

EFFECTS OF AN ASSESSMENT GRID ON THE ORAL PERFORMANCE OF  
STUDENTS FROM THE LANGUAGE CENTER AT THE UNIVERSITY OF  
NARIÑO

by

Diego Fernando de la Portilla Guerrero

Submitted to the Faculty of Human Sciences  
in partial fulfillment of the requirements for  
the degree of B.A. in the Department  
of Linguistics and Languages  
University of Nariño

2006

ACKNOWLEDGMENTS

"School improvement is more surely and thoroughly achieved when teachers engage in frequent, continuous and increasingly concrete talk about teaching practices... capable of distinguishing one practice and its virtue from another"  
Judith Warren

"No great improvements in the lot of mankind are possible until a great change takes place in their mode of thought"  
John Stuart Mills

I wish to thank teacher Ana Clara Sánchez Solarte for her guidance, advice and encouragement throughout this project. Dear teacher, this project would not have been possible without your help. It didn't matter how much work you had to do, you were always there with a big smile for me and a clear idea for my project. I admire you because you have taught me that being a teacher can be as exciting as any other job. You're the best and you know that! Thanks a lot.

I also would like to thank teachers Gabriel Vicente Obando Guerrero and Noé Vicente Araujo Quintero. Teacher Gabriel I have to say that I had discovered in you a great person, a teacher interested in his students and a person who is always willing to help. Best wishes for you!

Teacher Vicente, I will always remember you for your sense of humor, your knowledge of English and for being my first teacher at the university. Best wishes for you too!

Thanks to all people who had supported me during my career, to all those teachers, friends, and people who made all this five years at the university a memorable time in my live.

Thanks to my parents who had helped and encouraged me all the time. Mom and Dad, I always will be in debt with you two. I love you so much and I hope you are proud of me.

To my daughter and her mother who had been there always for me. I've tried my best during all this time, and it has been all for you. Las quiero mucho!

*TABLE OF CONTENTS*

LIST OF FIGURES.....	viii
ABSTRACT	
I. CHAPTER 1: INTRODUCTION TO THE STUDY	
Introduction to the problem .....	10
Problem Statement .....	11
Research Question .....	14
Hypothesis .....	14
Justification .....	15
Objectives .....	16
General Objectives.....	16
Specific Objectives.....	16
II. CHAPTER 2: LITERATURE REVIEW	
Evaluation, Assessment and Testing .....	17
Evaluation .....	18
Assessment .....	18
Testing .....	18
Reasons for testing .....	18
Finding out about Progress .....	19
Encouraging Students .....	19
Finding out about Learning Difficulties .....	19
Finding out about Achievement .....	20

Placing students .....	20
Selecting students .....	21
Finding out about proficiency .....	21
Characteristics of a Reliable Test .....	22
Practicality .....	22
Reliability .....	22
Student-Related Reliability .....	22
Rater Reliability .....	23
Test Administration Reliability .....	23
Test reliability .....	23
Validity .....	24
Content Validity .....	24
Criterion-related validity .....	24
Construct-related validity .....	25
Consequential Validity .....	25
Face Validity .....	26
Authenticity .....	26
Washback .....	27
Test Design .....	27
Objectives .....	28
Skills .....	28
Content .....	28
Format .....	29
Rubric .....	29

Materials .....	29
Marking .....	29
Communicative Language Teaching .....	32
Communicative Competence .....	34
Types of tests in CLT .....	37
Oral Assessment .....	39
Elements involved in oral assessment .....	41
Fluency .....	41
Accuracy .....	41
Formats used to assess oral production in Communicative approaches .....	42

### III. CHAPTER 3: METHOD

Design of the study .....	43
Procedure .....	43
Subjects .....	45
Materials .....	45
Instruments .....	45
Variables .....	45
Data Gathering and Analysis .....	46

### IV. CHAPTER 4: RESULTS

Results .....	49
---------------	----

V. CHAPTER 5: DISCUSSION AND CONCLUSIONS

Discussion .....55  
Conclusions .....60

REFERENCES

APPENDIX

*LIST OF FIGURES*

POLYGON OF FREQUENCIES

Students' performance First Oral Exam.....	68
Students' performance Second Oral Exam .....	69
Students' performances' comparison First and Second Exam) .	70
Students' average First Oral Exam.....	71
Students' average Second Oral Exam .....	72
Students' averages' comparison (First and Second Exam) ...	73
Students' performance Second Oral Exam (More common categories in terms of Accuracy and Fluency).....	74
Students' answers' percentages in the Questionnaire .....	75
Students' Total percentages in the Questionnaire .....	76



