers 50 subjects. The TOEFL (1995, version) was administered to elicit information about the participants' general language proficiency. Since the researcher intended to have different levels of language proficiency, the group was not homogenized after administering the TOEFL test.

2.2 Instrumentation

To answer the research question posed by the researcher, three instruments were used.

2.2.1 Proficiency test

2.2.2 The TOEFL

The TOEFL (1995, version) was administered to determine the examinees' level of language proficiency. The test consists of 100 items (structure and written expression: 40 items; vocabulary: 30 items and reading comprehension: 30 items). Due to administrative limitations listening comprehension section was not included. Estimated internal consistency measures (KR-21) of TOEFL test computed as 0.83.

2.2.3 The cloze test

Three teacher-made cloze tests were used in the present study. The tests were constructed by deleting every 7th word from the three passages which were taken from Practice and Progress (Alexander, 1968). As is customary the first few lines were left intact. Overall, the subjects were supposed to fill in 100 blanks. For scoring, the exact word method was applied by the researcher. The estimated Flesch reading ease for the texts was (Mean= 61.8) which indicates an average difficulty level (Flesch, 1948). The texts were then pretested with a sample group enjoying characteristics similar to the target group. Estimated internal consistency measures (KR-21) revealed that all of the three texts enjoy high reliability values.

Table 1. Reliability of cloze sub tests and the total test

Subtest	Mean	Variance	KR-21	K
1	13.5	22.65	0.73	26
2	15	33.87	0.79	30
3	25.89	97.81	0.91	44
Cloze test	56.18	114.o6	0.85	100

To establish the empirical validity of the cloze test, the Pearson product-moment correlation coefficient between the cloze test (n=100) and TOEFL was calculated as 0.80.

2.2.4 The C-test

According to Klein-Braley and Raatz (1994) and klein-Braley (1997) to construct the C-test one should choose a number of texts more than necessary. Considering the guidelines given by Klein- Braley (1997) a C-test developed using eight excerpts taken from three ELT books, viz. Practice and Progress (Alexander, 1968), Readers' Choice (Boudoin, 1977), and Expanding Reading Skills (Markestein & Hirasawa, 1982).In the C-test, texts were arranged from easy to difficult and each text consisted of 25 mutilations. To decide on the five final texts comprising the C-tests, the newly developed C-test was administered to a sample group enjoying characteristics similar to the target group. The estimated Flesch reading ease for the texts was (64.58) which indicates an average difficulty level (Flesch, 1948). Table 2. shows the characteristics of each

text (subtests) used in piloting phase.

Table 2. Characteristics of the C-test (early Version)

Subtest	Mean	Variance	KR-21	IF	K
1	9.53	30.62	0.87	0.42*	20
2	7	41.5	0.93	0.29	20
3	6.36	26.4	0.87	0.30	20
4	8.2	26.9	0.86	0.31	20
5	11.5	27	0.85	0.57*	20
6	7.54	30.90	0.89	0.35*	20
7	7.46	23.22	0.83	0.32*	20
8	9.5	41.39	0.92	0.47*	20
C-test	67.09	1746.2	0.97	0.37	160

Estimated internal consistency measures (KR-21) of individual texts (subtests) revealed that all of them enjoyed high reliability values (Table 5). In fact , item characteristics indices functioned as a criterion for selecting five subtests (5×20=100 test items) constituting the final version of the C-test which consisted of only those subtests which enjoyed an IF index equal or above 0.32. Furthermore, to adjust the overestimated reliability of the C-test resulting from the application of measures of internal consistency, using the individual texts as superitems in Cronbach alpha formula (Klein-Braley, 1997), its reliability was again estimated (0.87).

To establish the empirical validity of the C-test (final version), the Pearson product-moment correlation coefficient between the final version C-test (n=100) and TOEFL was calculated as 0.88.

2.3 Procedures

At the first phase, the TOEFL (1995 version) was administered to determine the examinees' level of language proficiency.

At the second phase (a week later), the cloze test was administered. The students were reminded that in the case of cloze test they are to use only one word to fill in a blank. They were assured that spelling errors would not count against them as long as the scorer could read the intended word. The exact word method was used to score the tests.

At the third phase (a week later), the C-test was administered. For the C-test as is the natural practice the acceptable word method was utilized.

