

## استخدام اللغة العربية في اختبار الاستيعاب القرائي في اللغة الانجليزية

جهاد حمدان

تركي دياب

**ملخص:** تهدف هذه الدراسة إلى تفحص الدور الذي يمكن أن تلعبه اللغة الأم (اللغة العربية) في تقويم الإستيعاب القرائي في اللغة الأجنبية (اللغة الإنجليزية)، وتعرض النتائج التي خلص إليها الباحثان من خلال استخدام نسختين من اختبار قرائي أحدهما صيغت أسئلته باللغة العربية، والأخرى باللغة الإنجليزية، أثر لغة الإختبار على تحصيل مجموعة تتألف من ستين طالباً من طلبة المرحلة الثانوية في الأردن في الإستيعاب القرائي في اللغة الإنجليزية. وبعد تلقي الطلبة دروسا تناولت مادة قرائية وتمارين في كتبهم المقررة مشابهة لتلك التي تم تضمينها في الإختبار، أعطيت مجموعتان نفس الإختبار القرائي مع فارق واحد هو ان أسئلة الإختبار كانت باللغة الإنجليزية بالنسبة للمجموعة الأولى وباللغة العربية بالنسبة للمجموعة الثانية. وقد أشارت نتائج الدراسة إلى أن الطلبة الذين تم اختبارهم بواسطة اللغة العربية قد تفوقوا في التحصيل على نظرائهم الذين تم اختبارهم بواسطة اللغة الإنجليزية. ومع ذلك فقد دلت النتائج على أن أداء الطلبة ذوي المستوى المتقدم في التحصيل القرائي في المجموعتين قد تساوا في أدائهم بغض النظر عن لغة الإختبار الذي أجابوا عنه. وعليه يرى الباحثان أن هذه النتائج الأولية قد تشجع الباحثين في الأردن على إجراء دراسات معمقة للظاهرة موضوع البحث في المستقبل.

## Using Arabic in Testing Reading Comprehension in English

Jihad Hamdan\*

Turki Diab\*

**Abstract** : This study aims at examining the role of the native language (Arabic) in assessing reading comprehension in the foreign language (English). It reports on the findings of a case study in which two test versions, one in Arabic and another in English, were used to assess the effect of the language of the test on the reading comprehension performance of 60 secondary school students in Jordan. After receiving instruction for one month using reading material in their prescribed textbook, two comprehension-groups were tested using the same reading test except that the language of the test was English for the first sub-group and Arabic for the second. The results showed that the subjects who were tested in Arabic outperformed their counterparts who took the English test version. However, the advanced subjects in the two groups did equally well regardless of the test version they took. The tentative results of the study may encourage researchers in Jordan to launch an in-depth investigation into the use of the native language as a possible alternative method in testing reading comprehension in English.

### **Introduction :**

Reading (in a foreign language) is a complex and an interactive process in which the reader is assumed to reconstruct a message encoded by a writer as a graphic display. This requires the reader to engage in a "

\* University of Jordan

psycholinguistic guessing game" where s/he examines a sample of language , predicts meaning , uses prior knowledge of the subject matter and tests the hypotheses s/he has made (cf. Goodman, 1971,1988; Coady 1979). The reading proficiency of Jordanian students in public schools, who study English for eight years , is often viewed as rather poor (cf. Al-Makhzoomy, 1986, p.20). This weakness may have been aggravated by the type and nature of reading comprehension tests in use, which is the basic concern of this paper.

