جامعة ديالى

أداء الطلبة العراقيين متعلمي اللغة الإنكليزية لغة أجنبية في مجال أشباه الجمل الاسمية: المشاكل والبرنامج العلاجي

رسالة

تقدمت بها الطالبة

شهد حاتم كاظم الطائي

إلى

مجلس كلية التربية /جامعة ديالى

وهي جزء من متطلبات درجة الماجستير/تربية

في طرائق تدريس اللغة الإنكليزية لغة أجنبية

إشراف

الأستاذ الدكتور عبد اللطيف علوان ألجميلي

الأستاذ المساعد الد كتورة الهام نامق خورشيد

ذو القعدة١٤٢٦/ هـ كانون الأول / ٢٠٠٥م

3.1 Elicitation procedures

An elicitation procedure is any procedure which causes a learner to make a judgement about the grammatical acceptability of a form or provokes him into generating a linguistic response based upon the grammar of his interlanguage (Corder ,1973:61)

For interlanguage research ,it is necessary to decide what kind of data to elicit ,whether textual or intuitional data.(Ibid:59).

For the present study ,the recognition and production techniques are used .This test is designed to find out how proficient our learners are in the handling of the NCs .

A pre requisite for any test is to specify the population from which the sample of the study is to be selected. Validity and reliability of the test are essential steps before its administration.
The following sections explain the procedure followed in the collection of the data

3.2 population

Population refers to any set of items ,individuals ,etc, which share some common and observable characteristics from which a sample can be taken . (Richards et al ,1992:282)

The Population of the present study is the fourth year students of the English Department at the College of Education/ University of Diyala.

The total number of the population is (141)distributed into three sections A,B and C.

3.3 Sample

Richards et al (1992:321) point out that populations may vary in their structures, which require the sample to represent these structures.

The sample of the present study which consists of (50)students (both males and females)is randomly chosen from the study population. See Table (1)

The following steps are followed in selecting the sample of the study :

1-Writing the names of the three sections on slips of paper and putting them in a container and putting one slip to represent the sample of the study which is section (A) of (52) students .

2-Drawing (2)fourth year students from section (A) to be added to section (B):Thus ,the number of sample is (50) and the population is (141).

3-The selected sample represents 35% of the whole population.

Besides another (50)students sat for the pilot test ,this choice is made on the premise that these students have studied the subject of NCs from which the data of the test are selected.

Table (1)

Description of the Sample& the Population of the Present Study

University	***College***	***Department of English Students***	***No. of Fourth Year Students***	***Sample***	***Population***
Diyala	Education	Section (A)	52	50	141
		Section (B)	48		
		Section (C)	41		

3.4 Selection of Material

The first step in test construction is to determine what material should be included in the test that meets the aims of the study which read as follows:

1- Investigating the ability of Iraqi EFL learners to identify and produce finite and non-finite noun clauses by form and function and finding out the area of difficulty in this respect.

2-Suggesting a remedial work for the alleviation of these difficulties. Since the test is designed for the fourth year students at the Department of English in the University of Diyala, the material has been carefully selected in such a way to ensure that the subject has already been introduced to the subjects of research. Most of the items selected are taken from the following sources

_Quirk et al(1985) **A comprehensive Grammar of the English language.**

_Quirk ,R. &Greenbaum, S.(1990) **A Student s Grammar of the English Language.**

_Quirk ,R.& Greenbaum ,S.(1973) **A University Grammar of English Language .**

_Stage berg ,N.(1981) **An Introductory English Grammar**.

In addition to the above mentioned sources ,some items are chosen from Internet sources.

3.5 Construction of the Test

Carroll(1968:46) defines educational test as a procedure designed to elicit certain behaviour from which one can make inferences about certain characteristics of an individual. Bachman(1990:20)agrees with Carroll in defining the test as a measurement instrument designed to elicit a specific sample of an individual's behaviour.(See alsoDarwesh& AL-Jarah, 1997:7)

Tests are normally designed to measure what the learner knows of the target language ,or, at least ,that part of it he has been taught.(Corder,1973:60) For further information c.f Ingram (1974:313) , Seliger and Shohamy (1989:126).