For the proficiency test a time limit according to the instruction imposed, but in the case of cloze test and C-test testees had enough time to attempt all items. The students finished cloze test at average of 70 minutes and C-test at an average of 60 minutes.

3. RESULTS AND DISCUSSION

The purpose of the present study was to investigate the probable differences between C-test and cloze test in terms of their prediction of language learners' proficiency as measured by other standardized tests like the TOEFL. A regression analysis through the stepwise method was run to investigate the degree of precision of C-test and cloze test in predicting the general proficiency of the students as measured by the TOEFL test.

Table 3, displays the model summary information. The C-test is entered as the first and most powerful predictor of the TOEFL. The R (Pearson Correlation Coefficient) is 0.88. If this r-value is squared (0.78), the resultant will be the common variance which shows the percent of TOEFL that can be predicted through C-test.

Table 3. Model Summary Information (the Amount of Adjusted R Square and R square)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.887(a)	.788	.783	5.44591	.788	177.951	1	48	.000
a Predictors: (Constant), C-test									
b Dependent Variable: Toefl									

Table 4, displays the results of the ANOVA test. The significant F-value (Sig. = .000 < .05) indicates that the regression model with the C-test as the only predictor of the TOEFL is an appropriate model.

Table 4. Regression Variance Analysis

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5277.641	1	5277.641	177.951	.000(a)
	Residual	1423.579	48	29.658		
	Total	6701.220	49			
a Predictors: (Constant), C-test						
b Dependent Variable: Toefl						

Table 5, contains the regression coefficient (B=0.793), and its constant (15.983). By using these two figures one can build the regression formula:

TOEFL = (C-TEST *0 .793) + 15.983

Table 5. Regression Equation

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	15.983	3.803		4.203	.000
	C-test	.793	.059	.887	13.340	.000
A Dependent Variable: Toefl						

Another important piece of information in Table (5) is the Beta-value which corresponds to the R-value (0.889). That is, if the C-test changes one unit, the TOEFL will change 0.887 units.

The last table shows the reason for exclusion of the cloze test. The results indicate that if the cloze test is entered into the model, it predicts only 22.3 percent of the TOEFL, and that a unit change in cloze test results in 0.191 (Beta) change in the TOEFL, and that all these figures are statistically non-significant because the t-value of 1.56 has a probability of 0.124 which is much higher than 0.05.

Table 6, Excluded Variables

Model		Beta In	T	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	cloze	.191(a)	1.568	.124	.223	.289
A Predictors in the Model: (Constant), C-test						
b Dependent Variable: Toefl						

Based on the observed statistics we can presume that C-test predicts the TOEFL much more accurately than cloze test.

4. CONCLUSION

To study the probable difference between C-test and cloze test in terms of their prediction of language learners' proficiency a regression analysis through the stepwise method was run to investigate the degree of precision of C-test and cloze test in predicting the general proficiency of the students as measured by the TOEFL test. The C-test is entered as the first and most powerful predictor of the TOEFL.

According to what we have in Beta-value which corresponds to the R-value of 0.889, we conclude that if C-test changes one unit, the TOEFL will change 0.887 units. On the other hand, if the cloze test is entered into the model, it predicts only 22.3 percent of the TOEFL, and that a unit change in cloze test results in .191 (Beta) change in the TOEFL, and that all these figures are statistically non-significant because the t-value of 1.56 has a probability of 0.124 which is much higher than 0.05.These findings by no means affect the value of cloze test, but seek our attention to C-test which in many cases has been neglected. C-test can be used along with cloze test in language proficiency tests.

As Klein-Braley and Raatz (1981) asserts, C-test is one of the most multi-function test types and suitable for different purposes. "It is an excellent testing method, as it provides a good and quick assessment of general language competence" (Dornyei and Katona, 1992, p. 12). In spite of the many advantages of C-test, it is a neglected language proficiency test in Iran. To the opinion of the researcher, C-test can be utilized along with cloze test in different language proficiency tests.

Considering the role and importance of C-tests in English language teaching, EFL teachers may use C-tests as an effective teaching tool (Klein-Braley and Raatz (1981). C-tests can be applied by teachers for teaching vocabulary and grammatical points in some reconstruction activities (Thornbury, 1997).