Testing reading comprehension in English in Jordan and probably in other Arab countries seems to have received little research attention. This may be partly due to the fact that testing foreign language skills is a highly complex and time-consuming activity which requires sophisticated skills not always accessible to many practitioners (cf. Mackay et al., 1979, p. ix). In the absence of informative research in this area , however, it may be assumed that Jordanian teachers of English resort to a limited number of options in handling reading comprehension tests. Some of them make use of the guidelines in the curriculum documents published by the Ministry of Education (1993, p. 32). Such guidelines suggest a variety of test formats, viz., open - ended questions , sentence paraphrasing, translation, guessing meanings of unfamiliar lexis , and objective types of test items. The curriculum also prescribes cloze tests, summarizing, note-taking, and replying to letters. These guidelines, however, may not be always fully utilized. In reality, the curriculum documents are inaccessible to many teachers; they are used as points of reference by supervisors and administrators. In addition, the curriculum itself lacks detailed instructions on how to prepare and implement reading comprehension tests. Some teachers tend to imitate tests prepared by more experienced

colleagues or simply photocopy exercise materials in textbooks and use them as tests.

To get insights into what test formats Jordanian teachers of English follow, we examined more than a hundred reading comprehension tests given to secondary school students. The tests showed one common format: a reading passage followed by a number of multiple choice, true/false, and wh-questions. A sample of 30 teachers were also asked why they chose such a format. Some of them reported that they were influenced by the practices of other teachers in addition to the fact that a similar format was used to test their own reading comprehension when they were students. The rest added that they were imitating the reading comprehension component in the General Secondary Examination (locally known as Tawjihi). A closer examination of such components since 1990 showed that reading comprehension was tested mainly through a similar format.

The Jordanian ELT curricula for the Basic Education Stage (1990) and for the Secondary Education Stage (1993) suggest an 'eclectic' approach to teaching. This approach allows for the use of a variety of methods and techniques that would help reinforce learning. In particular, the ELT curriculum (1990, p. 59) suggests using "translation" as one of the most practical ways of clarifying the meaning of certain words". Moreover, the ELT curriculum for the Secondary Stage (1993, p. 18) proposes "the use of the mother tongue in the English classroom ..., when it facilitates learning the language". However, personal experience has shown that Arabic is often used in teaching but not in testing. It seems that teachers resort to L1 because they believe that this facilitates the teaching / learning tasks but refrain from using it in testing probably because tests are in a sense official documents that can be

seen by educational bodies who are not generally in favour of integrating L1 in the EFL learning process. This may account for the fact that L1 disappears from L2 classes once an official visitor (e.g. school principal, head teacher, supervisor etc.) steps into the classroom.

The idea of using the mother tongue in the foreign language classroom has been found useful by many researchers. Wilkins (1974, p. 82) observed that explanations and instructions in L2 classes may be given in the students' native language if they are to be understood unambiguously. Cochran (1985) described a number of classroom strategies for teaching limited-English proficient students. One of these strategies is the use of native language literacy as a basis for second language skills.

Al-Absi (1991) investigated the effect of incorporating Arabic in the teaching of English to Jordanian students. The findings provided evidence in favour of this method. Uram (1992) suggested that the use of the mother tongue in ESL instruction would be useful only when all students have the same native language. Uram recommended specific techniques which an ESL teacher would like to adopt. Amongst these techniques are : alternating the native language and English in class; providing reading material in the native language followed by discussion in English; inviting proficient ESL speakers as guest speakers using both languages; and translating into the native language when explanation in the target language is ineffective.

As the ultimate goal of reading an academic text is comprehension, i.e., obtaining and utilizing knowledge encoded in written form, there is no reason, in principle, why it should be necessarily measured in the target language of the given text. Anecdotal evidence may

highlight this point. Many Jordanian graduates who studied in languages other than English tend to ascribe their failure in professional certification examinations conducted in English to the language and format of such examinations rather than to their weakness in professional skills or academic knowledge. Many of them believe that their achievement could have been better had they been tested in Arabic or at least in the language through which they acquired their specialized knowledge (Jordanian Medical Association: personal communication).