Jacob (1987:45)points out that a test is a procedure used to collect data on subject's ability or knowledge of certain disciplines. In second language acquisition research ,tests are generally used to collect data about the subject's ability in and knowledge of the

second language in area such as vocabulary, grammar , reading and general proficiency.

The test of the present study consists of three tasks, each of which consists 20 items .The targeted knowledge differs according to task ; Task 1and 2 are geared to elicit receptive knowledge whereas Task 3 is meant to elicit productive knowledge.

3.5.1 The Recognition Tasks

Elicitation Procedures are used to find out something specific about the learner's language ,not just to get him to talk freely. To do this, constraints must be placed on the learner so that he is forced to make choices within a severely restricted area of his phonological, lexical ,or syntactic competence .These constraints can be applied in two ways ;as an ordinary test, the first way by limiting the range of possible choices, as in a closed item recognition test, and the second way will be mentioned in ***3.5.2*** below . The choices in a recognition procedure will be based upon what learners at that stage are known , believed, or may be predicted to do (Corder, 1973:61).

Lado (1967:150) indicates that testing structure at the recognition level is a relatively simple matter which can be solved satisfactorily under any of the conditions found in testing.(See also AL-Jarah,1991:37)

In the present study ,the first task involves the subjects' responses concerning the identification of NCs from adv. and adj. clauses .

Task one includes (20) items, (7) of which are finite NCs which have different forms e.g.(**that –clause, wh-word** and **yes-no clause**),(7)of which are non-finite NCs which have different forms e.g.(**to-infinitive and ing-clauses**)and (6)items are divided equally between adj. and adv. clauses. See(Appendix 2).

In Task Two which consists of (20) items ,the subjects are asked to recognize the function of NCs according to the seven abbreviated forms that are given in the test. It involves(4) items a subject [SV] ,(3) as object [O], (3) as object complement [Co], (3) as object of preposition [OP],(2) as indirect object[Oi] and (2) as Appositive [Ap]

3.4.2 The Production Task

Freeman &Long (1990:35) point out that the production tests intend to measure the subject's ability to produce statements of certain linguistic features .It is based on the concept of elicitation where the subjects are asked to perform a language task based on certain stimulus(e.g. pictures, games, list of words ,etc.)in natural settings .The test can be developed to measure the route of acquisition grammatical usage , syntactic structures , etc. .The results are scored on the account subject production of correct items.(See also Nunan, 1992:68)

Tarone (1981:289) describes the production strategy as an attempt to use one's linguistic system efficiently and clearly, with a minimum of effort.

Lado (1967:164) indicates that testing structure of a foreign language on production level is more complicated than that of testing structure on a recognition level. The general technique to test production consists of eliciting from the student certain key structures which we wish to test.

As mentioned above ,Corder (1973) states that constraints can be applied in two ways in testing ,the second way is by restricting contextually the range of possible free choices as in an open-ended production test .The contexts for productive elicitation exercises is selected to elicit lexical items or syntactic forms which learners have already produced or may be predicted to produce in such contexts.

In the present study ,the Third Task elicits knowledge on the production level . It is made up of (20) items which the subjects are asked to replace finite NCs by a non-finite NCs , whenever possible, without changing the meaning of a sentence.

3.6 Test Validity

Heaton (1983:153) defines validity of a test as the extent to which it measures what it is supposed to measure and nothing else. Every test , whether it be a short, informal classroom test or a public examination ,should be as valid as the constructor can make it.(See also Harris, 1969:18;Bachman ,1990:279).

Thus, in Hughes's words (2003:26) a test is said to be valid if it measures accurately what it is intended to measure. We create language tests in order to measure such essentially theoretical constructions as reading ability , fluency in speaking, control of grammar and soon. (See also Ur , 1996:44)

Two types of validity are considered important, content validity and face validity, Therefore, both types have been adopted for the purpose of this study. Below is a brief explanation of the major features of both:

3.6.1 Content Validity

Undoubtedly, the most essential aspect for achievement tests is content validity, the extent to which the test adequately covers the syllabus area to be tested .To have content validity ,a test must reflect both the content and the balance of the teaching which leads up to it.(Deale ,1975:30).