As Dornyei and Katona (1992, p. 206) assert, "students can easily C-tests for each other, which could be made into a game or competition". On the other hand, reconstructing the mutilated texts in a way that learners do in C-test is much more like doing a puzzle. It is like playing a game and this fact by no means is a pitfall. Contrary to what some researchers propose about lack of face validity due to similarity of completing C-test and doing puzzle, in the researcher's opinion this idea is not only a demerit, but also a merit. Learners' mind engages in the activity, they have to look for words which have been passive for a period of time. They have to find the rules and vocabulary through different layers of their mind and at the same time enjoy doing it.

REFERENCE LIST

Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: an item difficulty perspective. *Modern Language Journal*, 76, 79-468.

Alderson, J. C. (1978). *The effect of certain methodological variables on cloze test performance and its implications for the use of cloze procedure in EFL testing*. Paper presented at the 5th International Congress of Applied Linguistics, Montreal.

Alderson, J.C. (1979a). The cloze Procedure and Proficiency in English as a Foreign Language. *TESOL Quarterly*, 13(2), 219-28.

Alderson, J.C. (1979b). The effect of the cloze test of changes in deletion frequency. *Journal of Research in Reading*, 2, 108-118.

Alderson, J. C. (1983). The cloze procedure and proficiency in English as a foreign language. In J. W. Oller (Ed.), *Issues in language testing research* (pp. 205-217). Rowley, MA: Newbury House. (Reprinted from *TESOL Quarterly,* 13(2), 219-227, 1979)

Alderson, J. C. & Urquhart, A. H. (1985a). The Effect of Students' Academic Discipline on Their Performance on ESP Reading Tests. *Language Testing*, 2(2), 192-204.

Alderson, J. C. & Urquhart, A. H. (1985b). This test is unfair: I'm not an economist. In P. C. Hauptman, R. Le Blanc, & M. Bingcham-Wesche (Eds.). *Second Language Performance Testing* (pp. 25-43). Ottawa: University of Ottawa Press.

Alexander, L. (1968). *Practice and Progress.* Longman, London.

Alptekin, C. and Atakan, S. (1990). Field dependence-independence and hemisphericity as variables in L2 achievement. *Second Language Research*, 6, 135-149.

Babaii, E. and Ansari, H. (2001). The C-test: a valid operationalization of reduced redundancy principle? *System*, 29, 209-19.

Babaii, E. and Jalali Moghaddam (2006). On the interplay between test task difficulty and macro-level processing in the c-test. *System*, 34, 586-600.

Bachman, L. F. (1982). The trait structure of cloze test scores. *TESOL Quarterly*, 16, 61-70.

Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, 19, 535-56.

Bensoussan, M., & Ramraz, R. (1984). Testing EFL reading comprehension using a multiple-choice rational cloze. *Modern Language Journal,* 68, 230-39.

Bradshaw, J. (1990). Test takers' reaction to a placement test. *Language Testing* 7, 13–30.

Chacevych, A., Delva, C., Houle, F., Moon, R., & Upshur, J. A. (1982). Language testing and content predictability. Paper presented at the third Annual Language Testing Research Colloquium, Honolulu, Hawaii.

Chapelle, C. & Abraham, R. (1990). Cloze method: what difference does it make? *Language Testing,* 7, 121–146.

Chapelle, C. A., & Roberts, C. (1986). Ambiguity tolerance and field dependence as predictors of proficiency in English as a second language. *Language Learning*, 36, 27-45.

Chavez-Oller, M. A., Chihara, T., Weaver, K. A., & Oller, J. W. (1985). When are cloze items sensitive to constraints across sentences? *Language Learning*, 35, 181-206.

Chihara, T., Oller, J., Weaver, K., & Chavez-Oller, M. A. (1977). Are cloze items sensitive to constraints across sentences? *Language Learning*, 27(1), 63-73.

Cleary, C. (1988). The C-test in English: left-hand deletions. *RELC Journal* 19, 26–38.

Cohen, A., Segal, M. and Weiss, R. (1984). The C-test in Hebrew. *Language Testing* 1, 221–225.

Duran, R. P. (1984). Some implications of communicative competence research for integrative proficiency testing. Ion C. Rivera (Ed.), *Communicative Competence Approach to Language Proficiency Assessment: Research and Application*. Clevedon, Avon: Multilingual Matters.