ABSTRACT

This research was carried out in order to know the influence of the use of a grid during an oral test among students. At the beginning of this research it was decided that apparently students perform in the same way no matter the use or not of a grid. Two oral tests were presented to the students of First level of English at the Language Center in the University of Nariño. Only during the second test some criteria were arranged with the students. Students' results were compared in order to know the influence of the grid in their performances. A questionnaire was also presented to them. Ninety percent of the students demonstrated through their answers a positive view towards the use of the grid. Eighty percent considered that the second test appeared to be less threatening to them and all students agreed that when the criteria selected to assess oral production is arranged in advance, their grades can be improved. Students improved their performance during the second exam but their grades were consistent with the first ones they obtained, proving that both tests were considerably well structured following the directions of important authors in language teaching such as Brown, Underhill, Weir, Heaton, Nunan, and so forth.

CHAPTER 1. INTRODUCTION OF THE STUDY

*Introduction to the Problem*

English Language Teaching (ELT) is a field of education that is constantly changing due to diverse factors involved with the conception of language. In order to accomplish the objectives of this paper, it appeared to be useful to consider how languages are taught - or at least how foreign language teaching is conceived - in an EFL setting like Pasto and which are the trends teachers follow.

In order to achieve the best performance of students in terms of communication, some teachers have selected Communicative Language Teaching (CLT) as the framework for developing their professional practice. CLT is a popular approach that offers interesting insights about language. However, this paper is focused only on one important aspect of CLT, which is assessment. How can a teacher establish who is communicatively competent? How does s/he support those selections?

Teachers tend to use oral exams as means to corroborate their assumptions about students' communicative competence. When exams are carried out, many elements related to test design come into play. However, one important aspect in tests' design appears to be ignored by teachers: that is marking. Teachers may incline to assess in a subjective way

and this might provide wrong information about the communicative competence of a student. Harrison (1983 p. 110-117) considers that what makes a test subjective is the way in which they are marked, because of the range of possible correct answers.

In terms of test design, marking is an important factor that must be considered. However, as it can be deduced from Harrison (1983), a good way to avoid subjectivity in oral tests is a well structured and theoretical-based marking that can measure specific aspects of language. Basically, what this paper intended was to offer students and teachers the design of a theoretical and experimental grid that could be used to measure in an objective way the oral performance of students during a test.

#### *Problem Statement*

Assessment is a very complex process that involves many different considerations. As it is stated by Heaton (1990 p. 9-22) one important aspect in assessment is the reason why a test is carried out. He presents a list of reasons that can be increased if each one of us were asked. However, some of those motives are concerned with checking students' progress. Heaton (1990) calls this "progress tests" (Heaton 1990 p. 9).

Teachers use tests to measure the quantity or the quality of what has been taught. As Heaton states, this is only one part of tests and some times the most negative but essential one. (1990 p. 13)

He also offers a reasonable clarification about the difference between progress and achievement tests when he states that: " Unlike progress tests, achievement tests should attempt to cover as much of the syllabus as possible". (Heaton 1990. p.14)

This means that teachers have a tendency to assess progress - short periods of learning - instead of achievement. Achievement tests are, according to Heaton, formal tests carried out at the end of a course or a scholar year, and these kinds of tests are liable to assess what it is expected to be taught but what has not necessarily been taught. (Heaton 1990 p. 14)

However, achievement tests attempt to cover longer periods of learning than progress tests, making them suitable for the purposes of this research. According to Brown (2004 p. 47-48) achievement tests do take into account students background, because they are based on the syllabus, or the curriculum in order to be applied. Moreover, achievement tests are focused on the objectives presented in those syllabuses or curriculums.

Heaton (1990 p. 14) considers achievement tests as some practices in which students' background is not taken into account. He refers to the TOEFL test as a kind of achievement test, and differentiates proficiency tests in terms of purposes. According to Heaton (1990 p. 17) proficiency tests aim for measuring the quality of candidates to be selected for certain course or task. Though Heaton's considerations are relevant, Brown's view is more closely related to the main objective of this paper and that is why his ideas have been selected as the basis for this paper.

Selecting a certain type of test is only the first step in the construction of a good test. Once teachers have decided which type of test is going to be used, some other aspects come into view. Now it is time to decide what and why to assess learners. However, this decision engages some other aspects, like marking.