This motivated us to hypothesize that the language of a reading comprehension test is a significant variable that affects the reader's performance. In this context, Weir (1990, p. 86) warns that "... test format may have an undue effect on the measurement of trait. It seems sensible, therefore to safeguard against possible format effects by including a variety of appropriate test methods in assessing competence in the various skills." Hence, L1 may turn out to be a possible testing method. Clarke (1972, p. 79) has actually included Arabic distracters in English vocabulary tests claiming that this technique "... provides a possible solution to the hard - to - find distracter problem". Hughes (1989, p. 129) on the other hand, argues that:

*The wording of reading test items is not meant to cause candidates any difficulties of comprehension. It should always be well within their capabilities, less demanding than the text itself. In the same way, responses should make minimal demands on writing ability.*

*Where candidates share a single native language, this can be used both for items and for responses.*

Furthermore, the idea of using L1 in testing L2 comprehension has received strong support from another authority on language testing. Weir (1993, p. 24) suggests that the test should be " candidate friendly, intelligible, comprehensive, brief, simple and accessible". Moreover, he argues that it is even preferable to give rubrics in the

candidates' first language as the aim of testing reading comprehension is to test the comprehension of a text rather than comprehension of a rubric. One way to make the test items simpler than the text, one may suggest, is to write them in the testees' native language in monolingual situations.

This study is meant as a contribution to the neglected domain of testing EFL reading comprehension in Jordan. It reports on the findings of a case study where two test versions, one in English and another in Arabic, were used to test the reading comprehension of secondary school students in Jordan. In particular, the study attempts to explore the role of L1 as a method of testing reading comprehension in L2. Section(2) below describes methodology and research design. The results are presented and discussed in Section (3).

### **Subjects**

The subjects were sampled from 85 first secondary (Scientific Stream) students at Sweileh Secondary Boys School , a public school in Amman II Education Directorate in Jordan. While the study was underway the subjects continued to receive their regular instruction in English in two separate classes. They used the same English language textbooks and were taught by the same Jordanian teacher. All subjects are native speakers of Jordanian spoken Arabic ( who also know Modern Standard Arabic). At the time of data collection , the subjects , whose mean age was 17, had received seven years of formal instruction in English as a foreign language (EFL) at the rate of 6 x 45 - minute classes per week. The results of a cloze test were used to determine the subjects' level of reading comprehension in English.

In an attempt to ensure the comprehensibility of the cloze test, the researchers selected its content from the Teacher's Book for the same grade. The test was originally a passage of approximately 300 words reporting a short tale

Arabic story ; the passage was new to the subjects. The researchers converted the passage into a 30-gap cloze test using the ' rational' deletion technique. This is accomplished by selecting words for deletion in accordance with certain discourse criteria (cf. Brown, 1994). The deleted words in the cloze test of the study belonged to different grammatical categories (i.e. nouns, verbs, adjectives, adverbs, articles and prepositions), and hence served different discourse functions.

The scoring procedure consisted of counting the number of exact or acceptable words restored to context. Those whose scores were below 49 /100 (n=15) were considered as low achievers(LAs) and therefore excluded from the study . Consequently, the subjects who proceeded to the study proper were 60 students whose scores ranged between 50 - 90 /100. One group (n=10) whose scores were between 50 - 69 / 100 were considered as intermediate achievers (IAs) and another group (n=50) who scored 70/100 and above were referred to as high achievers (HAs). Each group was further subdivided into two main groups, each consisting of 30 students: 5 HAs and 25 IAs.

The initial classification of subjects into three groups (HAs, IAs, and LAs) in terms of their scores on the cloze test does not have a theoretical base and has been adopted for practical considerations. For one thing, the grading system in Jordan considers 50/100 as the pass grade; those who score below 50/100 are often viewed as poor learners. In effect, the scores of 12 out of the 15 students who scored below 50/100 ranged between 15-35 /100. On a later check with the school administration, we found out that those students were weak in all school subjects. Therefore, they were excluded on the assumption that they would not contribute much to the study.

The level of achievement and the language of the test (see 2.3 below) are two important variables for this study. This may explain why the ten subjects who scored between 70-90 /100 were labeled HAs and those who scored between 50-69/100 IAs. A word of caution is due here. The terms HAs



and IAs are mere operational labels, and hence should be interpreted only with reference to the context of the study.

