

# Analysing Students' Performance in Spoken Language Tests Using the Criterion of Accuracy

Ayad Enad Khalaf

*The Directorate of Religious Education and Islamic Studies, Sunni Endowment.*

[dr.avadenad@gmail.com](mailto:dr.avadenad@gmail.com)

ARTICLE INFO
Submission date: 9/12/2019
Acceptance date: 16/3/2020
Publication date: 8/5/2020

## Abstract

Testing speaking, as any other skill, must depend on criteria that will help the tester to be accurate in testing. These criteria may include; range, accuracy, fluency, interaction, and coherence, as provided by the Council of Europe (2001) in the *Common European Framework of Reference (CEFR)* in order to facilitate the process of measuring and to be a base for all the testers. Although, testing student's performances in spoken language is still one of the difficult and essential missions for the teachers of ESL to assess their students. Spoken language test differs from the kinds of tests for two reasons; first, it is more difficult in design, administer and mark, second, it does not easily fit the conventional assumptions about people and testing. Spoken language testing is likely to be more complex than written language testing because, as Brown and Yule (1983) state: "spoken language often has incomplete sentences, non-specific words, and may contain relatively little information in any given chunk of language". In practice, the success of a spoken language test depends the ability of the tester to create the right topic, context and questions. In such a case, the students will feel relaxed and confident and the activity will succeed. This study assures on the principle of division and its importance in making the test objective not subjective. Dividing the test criteria and the main six levels into more subdivisions will make the tester more objective and give him more freedom and accuracy in giving the suitable marks for each student. In order to achieve acceptable results, the study tests eight undergraduate students, preparing a real conversation between each two to discuss a topic suitable for their level. This study shows practically the means and procedures by which testers can implement reliable spoken test. The results shows that a test can be successful if a tester; 1) has a good command of the language, 2) is familiar with the principles and criteria of language testing, 3) prepare all the requirements of the spoken test.

**Keywords:** Testing, spoken language, assessment criteria, CEFR.

## تحليل أداء الطلبة في اختبارات مهارة المحادثة باستخدام معيار الدقة

اياد عناد خلف

دائرة التعليم الديني والدراسات الإسلامية/ ديوان الوقف السني

[dr.avadenad@gmail.com](mailto:dr.avadenad@gmail.com)

## الخلاصة:

إن اختبار مهارة المحادثة، مثل أي مهارة أخرى، يجب أن يعتمد على معايير تساعد المختبر على أن يكون دقيقاً في الاختبار. وقد تشتمل هذه المعايير على التسق والدقة والطلاقة والتفاعل والاتساق، كما نص عليه مجلس أوروبا (2001) في النظام المرجعي الأوروبي المشترك (CEFR) من أجل تسهيل عملية القياس ولتكون قاعدة لجميع المختبرين. على الرغم من ذلك، لا يزال اختبار أداء الطلاب في مهارة المحادثة يمثل إحدى المهام الصعبة والأساسية لمعلمي اللغة الإنجليزية كلغة ثانية لتقييم طلابهم، وذلك لأنه يختلف عن الأنواع الأخرى من الاختبارات لسببين؛ أولاً، أنه أكثر صعوبة من حيث التصميم والإدارة والتقييم، وثانياً، أنه لا يلائم بسهولة الافتراضات التقليدية حول الأشخاص والاختبار. من المرجح أن اختبار مهارة المحادثة أكثر تعقيداً من اختبار اللغة المكتوبة لأنه، وفقاً لما ذكره براون ويول (1983): "تحتوي اللغة المنطوقة في كثير من الأحيان على جمل غير