A test is said to have content validity if its content constitutes a representative sample of the language skills, structures etc.. It is obvious that a grammar test, for instance ,must be made up of items relating to the knowledge or control of grammar. The test would have content validity only if it included a proper sample of the relevant structures. Just what are the relevant structures will depend, of course, up on the purpose of the test. (Hughes &Woods,2002:40)

To Heaton(1983:154) content validity depends on a careful analysis of the language being tested and of the particular course

objectives. The test should be so constructed as to contain a representative sample of the course, the relationship between the test items and the course objective always being apparent .

According to Mehrans and Lehmann (1973:290) content validity is particularly important for achievement tests .

To achieve content validity , a table of specifications is developed in which behaviours and content area to be tested are clarified .With reference to behaviors ,production and recognition tests are constructed.(AL-Hamash et al ,1982:65)

3.6.2 Face Validity

Face validity of the test is its surface acceptability to those involved in its development or use, but no more.(Mc-Namara,2000:50)See also(Thornbury,1999:141)

Hughes(2003:50)maintains that a test is said to have face validity if it measures what it is supposed to measure , a test which does not have face validity may not be accepted by candidates, teachers, education authorities or employers. (See also Heaton, 1983:153; Harris ,1969:21).

For this purpose , face validity is ensured by exposing the test to a jury of experts indifferent positions. The three tasks of the test are examined by a number of specialists*1 in English language teaching and linguistics .Each specialist is requested to point out his \her remarks and suggestions about the suitability of the test to the level of the students and to the aims of the study. The majority of specialists have verified the validity of test items.

Table (2)

Behavioural and content specification

Content area	***Weight***	***No. of test items***	***Behaviour***	***Note***
Task one			***R.***	
1-Finite NCs 2-Non-finite NCs 3-Adj. clauses 4-Adv. clauses		7(2,4,9,10,14, 16,18) 7(1,5,6,8,12,1 3,20) 3(3,15,19) 3(7,11,17)		A. to be able to identify NCs from adj.& adv. clauses and state whether NCs is finite or non-finite.
Task two			***R.***	
A-(SV) B-(Cs) C-(O) D-(Oi) E-(Co) F-(OP) G-(AP)		4(1,8,12,15) 3(4,18,19) 3(3,6,10) 2(9,16) 3(11,20,7) 3(5,13,2) 2(14,17)		B. to be able to recognize the function of NCs in the items according to the abbreviated form.
Task three			***P.***	
		20 items		C. to be able to replace the finite NCs by non-finite NCs, with out changing the meaning of a sentence .
Total		***60 items***		

3.7 Pilot Study

Bachman &Palmer(1996:234)state that pre-testing precedes operational testing in time ,and its sole purpose is to collect information about test usefulness in order to make revisions in the test itself and in the procedures for administering it rather than to make inferences about individuals*2.

In order to estimate the time needed, and to evaluate the test in terms of reliability and item analysis ,fifty students are chosen randomly as the subjects of the pilot study. Each student is given a copy of the test after being informal of the purpose of the test and the way to answer the items.

The pilot study has revealed that the time needed for the final administration of the test is fifty minutes.

After the first application of the test ,item analysis is carried out to check the difficulty level (DL) and discrimination power (DP)of each item in the light of the subjects' responses.

According to Madsen(1983:182)it is not completely accurate to think of very difficult or very easy items as weak items. The difficult items of grammar may appear as a result of not spending enough time in teaching them or having not been presented clearly enough. An easy item may show that all students have mastered that grammatical point which this item represents.

3.8 Item Analysis

Heaton(1983:172) indicates that item analysis is important for it provides information concerning the performance of the items comprising the test .By such information we can identify difficult items as well as those items which were answered correctly by the more able students. See also (Lado,1961:342; McNamara,2000:60)

Since item analysis is important ,both difficulty level and discriminating power of each item have been calculated.