### **Materials**

Prior to data collection, the subjects were introduced over a month to a study unit selected from their prescribed English textbook: **Revised Oxford Secondary English Course - Book One**. The reading comprehension text in the study unit is similar in nature to the one used in the data collection test. The latter is on 'refrigeration and refrigerators'. Both texts are examples of scientific register. The study unit contained a long reading comprehension text (1200 words) on 'telecommunications' followed by exercises on specific reading skills, e.g., skimming (either for single facts or main ideas), guessing meanings of key words, understanding inter or intra-sentential relationships indicated by connectives in addition to relating reference words to ideas stated in the text. The unit also included activities on grammar, vocabulary and writing.

### **The test**

The test consisted of a reading comprehension text on 'refrigeration and refrigerators', followed by nine questions, prepared by the researchers, which were meant to test a sample of sub-skills necessary for the promotion of the reading comprehension skill. The text was selected from a scientific English textbook (cf. Bolitho and Sandler, 1980, pp. 13-15). Such a text was similar in nature to a number of texts presented in the students' course book. The questions on the reading text appeared in two equivalent versions, one in English (Appendix 1) and the other in Modern Standard Arabic (Appendix II). The English version will be referred to as the L2 version, whereas the Arabic version will be referred to as the L1 version. The use of two versions is basic to this research since its primary objective is to examine if there is a significant relationship between the test version and the rate of reading comprehension in L2. Table (1) below provides a description (based on Munby, 1978) of each question in the test.

Table 1: Classification of test questions in terms of reading comprehension sub-skills  
(Based on Munby, 1978)

Ques.	Reading comprehension sub - skills
1	Skimming for basic ideas
2	Understanding explicitly stated information
3&4	Deducing meanings of unfamiliar lexis (with or without the help of options)
5	Extracting information which indicates discourse functions such as cause - effect, reason, condition and exemplification
6	Scanning for specific information (i.e., listing)
7	Understanding intra-sentential relationships through grammatical clues
8	Understanding inter - sentential relationship through reference words
9	Reading for the gist e.g. suggesting a title

The specific reading skills included in the test resemble those highlighted in the exercises included in the course book unit to which the subjects were exposed at the onset of the study. This seems basic to ensure the reliability of the elicitation tool. Hughes (1989, p. 117) suggests that such micro-skills may be recognized " ... as skills which we might well wish to teach as part of a reading course..." He adds that items which test such skills are appropriate in achievement tests.

Since a test can be unreliable because of the way it is marked, the researchers devised a special marking scheme including a clear answer key and adhered to it. To ensure the validity of the test, an earlier version was reviewed by a jury panel consisting of six experienced teachers of English who teach the same grade (first secondary). The jury were requested to judge the suitability of the test with reference to such variables as content, length, level of difficulty, interest and cultural bias (cf. *ibid*, 1989, p. 120). The final test versions were prepared in light of received feedback.

The Arabic test version used the same English reading text; only accompanying questions/ test items were translated into Modern Standard Arabic. The subjects taking this version were requested to provide answers in

Arabic. However, it is worth noting here that contexts for guessing word meanings were given in English but the subjects were asked to make selection from options worded in Arabic (See Question 3) or to provide the meaning of English words in Arabic (See Question 4). Similarly, the options for answering the reference question (Question 8) were given in English.

The English test version was given to Group 1 and the Arabic test version was given to Group 2. The subjects in both groups needed 50-60 minutes to complete the answers.

### **Hypotheses**

The following three hypotheses were formulated:

**Hypothesis 1** (H1): There are statistically significant differences between the mean scores of Group 1 who take the L2 test version and the mean scores of Group 2 who take the L1 test version.

**Hypothesis 2** (H2): There are statistically significant differences between the mean scores of the HAs and the mean scores of the IAs in both groups.