مكتملة، كلمات غير محددة، وقد تحتوي على معلومات قليلة نسبياً في أي جزء معين من اللغة. في الممارسة العملية، يعتمد نجاح اختبار مهارة المحادثة على قدرة المختبر على إيجاد الموضوع المناسب والسياق والأسئلة المناسبة. في مثل هذه الحالة، سيشرح الطلاب بالراحة والثقة وتتم العملية بنجاح. تؤكد هذه الدراسة على مبدأ التقسيم وأهميته في جعل هدف الاختبار عام وليس شخصي. إن تقسيم معايير الاختيار والمستويات الستة الرئيسية إلى أقسام فرعية سيجعل المختبر أكثر موضوعية ويمنحه المزيد من الحرية والدقة في إعطاء العلامات المناسبة لكل طالب. من أجل تحقيق نتائج مقبولة، تختبر هذه الدراسة ثمانية طلاب جامعيين معدة لهم محادثة حقيقية بين كل اثنين لمناقشة موضوع مناسب لمستواهم. أوضحت الدراسة عملياً الوسائل والإجراءات التي يمكن بواسطتها للمختبرين تنفيذ اختبار موثوق به للغة المنطوقة. أظهرت النتائج أن الاختبار يمكن أن يكون ناجحاً إذا كان المختبر (1) لديه إجابة باللغة، (2) على دراية بمبادئ ومعايير اختبار اللغة، (3) أعد جميع متطلبات الاختبار المنطوق بشكل صحيح.

**الكلمات الدالة:** الاختبار، المحادثة، معايير التقييم، النظام المرجعي الأوروبي المشترك.

## Introduction

Testing any skill must depend on principles that will help the tester to be accurate in testing. One of the important principles is that a tester should have a good command of the language and be familiar with the principles of language testing like reliability, validity, consistency, the test taker, construct, task and rating criteria. Spoken language testing is likely to be more complex than written language testing, because "spoken language often has incomplete sentences, non-specific words, and may contain relatively little information in any given chunk of language" (1). In spoken language, speakers use fillers and hesitation marks to give themselves time to speak. Therefore, learners of EFL/ESL (even native speakers) are unlikely to produce language with complete sentences without hesitations and errors unless they have prepared themselves for reading a prepared speech word by word.

Identifying the type of the test is very important for a tester as it will help to determine the kind of questions s/he is going to analyse. In this study, the analysis will be of spoken language test. Eight students were tested. Measurements of the students' spoken performance depend on criteria to put those students at suitable levels according to the CEFR. These levels start from the lowest level to the highest level respectively; A1, A2, B1, B2, C1, and C2.

## Literature Review

State that "the primary purpose of language test is to provide a measure that we can interpret as an indicator of an individual's language ability". Testing means to transfer something abstract like 'ability' or a skill of someone into something concrete i.e. marks. Testing language may be oral or written according to the objectives of the test. Speaking is one of the most difficult language skills to assess since it must be measured in live interaction (8). There is no single system measure that can give fixed results for different test takers, so testing of spoken language is different from one tester to another. It depends on the ability of the testers in measuring the ability of the learners. There must be some principled criteria to depend on in measuring speaking skills. These criteria may include; range, accuracy, fluency, interaction, and coherence, as provided by the (9) in the *Common European Framework of Reference* (CEFR) in order to facilitate the process of measuring and to be a base for all the testers. By using criteria, it is hoped that the gap between different testers will be minimised and this will lead to reliability.

In spoken language testing, learners have to interact with the testers. Oral interaction tasks may be affected by a range of dimensions, like: (2)

- The topic of the interaction.
- The level of formality (informal, consultative, formal).
- The number of the participants.
- The relative status of the participants with each other (strange, acquaintance, friend).
- The gender of the participants.

Each of these dimensions can make the test easier or more difficult, depending on the characteristics of the task and the personal characteristic of the learners (3).

### Statement of the Problem

This study deals with how to make our judgment for testing spoken language a reliable one in the sense that it should give consistent results (every time) for different test takers or in different situations. A detailed rating scale may not be used; instead the tester relied on the CEFR criteria and descriptors. Under these circumstances, testing may be less reliable than when more reliable detailed scales are used. Even when the tester tries to make his decision consistent, another problem will appear that is the differences between groups of testers. There are some principles that a tester should consider during testing a spoken test, these are reliability, validity, consistency, test taker, construct, task and rating criteria.

'Reliability' means "consistency of measurement of individuals by a test"(3).The reliability of a test depends on whether the test would produce the same results if given to the same student at a different time. Reliability therefore depends on the performance of the candidate as well as the accuracy and consistency of the marking. Thus, 'reliability' refers to the consistency of scores from one set of tests and test to another"(2). In this study, the reliability of giving credits to students will depend on dividing the ability of every one into five sub-divisions to reduce the differences between different testers or the same tester's results on another time.