The first step is scoring the test papers, the researcher has ranks the test papers from the highest to the lowest score. Test papers are separated into two subgroups, the upper consisting of 50 percent of the total group who receive the highest scores and the lower consisting of equal number of papers of those who receive the lowest scores.

To find out the DL for each item, the following statistical formula is used :

$$DL = \frac{\text{High in correct} + \text{Low in correct}}{\text{Total number of sample}} = \frac{Hi + Li}{N}$$

(Al-Dulaimy & Al- Mahdawi ,2000:54)

According to Madsen (1983:181) the accepted range of the level of difficulty is from (0.30 to 0.70) whereas other specialists like Bloom et al (1981:95) state that " a good spread of result can

be obtained if the average difficulty of the items is a round 50 to 60 percent and items vary in difficulty from 20-80 percent". It has been found that item difficulty of the present study was ranged between (0.21-0.80).

Specialists agree upon the principle that some items of a good test are to be of the kinds that can be answered by only exceptional testees. In this respect, Heaton(1983:173) confirms that difficult items will certainly separate the good student from the very good one ,which easy items separate the poor student from very poor one. In addition ,he believes that inclusion of difficult items is necessary to motivate the good students, and very easy items will encourage and motivate poor students.(See also Mc-Namara ,2000 :60)

The second step then, is to determine how well each item discriminates between high and low level examinees, for each item in the test should help to separate the proficient subjects from these who lack the tested skills or learning. (Harris,1969:105).

To Hughes(2003:226) a discrimination index is an indicator of how well an item discriminates between weak candidates and strong candidates. The higher its discrimination index ,the better the item discriminates in this way.

After the application of the formula of the item discrimination power it ranges between 0.30 and 0.68. According to Ebel's index of discrimination ,good classroom test items have

indexes of discrimination of (0.30) or more. (Ebel,1972:399)(See Table 3)

The discriminating power of an achievement test refers to the degree to which it discriminates between students with high and low achievement.(Gronlund,1965:267-268).

The statistical formula of DP of item used is :

$$DP = \frac{Ru - RI}{1\backslash 2\ T}$$

DP=discrimination power.

Ru=the number of pupils in the upper group who got the item right.

Rl=the number of pupils in the lower group who got the item right.

T=the total number of pupils included in the item analysis.

(Mehrens &Lehman,1973:192)

On the basis of the results of item analysis results, some items*3 have proved to be non-discriminative .Consequently some modifications are made, and are subsequently exposed to the jury members to ensure item validity. Later ,these items are given to the same (50) students to be answered. Results of applying the discriminating power and item difficulty formula have revealed that they are discriminative.

3.9 Test Reliability

Reliability refers to the consistency of scores obtained by the same persons when re-examined with the same test on different

occasions , with different set of equivalent items ,or under other variables examining conditions (Anastasi, 1976:103) .In addition to Anastasi ,Harris (1969:14)sees reliability as the stability of test scores. A test can not measure anything well unless it measures consistently. See also (Lado,1961:330; McNamara, 2000:61 and Thornbury,1999:141).

The concerns of reliability and validity can thus be seen as leading to two complementary objectives in designing and developing tests:

1-To minimize the effects of measurement errors.

2-To maximize the effects of the language abilities we want to measure. (Bachman ,1990:161)

To determine the reliability of the study test, Cronbach (1984:295) has developed the following general formula for estimating internal consistency which he called (coefficient alpha).

$$\& = \frac{K}{K-1}\left(\frac{1-\sum S^2 i}{S^2 X}\right)$$

Where:

&=Coefficient alpha

K=the number of items of the test

S^2 i=the sum of the variances of the different parts of the test.

S^2 x=the variance of the test scores.

Table (3)

The Difficulty Level and the Discriminating power of the Test Items .