Eckes, Th. & Grotjahn, R. (2006). A closer look at the construct validity of c-test. *LanguageTesting*, 23(3), 290-325.

Farhady, H. (1979b). The disjunctive fallacy between discrete-point tests and integrative tests. *TESOL Quarterly*, 13(2), 347-357.

Feldman, U., Stemmer, B. (1987). Think aloud a retrospective data in c-test taking: different languages different learners same approaches? In: Faerch, C., Kasper, C., (Eds.), *Introspection in Second Language Research* (pp. 251–267). Clevedon: Multilingual Matters.

Flesch, R. (1948). A New Readability Yardstick. *Journal of Applied Psychology,* 32, 321-333.

Fotos, S. S. (1991). The cloze test as an integrative measure of EFL proficiency: A substitute for essay on college entrance examination? *Language Learning*, 41(3), 313-36.

Grotjahn, R. (1986). Test validation and cognitive psychology: Some methodological considerations. *Language testing*, 3, 159-185.

Grotjahn, R. (1987). How to construct and evaluate a c-test: A discussion of some problems and some statistical analyses. In: R. Grotjahn, C. Klein-Braley and D.K. Stevenson (Eds.), *Taking their Measure: The Validity and Validation of Language Tests* (pp. 219-253). Bochum: Brockmeyer.

Hood, M. (1990). The C-test: A Viable alternative to the use of the cloze procedure in testing? In: L. Arena (Ed.), *Language Proficiency* (pp. 173-189). Plenum Press, New York.

Ikeguchi, C. (1998). Do different C-tests discriminate proficiency levels of EL2 learners? *JALT Testing & Evaluation SIG Newsletter*, 2, 3-8.

Jafarpur, A. (1995). Is C-test superior to cloze? *Language Testing* 12, 194–216.

Jafarpur, A. (1999). Can the C-test be improved with classical item analysis? *System*, 27, 79-89.

Jonz, J. (1990). Another turn in the conversation: what does cloze measure? *TESOL Quarterly*, 24, 61-83.

Kamimoto, T. (1992). An inquiry into what a C-test measures. *Fukuoka Women's Junior College Studies,* 44, 67–79.

Kamimoto, T. (1993). Tailoring the test to fit the students: improvement of the C-test through classical item analysis. *Fukuoka Women's Junior College Studies,* 30, 47–61.

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

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

Kobayashi, M. (2002). Cloze tests revisited: exploring item characteristics with special attention to scoring methods. *Modern Language Journal*, 86, 86-571.

Oller, J. W. (1979). *Language Tests at School: A Pragmatic Approach.* London: Longman.

Oller, J. W., & Conrad, C. A. (1971). The cloze technique and ESL proficiency. *Language Learning*, 21(2), 185-195.

Raatz, U. and Klein-Braley, C. (1981). The C-test — a modification of the cloze procedure. In: Culhane, T., Klein-Braley, C. and Stevenson, D., Editors, 1981. *Practice and Problems in Language Testing. University of Essex Occasional Paper* (pp.113-148). Colchester: University of Essex.

Sigott, G. and Kobrel, J. (1996). Deletion patterns and C-test difficulty across languages. In: Grotjahn, R., Editor, 1996. *Der C-test. Theoretische Grundlagen und Praktische Anwwndungen* (pp. 159–172).Bochum: Brockmeyer.

Spurling, S. & Ilyin, D. (1985). The impact of learner variables on language test performance. *TESOL Quarterly*, 19, 283-301.

Stansfield, C., & Hansen, J. (1983). Field dependence-independence as a variable in second language cloze test performance. *TESOL Quarterly*, 17, 29-38.

Tajeddin, Z. (1996). An Innovative Procedure for Investigating the Validity of Cloze Tests in Light of Their Trait Structures. Unpublished manuscript, Allameh Tabatabai University, Tehran, Iran.

Tajeddin, Z. (1997). *A Reappraisal of the Current Status of Cloze Test Validity: Is Field Independence a Source of Test Bias or Communicative Language Ability*? Unpublished manuscript, Allame

Tajeddin, Z. (2004). The Relationship between the Use of Language Learning Strategies and Performance on the Cloze. *IJAL*, 7(1), 63-79.

Thornbury, S. (1997). Reformulation and reconstruction: tasks that promote noticing. *ELT Journal*, 51(4), 326-335.