'Validity' refers to "the relationship between evidence from test performance and inferences drawn from that evidence about candidates' capacity to perform in the criterion" (3). In other words, the validity of a test depends on whether it actually tests what it claims to test. This may depend on the quality and quantity of the questions of a test whether they were sufficient to cover all the abilities that required to be tested or not. In this study, as the aim of the test was to test spoken production and spoken interaction, the tester divided his questions into three parts; the first part was direct questions for the students about general subjects and he tried to made them to speak as long as possible. Then, the second part was to make the students interact in a discussion to see their ability in initiating different types of dialogues or in turn taking. After that, in the third part, he gave them some suggestions and asked them again to discuss the advantages and the disadvantages of these subjects. In this way, the tester covered a range of abilities that required to be tested to reach to an accepted validity.

It is necessary in any test to identify the ability to be tested in order to achieve the validity in a test. These abilities are called constructs, which should be statements of the skill that is required to carry out for a spoken task. If a test measures the construct it claims to measure, the test will have what is called construct validity. In this study the constructs of the students' test is the ability to speak about definite subjects related with English learning like; memory, accuracy, fluency, pronunciation and grammar. The test takers are students of English department so, these subject are familiar with these topics.

The last important principle to be considered is 'task and rating criteria'. There are two kinds of rating scales; an analytic rating scale that describes the marks to be given to the test taker for different levels of performance according to a number of criteria, and the holistic rating scale that describes the marks to be given to the test taker for different levels of performance taken as a whole. This study will use the analytic rating scale which will describe the marks will be given for the students depending on their ability in speech (see next part).

### **Analysis and Discussion**

In analysing a spoken language test for each student, the same criteria will be used. These criteria are; range, accuracy, fluency, interaction and coherence that are provided by CEFR. Range means that the words will be classified according to the simplicity and complexity of every word used by the student. In other words, it refers to the ability of using simple or complex nouns, verbs, adjectives, adverbs or sentence models. Accuracy is related to the control of grammatical structures and using these structures appropriately, and also involves knowing how to pronounce words correctly. Accuracy refers to "how well the target language is produced in relation to the rule system of the target language"(4). Fluency is the ability to speak in a normal speed spontaneously without too much hesitation or too many pauses. (5) define 'fluency' as "the production of language in real time without undue pausing or hesitation". Interaction also is important, it refers to the ability to start or to stop a conversation, take turns when suitable time, keeping the conversation going, invite others in a dialogue. Finally, coherence refers to the ability of the speaker in using connectors and cohesive devices to link utterances into clear and logical discourse structure. This also includes the coherence of the meaning for the conversation to be one unit form.

This study concentrates on the accuracy criteria, which can be divided into smaller parts like, tense structures, word order, pronunciation, vocabulary and sentence grammar. This division will help in measuring the ability of each student in their command of these parts of speech. When assessing pronunciation, testers need to be conscious of the articulation of words and longer stretches of language in discourse rather than in isolation(6). In assessing grammar, we should consider grammatical knowledge as "when examinees understand or produce utterances that are grammatically precise and contextually meaningful"(7).

Deciding the level of every test taker in this study will depend on a number of steps. Firstly, the spoken language of every test taker will be analysed depending on accuracy criteria. The analysis will include the ability of the students to create accurate utterances grammatically and also the ability of pronouncing these sentences. Secondly, depending on the analysis, each student will be given credits according to his ability. Thirdly, the credits of every student will be put in a scale of levels, which is prepared already, to be given the suitable level.

### **Step 1: Spoken language analysis:**

In this step the spoken language of all students will be analyse to see their strong points and weakness in their ability of speaking production and speaking interaction and this analysis will examine in detail the following five things; tense structure, word order, grammar, vocabulary and pronunciation for every student.