Task One Items	***Item Difficulty***	***Item Discriminating Power***
1	0.30	0.39
2	0.33	0.40
3	0.25	0.41
4	0.32	0.35
5	0.26	0.44
6	0.32	0.64
7	0.30	0.50
8	0.35	0.43
9	0.28	0.42
10	0.50	0.37
11	0.22	0.40
12	0.42	0.48
13	0.45	0.51
14	0.25	0.40
15	0.28	0.38
16	0.21	0.42
17	0.22	0.36
18	0.26	0.38
19	0.24	0.36
20	0.30	0.36
Task Two Items	***Item Difficulty***	***Item Discriminating Power***
1	0.34	0.52
2	0.60	0.46
3	0.75	0.40
4	0.64	0.40
5	0.58	0.52
6	0.25	0.68
7	0.53	0.40
8	0.54	0.68

9	0.68	0.48
10	0.62	0.36
11	0.72	0.50
12	0.54	0.41
13	0.43	0.46
14	0.76	0.40
15	0.50	0.53
16	0.80	0.41
17	0.78	0.40
18	0.64	0.45
19	0.79	0.41
20	0.24	0.41
Task Three Items	***Item Difficulty***	***Item Discriminating Power***
1	0.28	0.52
2	0.76	0.38
3	0.74	0.40
4	0.33	0.53
5	0.52	0.48
6	0.42	0.40
7	0.66	0.36
8	0.25	0.30
9	0.20	0.40
10	0.28	0.40
11	0.21	0.41
12	0.68	0.48
13	0.26	0.41
14	0.80	0.40
15	0.68	0.32
16	0.52	0.56
17	0.78	0.32
18	0.40	0.66
19	0.75	0.42
20	0.24	0.52

The reliability coefficient is obtained through the use of **cronbach alpha** for the whole test .A coefficient Alpha of 0.85 is considered acceptable as a reliability index (Carroll and Hall, 1985:188).Nunnally (1972:226) sees that the coefficients above the level of (0.50) and reaches more than (0.68) statistically acceptable.

The statistical analysis has shown that reliability coefficient of Task One is (0.86)Task Two is (0.84) and of Task Three is (0.85) . This means that the test is highly reliable.

3 .10 Test Administration

After securing the test objectivity ,validity and reliability, and making the necessary changes ,both the Recognition Test and the Production Test were administrated simultaneously on 9 th of May , 2005 to the fourth year students at the Department of English\ College of Education \ University of Diyala during the academic year 2004 - 2005.

The test is given to the fourth year students under the same conditions. What they are required to do in each task is given to them in English and explained in ,i.e Arabic ,in order to clarify the ambiguities that they might meet when answering the questions. They were encouraged to ask for any information that has to do with the questions.

Carroll(1980:16) maintains that a good test is expected to provide as much information as is required with minimum

expenditure of time ,effort and resources ,moreover , the students were asked to write their responses on the test papers so as not to waste time and effort.

They were also asked to write their names on the test papers in order to encourage them to participate seriously and do their best .

3.11 Scoring Scheme of the Test

The main test is of two parts :a recognition and production part. It consists of sixty items. Each item is assigned one mark or two marks so that the total score of the test is 74 marks.

The students' responses are categorized according to two criteria correct or incorrect .Task One is different from the other tasks in scoring ,since it includes two parts ,part one is designed to measure the identification of NCs from adjectival and adverbial clauses , so that one mark for recognition of the NC.

In part two ,one mark is given to the correct answer for recognition of finiteness and nonfiniteness, zero for incorrect ones, while the items that are avoided are considered wrong .It is worth noticing that part two is out of (14),adjectival and adverbial clauses are excluded in scoring part two.

The second and third tasks are different in scoring ,one mark is given for the correct item and zero for the incorrect and avoided items.

3.12 Statistical Methods

The following statistical methods are used in the present study:

1-Cronbach's alpha formula is used to find out the reliability of the test .