Once a teacher has decided what to assess, why and in which form, s/he must consider how s/he is going to score or mark students' answers and what aspects of language are going to be assessed. As it has been explained before, subjectivity may constitute a considerable problem when marking our learners' exams and when issuing our opinions about their communicative competence. Thus, this paper

intended to implement a grid for reducing subjectivity and for improving the quality of tests administered to students. Heaton (1990 p. 68) proposes a well structured marking or rating scale and though this author is inclined to assess oral production through a rating scale, Underhill (2000 p. 94-103) proposes a system that involves both a rating scale and marks. This author introduces a significant element in oral production assessment, what is called a grid or chart for rating and marking students' performance. Underhill (2000 p. 95) combines a rating scale that is according to Heaton (1990 p. 69) "a short description for a grade in a scale", with a marking system. Basically, the grid is going to be an instrument that incorporates categories or descriptions of language performance and a grade or mark that represents that performance.

*Research Question*

Does the use of a grid contribute to improving the performance of students when they are orally assessed?

*Hypothesis*

For this research a null hypothesis was used:

The use of a grid has neither positive nor negative effects in the students' performance when they are assessed orally.

*Justification*

Teachers always carry out assessment, and it may be one of the central activities teachers perform in their professional practice. Assessment though, may be perceived as threatening or stressful by students and teachers sometimes lack the knowledge or the time to create grids that facilitate and enhance assessment. Creating a grid is important because they save teachers time and they might serve as tools for avoiding subjectivity in the marking and grading of oral exams.

Although they appear to be only useful for teachers, grids are useful for students too. Weir (1995 p. 25-26) explains that both teachers and students must know what aspects are going to be assessed and how those aspects are going to be evaluated. This might foster the students' motivation and make tests look less threatening for learners.

Grids also can be used in order to perform an oral assessment practice that can lead to a transparent, objective and enjoyable experience for the teachers and for his/her learners.

Another aspect that justifies undertaking this research is that creating and using a grid can save time and specify almost all language aspects while it is carried out.

The validity of the test can also be positively affected if it is combined with a grid, since it can show that the test measures what it is intended to measure; the design of a marking-rating scale system - a grid - is an option that must be applied in our setting to observe its results and for informing other teachers about the potential advantages of this tool.

### *Objectives*

#### *General Objective*

To determine the effects of applying a grid designed within our context on the students' performance during an oral exam.

#### *Specific Objectives*

To compare the results obtained by the students when they know what aspects of language are assessed and when they are not given criteria before the assessment.

To design a grid based on current literature about evaluation, assessment and testing.

To create an instrument that is contextualized to the setting and the characteristics of the students with whom we interact.



II. CHAPTER 2: LITERATURE REVIEW

*Evaluation, assessment and testing*

As it can become confusing to interpret the differences of these three terms, it is necessary to define each one of them in order to avoid using them as equivalent words.

*Evaluation*

Richards (1994 p. 17) argues that evaluation is concerned with determining the effectiveness of a program. It means that evaluation can be considered as tool through which teachers can decide whether or not a program is desirable for obtaining certain objectives. Richards (1994 p.17) also considers that evaluation means to verify the way in which a program works, and he supports his idea by quoting Popham (1975) and Jervis and Adams (1979):

Evaluation is concerned with gathering data on the dynamics, effectiveness, acceptability, and efficiency of a program to facilitate decision making. (Popham 1975; Jervis and Adams 1979 quoted in Richards J.C. 1994. *The Language Teaching Matrix*. New York: Cambridge University Press.)

Moreover, Worthen and Sanders (1973 p. 19 quoted in Richards 1994 p. 17) believe that:

Evaluation is the determination of the worth of a thing. It includes obtaining information for use in judging the worth of a program, product, procedure, or objective, or the potential utility of alternative approaches designed to attain specific objectives. (Worthen and Sanders 1973 p. 19 quoted in Richards J.C. 1994. *The Language Teaching Matrix*. New York: Cambridge University Press.)

*Assessment*

Brown (2004 p. 4) argues that this term is in some cases misunderstood and confused with testing. He refers to assessment as:

"... an ongoing process that encompasses a much wider domain. Whenever a student responds to a question, offers a comment, or tries out a new word or structure, the teacher subconsciously makes an assessment of the students' performance. (Brown H.D. 2004. *Language Assessment. Principles and Classroom Practices*. White Palms: Pearson Education)

*Testing*

Brown (2004 p. 4) also states that tests are only a part of assessment. He believes that:

Tests,...are a subset of assessment; they are certainly not the only form of assessment a teacher can make. (Brown H.D. 2004. *Language Assessment. Principles and Classroom Practices*. White Palms: Pearson Education)

He also considers that tests are "useful devices" a teacher can use to assess students, but those devices are only a part of many different kinds of practices that can be carried out in order to assess students' performance.