## STUDENT (1):

### Tense structure and word order

The analysis showed that she used basic sentence patterns and she had a few words. She spoke very little about 160 words and most of her sentences are uncompleted for example:

- *she is studying....*
- *and she knows that I can....*
- *I m not fine because.....*
- *accuracy is important to listen and understand .....*

She used only the present simple and past simple tense in most of her sentences for example:

- *Memory part of my brain.*
- *I was so excited.*
- *The first friend is Ceyda.*
- *she makes me and she knows that I can.*

### Grammar and vocabulary

Student (1) formed few questions and all of them were grammatically incorrect for example:

- *Did you like here?*
- *How is your first day in university? ( without; the, this)*

Therefore, she tried to repeat the questions of her friend saying " *what about you*". This showed her weak ability in initiating a dialogue. Her answers for the questions were sometimes unclear because she used unsuitable vocabularies and very short sentences for example:

- *Its provides to this understand and to reading, writing.*
- *"the university was insection.*
- *I tell to improve this.*
- *Accuracy is important to listen and.*

Her grammar in general was poor and she made grammatical mistakes like:

- *Memory part of my brain.*
- *It is provides to this understand.*
- *It may be can nice be nice.*
- *Accuracy is grammatical correctly words.*

## Pronunciation

The pronunciation of student (1) was good and she could pronounce the words correctly. The pauses between words were few but she had many long pauses at the end of her sentences because she did not have enough vocabularies to complete her speech.

## STUDENT (2):

### Tense structure and word order

The analysis showed that he used basic sentence patterns and he had a few words. He spoke about(344) words. Most of his sentences were completed and he used different types of tense structure like the present simple, past simple tense and present perfect tense in most of his sentences. This led to clarity in most of his speech for example:

- *my memory is err not bad I can easily remember the numbers.*
- *Isparta is a new place for me*

- *I like university because this university presents us so many advantages like library like tickets...*

He had very few pauses but he made a lot of repetition during his speech like:

- *grammars in is important than important for speaking correctly.*
- *Because when I am when I was I am speaking I am so excited.*
- *because thanks to thanks to accuracy.*
- *my words can easily can be easily understood.*
- *what's your, what's your idea?*
- *repeating rules are so many in high in high class.*

### **Grammar and vocabulary**

Student (2) made few questions and most of them were grammatically correct for example:

- *How was your first day at university?*
- *who was your first friends at university?*
- *what do you think about practising with friends?*

Although he made some mistakes in making questions but they were very few like:

- *So, what you think of the advantages and disadvantages of repeating rules?*

In general, he had a good command of grammar. Even when he made some mistakes, he tried to correct them by repeating the words again. His mistakes were simple and did not cause any misunderstanding from the listener. His communication was good. He could initiate questions and help his partner in her pauses.

### **Pronunciation**

The pronunciation of student (2) was good and he could pronounce the words correctly. The pauses between sentences were very few but he had many short pauses between words in general because he was speaking very slowly. In addition, he had a lot of repetition in his speech.

### **STUDENT (3):**

#### **Tense structure and word order**

The analysis showed that she used basic sentence patterns and she had few words. She spoke more than her partner student (4) did, about (485) words, but many of her sentences were not understandable for example:

- *I remember important for men for example mother telephone numbers.*
- *Nigde university er is er famous er only doors.*
- *my family erlifes city.*
- *Grammar is sentence of combination er grammar is a person learn grammar.*

She used the present simple and past simple tense in most of her sentences for example:

- *memory is par er part of brain.*
- *memory is important er for people.*
- *I went to Nigde with my boyfriend*

In general, she used simple structures and words. She could make her partner to understand her speech although there were many pauses and mistakes in arranging words in a sentence. She could link groups of simple words with simple connectors like and, but, and because. Nevertheless, sometimes her speech lacks of cohesion and coherence. She did not use cohesive ties like conjunctions and this led to ambiguity in some sentences in her speech, for example:

- *I have got good memory er I remember for er important for I remember important for men...*
- *Grammar is sentence of combination er grammar is a person learn grammar if a person if a person learn language you er a person must learn grammar.*
- *for example you know I when I I am speaking English I can er I can't er I er I can my sentence haven't rulers haven't rulers...*
- *I speak difficult English or language.*

#### **Grammar and vocabulary**

Student's (3) speech had many well-formed sentences but she made some mistakes in some sentences, like:

- *I really I want go to er swim course.*
- *I can my sentence haven't rulers.*
- *I say them I I say can I say things them.*
- *OK what did you er last holiday?*

Even when she tried to repeat her phrases to correct them, she still made mistakes. This showed her weak ability to initiate a dialogue or create a stretch of sentences.