2-One Sample t-test is used to find out the level of the sample testees in recognition and production skills :The following formula is used :

$$t = \frac{\overline{X} - M}{S / \sqrt{n}}$$

Where:-

T= t-test

X= mean

M= theoretical mean

S= standard Deviation

N = number of subjects

{Madsen,1983:170}

3-Pearson correlation coefficient formula is used to find out the correlation coefficient among all three tasks .

$$r = \frac{N.\sum X.Y - (\sum X)(\sum Y)}{\sqrt{[N.\sum X^2 - (\sum X)^2][N.\sum Y^2 - (\sum Y)^2]}}$$

Where:

r = the correlation coefficient.

N= the whole number of the tests.

X = the scores of the first question.

Y = the scores of the second question.

(Glass& Stanley ,1970:114)

4- T-test for testing the significance of the correlation coefficient is used .

$$T = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

Where :-

r= the correlation coefficient.

N= the whole number of the testees .

(Uda& Al-Khalili ,2000:305)

5-The standard deviation is estimated by applying the following formula:-

$$S = \sqrt{\frac{(x-\overline{x})^2}{N}}$$

(Al-Bayati & Athnaciyus ,1977:162)

6- ***Percentages*** of correct and incorrect answer of each item and subjects have been used in order to find out the distribution of both correct and incorrect answer made by the subjects of the present study .

(Runyon, 1973 :28)

Notes To Chapter Three

*1-The following are names of the members of the jury arranged alphabetically and their scientific degree.

1-Prof. Kadhim Al-Jawadi ,Ph.D. College of Arts,University of Baghdad .

2-Prof. Sabah S. Al Rawi ,Ph.D. College of Languages, University of Baghdad .

3-Assist Prof. Abdullh Salman Abbas ,Ph.D. College of Education, University of Diyala .

4- Assist Prof. Ali Khudheir ,Ph.D. College of Arts , Al-Mustansiriya University .

5- Assist Prof. Firas Awad ,College of Education /Ibn Rushd, University of Baghdad .

6- Assist Prof. Irfan Saeed ,Ph.D. College of Arts , Al-Mustansiriya University .

7- Assist Prof. Khail Ismaeel Rjaya ,Ph.D. College of Education, University of Diyala .

9- Assist Prof. Lemia Al-Ani, College of Education /Ibn Rushd, University of Baghdad .

10- Assist Prof. Muayyad M. Saed ,Ph.D. College of Education / Ibn Rushd ,University of Baghdad .

11- Assist Prof. Munthir Manhal ,Ph.D. College of Languages, University of Baghdad .

12- Assist Prof. Umran M. Mahood ,Ph.D. College of Education/ Ibn Rushd ,University of Baghdad .

13-Instructor Ayad Hameed, Ph.D. College of Education, University of Diyala .

14- Instructor Sami Al- Ma'muri ,Ph.D. ,University of Diyala.

The statistical methods have been suggested by the following two experts:-

1- Assist Prof. Ihsan A. Al Dulaimy, Ph.D. Ibn Al Hathem , University of Baghdad .

2- Assist Prof. Salih Mehdi Salih ,Ph.D. , College of Education, Al-Mustansiriya University .

*2- The term **pretest** can be used to refer to the procedure itself ,as well as to the activity ,so that **pretest** is commonly used as either a noun or a verb .Other terms that are commonly used more or less synonymously with **pretest** include **pilot ,trail and try-out .**

*3- The following are non-discriminative items ,these items replaced by another items that are seemed more suitable for the learners' level .

1- The first man who landed on the moon was an American scientist . (omitted item)

-The old carpenter who had been laying the floor stood up and straightened his back . (added item)

2- As she was short of money ,she decided not to buy a car. (omitted item)

-He looked at the toad as if it were poisonous . (added item)

3-Today he received a telegram specifying all the details. (omitted item)

-The day when my passport was stolen was Friday **.** (added item)

4-If the weather permits we shall go out for a walk tomorrow. (omitted item)

- A dog will run away if looked steadily in the eyes . (added item)

5-Having lived in this city for five years ,you ought to know your way . (omitted item)

-The accident happened where the two wards cross. (added item)

6- The first man to land on the moon was an American. (omitted item)

-I saw the girl whose dog you walked. (added item)