At this point, it is necessary to take a look of some reasons that can be established when a test is placed.

*Reasons for Testing*

Heaton (1990 p.9) presents some reasons for testing that should be considered when developing this research. He believes that a test is carried out according to a purpose. The differences that can be found among those purposes

create a reason for a test to be prepared. He categorises seven different reasons in which tests can be classified.

*Finding out about progress*

According to the author this is the most common reason for a test to be carried out. In progress tests teachers measure short periods of learning and the way in which students have mastered the aspects of language that have been taught recently. (Heaton 1990 p. 9)

*Encouraging students*

Tests can serve as a means for increasing students' motivation towards their own learning process. Heaton (1990) claims for tests to be a tool that can be used for students to check their progress and the goals they have achieved.

*Finding out about learning difficulties*

Heaton (1990 p. 11) states that tests can guide teachers to identify students' weaknesses and necessities. To identify the problems students can have in the learning of English is the first step in correcting or giving appropriate information to the learners. To find difficulties and weaknesses among the students can serve to evaluate the teaching process as well as the syllabus and the materials used in the course (Heaton 1990. p. 13)

*Finding out about achievement*

At this point it is essential to consider the ideas established by Brown (2004) and McNamara (2000) who consider achievement tests in a different view of that of Heaton. While Heaton (1990 p. 14) claims for achievement tests as those kinds of tests that measure the level in which students have mastered what have been taught or what should be taught during a course, Brown (2004 p. 48) and McNamara (2000 p. 7) understand achievement tests as those tests that are constructed taking into account the goals of the course as well as the process (McNamara 2000) and the particular time in which a teaching-learning process has been developed (Brown 2004). This means that according to Brown and McNamara achievement tests do take into account students' background as an important feature in this kind of exams. On the other hand, Heaton's perception of achievement tests does not include students' background, therefore, according to him (Heaton 1990 p. 14) the Test of English as a Foreign Language (TOEFL<sup>®</sup>) is an achievement test. However, Brown (2004 p. 45) considers that this exam (TOEFL<sup>®</sup>) is a typical proficiency test.

*Placing students*

Tests are according to Heaton (1990 p.15) a useful tool to establish groups according to the students' "language

ability". Brown (2004 p.45) considers that placement tests are used to place students in a certain level according to their performance taking into account the materials that are going to be covered in a course.

*Selecting students*

Tests can be used to decide which student or candidate is qualified to complete certain type of task.

Heaton (1990 p. 16) presents this purpose as one in which many candidates' performances are compared in order to select the best one.

*Finding out about proficiency*

Brown (2004 p. 45) considers proficiency tests those tests in which no attention is paid to the language a student has learned during a course, or the goals achieved through a syllabus, or even to the curriculum developed during the teaching process. Proficiency tests measure "overall ability" (Brown 2004 p. 45). For these reason, it can be said that proficiency tests do not take into account students' background.

Moreover these reasons, McNamara (2000 p. 68) presents tests as a means of "social and cultural exclusion" and as a tool to accomplish educational and social policies. Whatever reason is proposed, tests should be constructed in a reliable way.

*Characteristics of a reliable test*

According to Brown (2004 p. 19) there are five major characteristics that can identify a reliable tests. These characteristics are: practicality, reliability, validity, authenticity and washback.

*Practicality*

According to Brown (2004 p. 19) practicality in testing means that the test:

- is not expensive.
- stays within appropriate time constraints.
- is relatively easy to administer, and
- has a scoring/evaluation procedure that is specific and time-efficient. (Brown H.D. 2004. *Language Assessment. Principles and Classroom Practices*. White Palms: Pearson Education)

*Reliability*

Tests that are reliable are "consistent and dependable" (Brown 2004. p.20). This means that the results obtained by the students when developing the same exam in different periods of time should give similar scores. However, there are some factors that can affect the reliability of a test. Brown (2004) presents some of them.

*Student-Related Reliability*

It refers to the conditions of a student while carrying out a test. Those conditions are related to a "physical or psychological factors" such us "temporary illness, fatigue

and anxiety" that can decrease the performance of a student during a test. (Brown H.D. 2004. p. 21)

#### *Rater Reliability*

Brown (2004 p.21) suggests "human error, subjectivity and bias" as factors influencing scoring process. He presents two different rater-reliability categories: the inter-rater reliability referring to the inconsistency of two or more scores in the same test; the intra-rater reliability referring the way in which the teacher award scores to the students. The last one can be affected by bias and unclear scoring procedure. The use of "an analytical scoring instrument" is according to J.D. Brown (1991 quoted in Brown D.H. 2004 p. 21) a way to increase rater reliability.