- *I remember for er important for I remember important for men for example mother telephone numbers and er computer er computers and memory is important er for people er because erer if er because er ..*

#### **Pronunciation**

The pronunciation of student (3) was good in general but she made many mistakes in pronouncing some words that led sometimes to unclear sentences. For example, she pronounced the following words as it shows down:

- *Then / dan/, with /wid/, live/ laif/, famous/ fimas/, city/ sita/, holiday/ holdei/, want/want/, rules/ rulas/, speaking/ sepi:kng/.*

Moreover, she had many pauses between words and some long pauses at the end of her sentences because she didn't have enough vocabularies to complete her speech. The pronouncing of some words, sometimes led to unclear sentences, for example:

- *my uncle lifesmarmariser.*
- *Becauseer my family erlifes city.*

#### **STUDENT (4):**

##### **Tense structure and word order**

The analysis showed that he used basic sentence patterns and he had very few words and he could not make even one complete sentence. He spoke very little about 185 words and most of his sentences were uncompleted. He could not answer many questions, for example:

- *I remember for I remember something ....*
- *I have money my family always .....*
- *Grammar speaking ... er ... grammar is ....*
- *I think er learn rules is disadvantage because ..... because because we...*

He used only the present simple and past simple tense in all his sentences for example:

- *Memory er I have bad memory (laughs)*
- *I wake up I woke up and er I had breakfast*
- *I have peace with my family er I I can do something*

##### **Grammar and vocabulary**

Student (4) formed few questions but all of them were grammatically correct, for example:

- *what did you do last holiday?*
- *How was it?*
- *What did you watch?*

His grammar in general was good and he did not commit many grammatical mistakes.

### **Pronunciation**

The pronunciation of student (4) was good and he could pronounce the words correctly. The pauses between words were few but he had many long pauses at the end of his sentences because he did not have enough vocabularies to complete his speech.

### **STUDENT (5):**

#### **Tense structure and word order**

The analysis showed that he used basic sentence patterns and he had a few words. He spoke about (390) words. Most of his sentences were not well structured and he used very simple types of tense structure like the present simple, past simple tense in most of his sentences. His lexical background is very weak. This led to a lack of clarity in some of his sentences for example:

- *for example er some er games in the in the playing village.*
- *we can entry in the exam and conclusion it was best.*
- *Definition is important for speaking a language.*
- *advantages is we can we can understand er together.*

He had very few long pauses because he lacks of many important words, but he made few repetitions during his speech.

#### **Grammar and vocabulary**

In general, student (5) had weak ability in forming well sentences. When he made some mistakes, he tried to correct them by repeating the words again. His mistakes were simple and did not make any misunderstanding for the listener. He could form some correct phrases and questions but he formed some incorrect questions, for example:

- *say me please?*
- *Can you say me advantages?*
- *Do you listen radio programme in the English?*
- *You think is there advantages listen to radio in English?*

### **Pronunciation**

The pronunciation of student (5) was not very clear. He could pronounce the words correctly except for some words like:

- *Child /tʃild/, won /von/, here /hə/, think /tink/, now /nəu/, exam /i:gzam/.*

He had many short pauses between words in general because he speaking very slowly. In addition, he had a lot of repetition in his speech.

### **STUDENT (6):**

#### **Tense structure and word order**

The analysis showed that she used basic sentence patterns and she had very few words. She spoke more than her partner did, about (375) words, but she made a lot of repetition in her speech. Some of her sentences were uncompleted and she did not answer some questions, for example:

- *Yes yes I remember I can remember my ...*
- *Er pronunciation is er that's all. (no answer for the question)*



- (laughs) memorising ... (no answer for the question)

She used only the present simple and past simple tense in most of her sentences for example:

### Grammar and vocabulary

Student (6) formed only one question and it was grammatically correct like:

- *do you remember the best friend in your high school?*

Her grammar in general was poor. She made some grammatical mistakes. Her answers for the questions were sometimes unclear because she used unsuitable vocabularies and very short sentences, for example:

- *Er remember brain about all of memory er memory is ... remembering (laughs).*
- *because I can meet each other.*

She also did not use connectors to join many of her sentences, like:

- *I remember I can remember my ... I suppose I don't have good memories (laughs). I am I am excited er first day at high school.*

### Pronunciation

The pronunciation of student (6) was good and she could pronounce the words correctly except for some words like:

- *Think /tink/, disadvantage /dizadv.../.*

The pauses between words were few but she had many long pauses at the end of her sentences because she did not have enough vocabularies to complete her speech or to answer the questions.