**Hypothesis 3** (H3): There is some type of relationship between the achievement level of the learner (i.e. HA or IA) and the type of test version s/he takes. This interaction hypothesis (H3) is of special importance since the score of the learner is supposed to depend on his/her achievement level in English and whether or not s/he is tested in English or in Arabic. Part of the scope of this hypothesis can be expressed symbolically as follows:

- H3:- means of HAs in Group 1 > means of HAs in Group 2  
- means of IAs in Group 1 > means of IAs in Group 2  
- means of HAs in Group 1 < means of HAs in Group 2  
- means of IAs in Group 1 < means of IAs in Group 2

## Study Design and Statistical Analysis

The study provides an example of a 2 x 2 factorial design; a factorial analysis of variance is available. The paradigm of the study is presented in Figure 1 below:

Figure1: Treatment of reading comprehension

Group & Achievement Level	Test Version	
Group 1: HAs ( n=5) IAs ( n=25)	English	Arabic NA* NA
	SCORES	
Group 2: HAs ( n=5) IAs ( n=25)	NA	NA

\* NA: Not applicable

A two-way ANOVA was used to test any statistically significant differences between the means of the main effects (i.e. test version and achievement level) and the interaction between them.

## Findings and Discussion

Table (2) below sums up the mean scores and standard deviations of accurate responses for each test version and achievement level of subjects.

Table 2: Means and standard deviations of accurate responses in terms of test version and achievement level

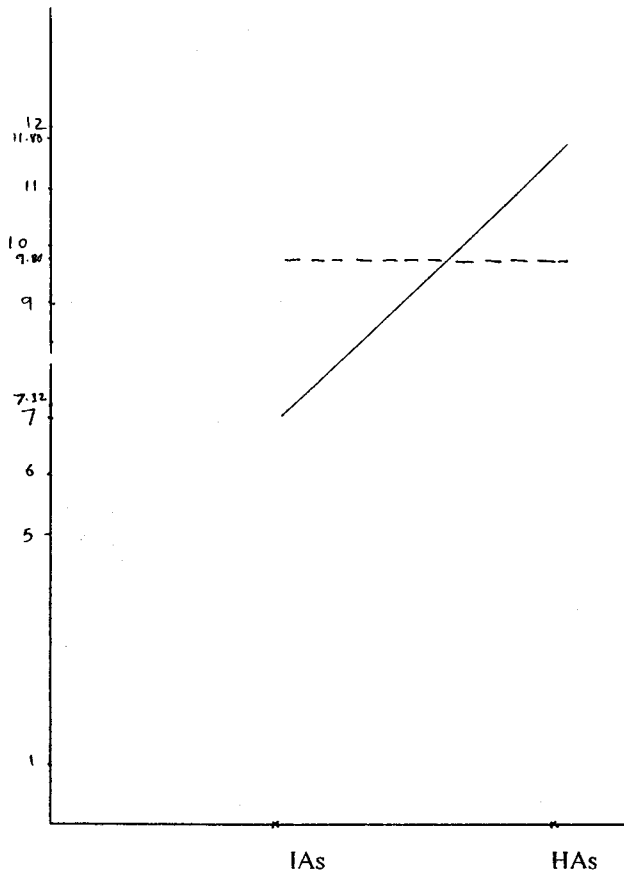
Variable	M	SD
<b>Test Version</b>		
L2 Version(English)	8.07	2.33
L1 Version (Arabic)	9.80	3.20
<b>Achievement level</b>		
HAs on L2 Version	11.80	1.30
HAs on L1 Version	9.80	3.56
IAs on L2 Version	7.32	1.68
IAs on L1 Version	9.80	3.20

Table (2) shows that there are differences in the means and standard deviations of accurate responses assigned to each test version as one whole. Similar differences also exist between HAs and IAs who took the

same test version. Moreover, there are differences in achievement level between HAs and IAs who took the L2 test version. However, such differences disappear between the two achievement level subgroups who took the L1 test version.