#### *Test Administration Reliability*

Related to the conditions in which a test is administered. It means the locations, adequate use of sources and correct function of materials (Brown 2004)

#### *Test reliability*

Brown (2004) refers to this factor taking into account the time a test takes to be performed. He argues that tests should avoid becoming a race against time, because this can affect the performance of students. (Brown 2004. p. 22)

*Validity*

Brown (2004) considers validity as one of the most important aspects when assessing (p. 22) He argues that validity is measured in the level in which the test reflects the purpose for which it has been designed. He explains that a reading ability test must measure reading ability; no other things (Brown 2004 p. 22). Brown presents five aspects concerning validity.

*Content Validity*

The test must be designed to focus on the ability that is intended to be measured. Brown (2004) presents a clear example that defines content validity:

If you are trying to assess a person's ability to speak a second language in a conversational setting, asking the learner to answer paper-and-pencil multiple-choice questions requiring grammatical judgments does not achieve content validity. A test that requires the learner actually to speak within some sort of authentic context does. (Brown H.D. 2004. *Language Assessment. Principles and Classroom Practices*. White Palms: Pearson Education)

*Criterion-related validity*

This terms are used to define the "extend to which the criterion of the test has actually been reached" (Brown 2004 p. 24). Criterion-related validity is demonstrated - in the case of teacher-made classroom assessments- in a better way by comparing the results obtained "of an assessment with results of some other measure of the same criterion" (Brown



2004. p. 24). Brown (2004) also defines two different categories in which criterion-related validity can fall. The concurrent validity that means that the results are accompanied by other "concurrent performance beyond the assessment itself" (Brown 2004. p 24) and the predictive validity which is necessary in some sort of tests like placement, admissions, and so forth whose purpose is to predict how well someone is going to perform some activities in the future. (Brown 2004. p 24)

#### *Construct-related validity*

As Brown (2004 p. 25) explains, construct validity is an important feature sometimes neglected in classroom tests. The idea of the construct validity refers to the need of implementing some theoretical basis that serve as means for "explaining phenomena in our universe of perceptions" (2004 p. 25).

#### *Consequential Validity*

Expressing all the consequences that can derive from applying a test. McNamara (2004 p.53-54) considers the consequential validity as the effects of tests, and he argues that teachers should try to foresee that consequences, so that they can "anticipate them and investigate their effect on the validity of test score" (McNamara 2000 p.54)

*Face Validity*

Brown (2004 p. 26) offers an important view of what he calls face validity. These terms are used to name the impression caused by the test. He establishes that face validity refers to "the degree to which a test looks right, and appears to measure the knowledge or abilities it claims to measure" according to the subjective view of those involved in the test. (Brown 2004 p. 26) This means that in the case of students, they should feel that the test is testing what it was supposed it had to assess. (Brown 2004 p. 26) According to Brown (2004), in order to increase the face validity of a test, it is useful that learners find in it:

- a well constructed, expected format with familiar tasks.
- a tests that is clearly doable within the allotted time limit,
- items that are clear and uncomplicated,
- directions that are crystal clear,
- tasks that relate to their course work (content validity) and,
- a difficulty level that presents a reasonable challenge. (Brown H.D. 2004. Language Assessment. Principles and Classroom Practices. White Palms: Pearson Education)

*Authenticity*

Bachman and Palmer (1996 quoted in Brown 2004 p. 23) define authenticity as

..."the degree of correspondence of the characteristics of a given language test task to the features of a target

language task" (Bachman and Palmer 1996. quoted in Brown 2004. Language Assessment. Principles and Classroom Practices. White Palms: Pearson Education)

Brown (2004 p. 28) also presents the necessity of authenticity as a task associated to what he denote as "real world". Moreover, he presents some characteristics that can in which authenticity may be present in a test:

- The language in the test is as natural as possible.
- Items are contextualized rather than isolated.
- Topics are meaningful (relevant, interesting) for the learner.
- Some thematic organization to items is provided, such as through a story line or episode.
- Tasks represent, or closely approximate, real world tasks. (Brown H.D. 2004. Language Assessment. Principles and Classroom Practices. White Palms: Pearson Education)

#### *Washback*

Brown (2004 p. 29) considers washback as a feature used to know the consequences that a test have on "instruction in terms of how students prepare for a test". But he also, comments the possibilities of proportionate appropriate feedback to the students based on their own mistakes. It means the effects of a test in the learning process.