### STUDENT (7):

#### Tense structure and word order

The analysis showed that he had sufficient range of language to be able to give clear descriptions. He could express his views about memory, fluency. He used basic sentence patterns and he had a lot of words. He spoke about (530) words. Most of his sentences were completed and he used different types of tense structure like the present simple, past simple tense, future simple, and present perfect tense in most of his sentences. This led to clarity in most of his speech.

He had no pauses and he rarely made repetition during his speech. Even when he made a repetition, he tried to correct his mistakes, like:

- *I don't like politics political books.*

### Grammar and vocabulary

In general, he had a good command of using grammatical structures. Even when he made some mistakes, he tried to correct them by repeating the words again. His mistakes were simple and did not cause any misunderstanding from the hearer. His communication was good. He could initiate questions and helped his partner in her pauses. He made few questions and most of them were grammatically correct for example:

- *Is there any book that we read it common?*
- *What do you think is it better to practice with your friends or memorise the words?*

He made very few mistakes in making questions. Most of his mistakes were in adding or omitting the letter (s) of plural or the definite article (the) to some words, like:

- *I think Harry Potter the books I just read it last week.*
- *it has lots of interesting.*
- *I'm the one of the character(s) of the book.*

- fluency is the talking without gaps.
- I have more fluency speaking.

### Pronunciation

The pronunciation of student (7) was very clear and he could pronounce the words correctly. He also could produce a stretch of language fluently, but during his speech, he gave an incoherent sentence in meaning which led to contradiction, like:

- I have to learn some old words which is used to use in old ages so modern novels is about our days I think it's better to read our days novel.

But, generally, most of his speech was good and clear.

### STUDENT (8):

#### Tense structure and word order

The analysis showed that student (8) had enough language and sufficient vocabulary to express simple subjects like memory or fluency. He used basic sentence patterns and he had well stock of lexical words. He spoke about (290) words. Most of his sentences were completed. He used tense structures like the present simple and past simple tense in all of his sentences. Most of his sentences were clear. He had very few long pauses, but he made a lot of repetition during his speech like:

- it may depend it may change because of my situation.
- because I don't ..I don't interested in magazine.
- because I don't... I don't practice with my friends.

#### Grammar and vocabulary

Student (8) formed few questions, some of them were grammatically correct for example

- about which countries: our country or foreign countries?
- What do you think about practising with friends?

But he had some mistakes like:

- what do you like kinds of music?

In general he had good command in using grammatical sentences. Even when he made some mistakes, he tried to correct them by repeating the words again. His mistakes were simple and didn't cause any misunderstanding from the hearer. His communication was good. He could initiate questions and communicated with his partner.

### Pronunciation

The pronunciation of student (8) was good and he could pronounce the words correctly, except some words like:

- sing / sink/, with / wid/

He made some short and long pauses between sentences. He spoke very slow.

### Step 2: Scale of Credits

According to the analysis of the spoken language of the students, and depending on the accuracy scale of CEFR, the ability of each student in command tense structures, word order, pronunciation, vocabulary and sentence grammar will be shown in credits. Table, no (1) shows the credits that were given for his/her ability.

Table(1):Credits of Students for Accuracy Scale

Scale	Word order20%	Tense structure20%	Pronunciation 20%	Sentence grammar20%	Vocabulary 20%	Total 100%
Student (1)	12	11	13	9	11	56
Student (2)	12	12	12	13	11	60
Student (3)	12	10	8	9	12	51
Student (4)	8	10	11	10	8	47
Student (5)	10	9	8	9	10	46
Student (6)	9	9	11	11	11	51
Student (7)	16	14	16	14	17	77
Student (8)	12	15	12	12	12	63

As the table shows, the accuracy scale of testing spoken language can be divided into five sub-scales. Each one of these sub-scales covered 20% of the total scores. The next step and according to another division of the six levels of the CEFR, every student will get his suitable level, A1, A2, B1, B2, C1, C2. To match every student with an approximate level, these six levels will be sub-divided into more divisions. Table (2) shows how these levels are divided into smaller parts.