For the reader's convenience, the relationship between students' achievement level and the language of the test is presented diagrammatically in Figure 2 below.

**Figure 2: Diagrammatic representation of the relationship between achievement level and language of the test**



Achievement Level: HAs = High Achievers

IAs = Intermediate Achievers

Language of the test: English(L2) Arabic (L1) -----

ANOVA was conducted to test whether there were any statistically significant differences in the means of accurate responses that can be attributed to main effects (i.e., test version and achievement level) or the interaction of test version (i.e.,  $n=2$ ) and achievement level (i.e.,  $n=2$ ). The results are presented in Table (3) below:

Table 3: Two -Way ANOVA of main effects and interaction  
( From data of Table 2)

Source	DF	SS	MS	F value	Pr>F
Test Version (TV)	1	45.07	45.07	6.80	0.0117
Achievement level (AL)	1	41.81	41.81	6.31	0.0149
Interaction (TVxAL)	1	41.81	41.81	6.31	0.0147
Residual	56	128.69			
Total	59	257.38			

Table (3) shows significant effects for test version and achievement level. It also shows a significant interaction between test version and achievement level. That is to say, the three hypotheses of the study are confirmed.

Table (2) above shows that the mean score of subjects on the L2 test version is 8.07, whereas the mean score of subjects on the L1 test version is 9.80 . The difference between the two mean scores is statistically significant as indicated in Table (3) above. That is, the subjects whose reading comprehension skill was tested via L1 (Arabic) outperformed their counterparts whose reading comprehension skill was tested in L2 (English). This result suggests that the language of the reading comprehension test plays an important role in determining the rate of comprehension. Further, one may argue that a test of L2 reading comprehension via L1 may better reveal L2 learners' comprehension than a test of the same content but written in L2. After all, L2 learners who have access to L2 through formal instruction only do not necessarily need to demonstrate their understanding of L2 texts through L2. Once they leave the L2 classroom , the overwhelming majority of such L2 learners resort to communication in L1

about all sorts of topics including those read or discussed in the L2 class.

In line with H2 , Table (3) shows that the HAs did significantly better than the IAs. Notwithstanding, a significant interaction between test version and achievement level was revealed. In other words, the mean scores of subjects are influenced by two factors, namely, 1) their general achievement level in English and 2) the test version (L1 or L2) they were given.

To identify the sources of interaction between the language of the test (L1 or L2) and achievement level (HAs and IAs) revealed by ANOVA, a number of paired comparisons were carried out. Table (4) below shows the results of paired comparisons of subjects' responses across test versions.

Table 4: Paired comparisons for responses in terms of achievement level across test versions

Means	1	2	3	4
1. IAs in L1 9.80	—	ns	*	ns
2. HAs in L1 9.80		—	*	ns
3. IAs in L2 7.32			—	*
4. HAs in L2 11.80				—

\* P= 0.05

ns = not significant

A careful examination of Table (4) above shows the following:

1. There are no statistically significant differences between the mean scores of the IAs and those of the HAs who took the L1 test version (Arabic) . Both groups did equally well.
2. The mean scores of the IAs who took the test in Arabic are significantly different from those of the IAs who took the test in English. The difference is in favor of the former group.

3. The mean scores of the IAs who took the test in Arabic are not significantly different from those of the HAs who took the test in English.
- 4 The mean scores of the HAs who took the test in Arabic are significantly different from those of the IAs who were tested in English. The difference is in favour of the former group.
5. There are no statistically significant differences between the mean scores of the HAs; both sub-groups did equally well.
6. The mean scores of the IAs who took the test in English are significantly different from those of the HAs who took the same test version. The difference is in favour of the latter.