Now all these elements have been considered, it should be useful to revise the way in which a test is designed.

#### *Test design*

Harrison (1983 p. 16) presents a seven steps procedure that can lead to the construction of a test. However, more

items can be included as those presented by Brown (2004), Heaton (1990) and Weir (1993).

Harrison considers that a good test construction starts with the specification of some items. He also states that "a good specification ... is the result of careful judgments" instead of the "precise definitions". (Harrison 1983 p. 16)

Moreover the aspects considered above about the characteristics of a reliable test, an outline that searches for specific aspects is desirable.

The outlining for specification has been designed according to the author in seven steps.

#### *Objectives*

The objectives vary according to the type of test and to the necessity expressed by the teacher to check some aspects of language (Harrison 1990 p. 18).

#### *Skills*

Harrison states that "since skills do not operate in isolation from each other" it is necessary to establish what skill teachers want to assess and identify the type of test that can be applied in order to test certain skill (Harrison 1990 p. 18)

#### *Content*

The content of a test is according to Harrison (1990 p. 18) defined by the objectives placed at the beginning

of the specification.

*Format*

Referring to the "mechanics of the test" (Harrison 1990 p.20) and their implications.

*Rubric*

It is necessary to place a set of clear instructions on the test. This is a very important feature in order to maintain the validity of the test (Harrison 1990 p. 20)

*Materials*

Harrison (1990 p. 21) considers materials as a problem of administration but an aspect that must be taken into account. It is important to consider the facilities when using some materials, the time and the length it take to use them (Harrison 1990 p. 21).

*Marking*

At this point, Harrison's view about marking is reduced to the discrimination of what he calls objective and subjective tests (Harrison p.22). However, as this paper is intended to measure the performance of students in oral skills when they know what aspects of language are assessed, it is necessary to consider other concepts.

First of all, as it was stated by Heaton (1990 p. 67) testing speaking becomes difficult because the method used

to award marks. Oral tests tend to be qualified in a subjective way. Harrison (1983 p. 110-117) argues that subjectivity in tests is related to the way in which tests are marked and not to the type of test.

Secondly, considering the numerous problems with subjectivity in oral tests, Heaton (1990 p. 68) proposes an option to avoid this aspect in tests by offering a well structured marking or rating scale.

Even though Heaton (1990 p. 68) is inclined to assess oral production through a rating scale, Underhill (1987 p. 94-103) proposes a system that involves both a rating scale and marks. This author introduces a significant element in oral production assessment, what is called a grid or chart for rating and marking students' performance. Underhill (1987 p. 95) combines a rating scale that is according to Heaton (1990 p. 69) "a short description for a grade in a scale", with a marking system. Basically, the grid is going to be an instrument that incorporates categories or descriptions of language performance and a grade or mark that represents that performance.

This grid is going to be the tool through which teachers can justify their decisions and students can correct their mistakes. The importance of developing such a tool is according to Underhill (1987 p. 95) to establish

what he calls a "protocol" which is going to be useful to "save time and uncertainty". Moreover, as the author states, the marking system specifies in advance how to assess or how to categorized students according to their performance by given specific directions towards what it is going to be assessed.

As Weir (1993 p. 25-26) states, not only the teachers but also the students need to know what is going to be assessed and in which form those aspects are going to be rated or marked. The problem among teachers is concerned with the criteria selected when assessing oral production.

Weir (1993 p. 26) proposes that examiners have to standardize the way in which they mark a student's performance. In this way, teachers can create a consensus in terms of rating, marking and prioritizing aspects of learning when a test is carried out.

And in third place, it is important to consider the recent conceptions about assessment and the way in which tests are designed. Weir (1993 p. 30) proposes a three-step procedure to design an oral test. According to the author, it is necessary to decide whether or not an activity is useful for a test, the conditions in which a test is carried out and the output in terms of quality according to the criteria selected for assessing spoken interaction. These

features create a framework to analyze the possibilities of a test to be performed.

As Richards states (1994 p. 16) the criterion selected for a test, creates a criterion-referenced test that is according to the author:

a test which measures a student's performance according to a particular standard or criterion which has been agreed upon. The student must reach this level of performance to pass the test, and a student's score is therefore interpreted with reference to the criterion score, rather to the scores of other students (Richards, Platt, and Weber quoted in Richards J.C. 1994. *The Language Teaching Matrix*. New York: Cambridge University Press)

Now all these aspects have been presented and clarified, it is compulsory to define the aspect of language that is intended to assess and the setting - referring to the approach - in which this research is going to be performed.