Table(2): Sub-division of CEFR Levels

Marks	Sub- levels	Levels
95-100	C2.2	C2
90-94	C2.1	
85-89	C1.2	C1
80-84	C1.1	
75-79	B2.3	B2
70-74	B2.2	
65-69	B2.1	
60-64	B1.3	B1
55-59	B1.2	
50-54	B1.1	
45-49	A2.3	A2
40-44	A2.2	
35-39	A2.1	
30-34	A1.3	A1
25-29	A1.2	
20-24	A1.1	

As the table above showed, levels A1, A2, B1 and B2 were subdivided into three levels. These levels have been compared with a suitable credit from 20-100%. Any score lower than 20 will be ignored taking in account that any student who would take less than 20%, is outside this scale.

### Step 3: Giving the Suitable Level.

As a result from the steps of the analysis, the criteria and the conformity between the credits and the levels the students were put in the following levels:

Table (3): Final Results

Name	Credits	Level
Student (1)	56%	B1.2
Student (2)	60%	B1.3
Student (3)	51%	B1.1
Student (4)	47%	A2.3
Student (5)	46%	A2.3
Student (6)	51%	B1.1
Student (7)	77%	B2.3
Student (8)	63%	B1.3

## Findings

As Table (3) showed, each student was put in his/her suitable level. For example, student (2) was put in level B1.2 because he had enough language to get by and sufficient vocabulary to express self with some hesitation on topics such as memory or to speak about the first day at the university. His accuracy was not bad as he used reasonably accurately a repertoire of frequently used routines and patterns associated with more predictable situations. He also could keep going comprehensibly, even though pausing for grammatical and lexical planning and repair was very evident, especially in longer stretches of free production. He could initiate, maintain and close simple face-to-face conversation on topics that are of familiar or of personal interest. He could repeat back part of what someone had said to confirm mutual understanding. Finally, he could link a series of shorter, discrete simple elements into a connected, linear sequence of points.

Another student, for example was student (7). He was put at level B2.3 because he had a sufficient range of language to be able to give clear descriptions; he made good expression viewpoints on most general topics, like a book he had read or about memory, without much conspicuous searching for words, using some complex sentence forms to do so. His command for grammar was good and showed a relatively high degree of grammatical control. He did not make errors that cause misunderstanding, and could correct most of his mistakes. He could produce stretches of language although he was hesitant because he was searching for patterns and expressions to complete his speech. There were very few noticeably long pauses in his speech. He could initiate discourse and took his turn when appropriate and ended conversation after he gave sufficient description. He was helpful for his partner during the discussion. Finally, he could use a limited number of cohesive devices to link utterances into clear, coherent discourse. For all these reasons, he was given B2.3 level.

## Conclusion

Although speaking is a difficult language skill to assess since it must be measured in live interaction, the use of specific criteria makes the evaluation of a spoken test easier for the testers and can lead to reliability. As this study showed, the use of scales increases reliability in the test of this study through number of steps starting from the analysis of spoken language, then measuring the ability of test takers by giving them credits and finally, making a balance between the credits obtained by students and the six levels of CEFR. This process also requires dividing the levels into six levels, and then these levels need in turn to be sub-divided to accommodate differences between test takers. This study assured on the principle of division and its

importance in making the test objective not subjective. Dividing the test criteria and the main six levels into more subdivisions will make the tester more objective and give him more freedom and accuracy in giving the suitable marks for each student. It should be taken in the consideration that the tester should be familiar with all the skills of language and to be familiar with skills and principles of language testing like reliability, validity and consistency to be able to give the right level for test takers.

#### **CONFLICT OF INTERESTS**

**There are no conflicts of interest**

#### **References**

1. Brown, G. & Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.
2. Lyle F. Bachman & Adrian S. Palmer. 1996. *Language Testing in Practice*. Oxford: OUP.
3. McNamara, T. F. (2000). *Language testing*. Oxford: Oxford University Press.
4. Skehan, P. 1996b. 'Second Language acquisition research and task-based instruction' in J. Willis and D. Willis (eds.): *The Challenge and Change in Language teaching*. Oxford: Heinemann.
5. Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
6. McKay, P. (2006). *Assessing Young language learners*. Cambridge: Cambridge University Press.
7. Purpura, J. (2004). *Assessing Grammar*. Cambridge: Cambridge University Press.
8. Luoma, S. (2004). *Assessing Speaking*. Cambridge: Cambridge University Press.
9. Council of Europe. (2001). *Common European Frame Work of reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.