In view of the foregoing, it may be argued that the language of the reading comprehension test seems to play an important role in determining the understanding rate of comprehension within and across the two major study groups. While the HAs who were tested in English did significantly better than the IAs who were tested in the same language, as one would naturally expect, both the HAs and the IAs who were tested in Arabic did equally well. This finding motivates a conclusion that learners of intermediate achievement level may show a better understanding of an L2 reading comprehension text if the test items are written in their native language, and if they are required to answer them in the same language. This tentative conclusion is supported by the fact that the IAs who took the L1 test version outperformed their counterparts who were tested in L2.

On the other hand, the findings of the study indicate that the language of the test does not influence the rate of comprehension of high achievers. They will continue to do well regardless of the language of the test.



Based on the findings of this study, L1 is worth being considered as a possible alternative method in testing L2 reading comprehension. In particular, this alternative method may prove relevant to evaluating the professional knowledge and skills of students majoring in science or technology where the medium of instruction is English. In such contexts, the main objective of testing is to measure understanding of English technical texts. The language of the test should not add an additional burden to demonstrating comprehension. If this turns out to be correct, it may be reasonable to allow such test-takers to show their understanding of basic concepts and processes in their fields of specializations through comprehension tests written in their native language. To validate this more research is needed.

## References

- 1- Al-Absi, I. (1991). "The Effect of Using the Native Language as a Method in Teaching English as a Foreign Language on Developing Learners' Lexical Proficiency," M.Ed. Thesis, University of Jordan, Amman, Jordan.
- 2- Al-Makhzoomy, K. (1986). Teaching reading comprehension to secondary students in Jordan: suggestions for improvement. **DIRASAT** 13, pp. 19-29 .
- 3- Al- Mutawa, N. and Keilani, T. (1989). **Methods of Teaching English to Arab Students**, London: Longman .
- 4- Bolitho, R. and Sandler, P. (1980) .**Study English for Science**, London: Longman.
- 5- Brown, H. (1994). **Principles of Language Learning and Teaching**, 3rd edition. London: Prentice-Hall.
- 6- Clarke, M. (1972). Arabic distracters for English vocabulary tests. **ELT Journal** 27, pp. 77-90.
- 7- Coady J. (1979). A psycho - linguistic model of the ESL reader. In Mackay, et al., (eds.) **Reading in a Second Language**, pp. 5-12. Rowley, Massachusetts: Newbury House.
- 8- Cochran, C. (1985). **Effective Practices for Bilingual / ESL Teachers. Classroom Strategies for Limited English Proficient Students**, New Jersey State Dept. of Education, Trenton, Division of Compensatory / Bilingual Education.
- 9- Goodman, K. (1988). The reading process. In Carrell, et al., (eds.) **Interactive Approaches to Second Language Reading**, pp. 11-21 Cambridge: Cambridge University Press.
- 10- Hughes, A. (1989). **Testing for Language Teachers**, Cambridge: Cambridge University Press.
- 11- Larsen-Freeman, D. (1986). **Techniques and Principles in Language Teaching**, Oxford: Oxford University Press.
- 12- Mackay, et al., (eds.). (1979). **Reading in a Second Language**, Rowley, MA: Newbury House.
- 13- Ministry of Education. (1990). **English Language Curriculum and its General Guidelines for the Basic Education Stage**, Amman: Jordan.
- 14- \_\_\_\_\_ . (1993). **English Language General Guidelines and Curricula for the Secondary Education Stage**, Amman: Jordan.
- 15- Munby, J. (1978). **Communicative Syllabus Design**, Cambridge: Cambridge University Press.

- 16- Uram, A. (1992). "The What and How of Bilingual ESL," Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages. Vancouver, British Columbia, Canada.
- 17- Weir, C. (1990). **Communicative Language Testing**, London: Prentice Hall.
- 18- \_\_\_\_\_ (1993). **Understanding and Developing Language Tests**, London: Prentice Hall.
- 19- Wilkins, D. (1974). **Second Language Learning and Teaching**, London: Edward Arnold.
- 20- Zughoul, et al., (1988). **Revised Oxford Secondary English Course: Book one**, Amman, Jordan: Ministry of Education.