#### *Communicative Language Teaching (CLT)*

Richards and Rodgers (1986 p. 64) present a review of CLT that must be considered in this document. However, this review of the literature is going to be focused on the theory surrounding CLT approach and the definition of communicative competence.

As it is stated by Richards and Rodgers (1986 p. 69) CLT tends to develop what Hymes called communicative competence. Hymes' view is closer to be a reaction against Chomsky's view of competence by including the term



communication (Richards and Rodgers 1986 p. 69). Chomsky's view of competence was related to the "abstract abilities speakers possess that enable them to produce grammatically correct sentences" (Richards and Rodgers 1986 p. 70).

However, Hymes stated that such a view of competence needed to incorporate two major aspects as communication and culture (Richards and Rodgers 1986 p. 70). Richards and Rodgers (1986 p. 71) state that CLT proponents had a similar point of view of learning a second language which is "acquiring the linguistic means to perform different kinds of functions". The theory proposed by Hallyday was benefited by CLT. Hallyday's view of language is that:

Linguistics... is concerned... with the description of speech acts or texts, since only through the study of language in use are all the functions of language, and therefore all components of meaning, brought into focus. (Hallyday 1970 p. 145 quoted in Richards and Rodgers 1986. *Approaches and Methods in Language Teaching*. New York: Cambridge University Press)

Widdowson (1978) presents a relationship between the "linguistic systems and their communicative values in text and discourse". He underlines the importance that speakers' ability for using the language has in order to accomplish different purposes (Widdowson 1978, quoted in Richards and Rodgers 1986 p. 71)

CLT comports, according to Richards and Rodgers (1986 p. 71), three basic principles that have been discerned

among the practices carried out in this approach. The first principle is concerned with communication: any activity involving "real communication promotes learning" (Richards and Rodgers 1986 p. 72). A second principle is related to the tasks: According to Johnson (1982 quoted in Richards and Rodgers 1986 p. 72) those activities that aim for using language to complete meaningful tasks encourages learning. And the last one is called the "meaningfulness principle": It is basically that "language that is meaningful to the learners supports the learning process" (Richards and Rodgers 1986 p. 72)

In CLT, the communicative view of language implies that:

1. Language is a system for the expression of meaning.
2. The primary function of language is for interaction and communication
3. The structure of language reflects its functional and communicative uses.
4. The primary units of language are not merely its grammatical and structural features, but categories of functional and communicative meaning as exemplified in discourse. (Richards, J. & Rodgers, T. 1986. Approaches and Methods in Second Language Acquisition. New York: Cambridge University Press. p. 71)

#### *Communicative Competence*

According to Richards and Rodgers (1986 p. 70) Hymes uses the term communicative competence to designate "what a speaker needs to know to be communicatively competent in a speech community" and in his view

"a person who acquires communicative competence acquires both knowledge and ability for language use respect to:

1. whether (and to what degree) something is formally possible;
2. whether (and to what degree) something is feasible in virtue of the means of implementation available;
3. whether (and to what degree) something is appropriate (adequate, happy, successful) in relation to a context in which it is used and evaluated;
4. whether (and to what degree) something is in fact done, actually performed, and what its doing entails.

(Richards, J. & Rodgers, T. 1986. Approaches and Methods in Second Language Acquisition. New York: Cambridge University Press. p. 70)

Canale and Swain (1980. Quoted in Richards and Rodgers 1986 p. 70) identify four dimensions of communicative competence.

They consider that grammatical competence, sociolinguistic competence, discourse competence and strategic competence are aspects that are circled in a bigger term that is communicative competence.

The *grammatical competence* refers to the "grammatical and lexical capacity" a speaker has. It is connected to the idea of Chomsky about competence. (Canale and Swain 1980. Quoted in Richards and Rodgers 1986 p. 70)

*Sociolinguistic competence* is a term used for ... an understanding of the social context in which communication takes place, including role relationships, the share information of the participants, and the communicative purpose for their interaction" (Richards, J. & Rodgers, T. 1986. Approaches and Methods in Second Language Acquisition. New York: Cambridge University Press. p. 70)

The *discourse competence* is concerned with the connection of messages and the meaning of them in a whole context.

And the *strategic competence* refers to

...“the coping strategies that communicators employ to initiate, terminate, maintain, repair, and redirect communication” (Richards, J. & Rodgers, T. 1986. *Approaches and Methods in Second Language Acquisition*. New York: Cambridge University Press. p. 70)

Seville-Troike (1996. *The ethnography of communication*. In Hornberger H. & McKay S. Ed. *Sociolinguistics and Language Teaching*. New York. NY: Cambridge University Press) considers some important aspects about communicative competence.

According to Seville-Troike, communicative competence involves to know “not only the language code”, it also includes to know “what to say, to whom, and how to say it appropriately in any given situation” (1996. *The ethnography of communication*). She argues that communicative competence is extended to:

...“both knowledge and expectation of who may or may not speak in certain settings, when to speak and when to remain silent, whom one may speak to, how one may talk to persons of different statuses and roles, what non-verbal behaviors are appropriate in various routines, ...” (Seville-Troike (1996) *The ethnography of communication*. In Hornberger H. & McKay S. Ed. *Sociolinguistics and Language Teaching*. New York. NY: Cambridge University Press)

In other words, to be communicatively competent means to be able to put the four dimensions -mentioned earlier- together. However, because the necessity to interact in a "speech community", it is important to distinguish some factors that can affect the decision of sequencing and selecting the language of a curriculum. (Seville-Troiike 1996. The ethnography of communication).

..."the content of what a speaker needs to know depends on the social context in which he or she is or will be using the language and the purposes he or she will have for doing so. (Seville-Troiike (1996) The ethnography of communication. In Hornberger H. & McKay S. Ed. *Sociolinguistics and Language Teaching*. New York. NY: Cambridge University Press)

Seville-Troiike considers that in order to select and sequence the language in a better way, it is important to consider the following aspects that are only going to be mentioned in this paper: Linguistic Knowledge, Interaction Skills and Cultural Knowledge. (Seville-Troiike 1996. The ethnography of communication).

#### *Types of tests in CLT*

In order to understand the development of communicative tests, it is important to consider some aspects.

According to McNamara (2000 p. 16) lately communicative language tests tend to have two important features. He comments that:

They were performance tests, requiring assessment to be carried out when the learner or candidate was engaged

in an extended act of communication, either receptive or productive, or both. They paid attention to the social roles candidates were likely to assume in real world settings, and offered a means of specifying the demands of such roles in detail. (McNamara T. 2000. *Language Testing*. Hong Kong: Oxford University Press)

Such a view of language as the one proposed by Hymes, produced that language was associated to a sociological perspective in which it focused "on the external, social functions" (McNamara 2000 p. 17)

Whatever aspects of language are intended to be assessed in a communicative setting, Weir (1993 p. 28), Brown (2004 p. 10), and McNamara (2000 p. 20) converge in the necessity of including real-world tasks or real-world situations as the framework in which a communicative test should be developed.

Moreover, Weir (1993 p. 28) considers that it is necessary in order to increase the validity of the instrument that communicative tests have clear specifications about:

... the operations students will have to perform in a target situation, the conditions under which those tasks will be performed and the quality of output that will be necessary. (Weir C. 1993. *Understanding and developing Language Tests*. Prentice Hall International)

As he considers that a major focus must be placed on the skill that is going to be assessed, it is relevant to consider some techniques in oral assessment.

*Oral Assessment*

According to McNamara (2000 p. 13) in order to assess communicative competence, two major types of tests were developed: the discrete point tests and the integrative and pragmatic tests.

The first type of tests was devoted to assess "candidates' knowledge of the grammatical system, of vocabulary, and of aspects of pronunciation" (2000 p. 14). This kind of test assessed "aspects of language in isolation" (2000 p. 14).

The integrative tests, on the other hand, integrate various aspects of language -such as grammar, vocabulary and pronunciation- "with an understanding of context" (2000 p. 15).

Once a teacher has selected what type of assessment is more convenient, it is important to consider what type of speaking is going to be measured.

Brown (2004 p. 141) considers that exist five "different performance assessment tasks" in speaking. The *imitative* task which corresponds to simply imitate the sounds but without conveying any meaning; the *intensive* task in which students are asked to perform some patterns but there is not a high level of communication; the *responsive* which corresponds to very limited conversations

characterized by "small talk, simple requests and comments" (Brown 2004 p. 141); the *interactive* task that involves a more complex interaction while including more people and becoming "pragmatically complex" (Brown 2004 p. 142); and the *extensive* task or *monologue*.

Brown also considers the existence of micro-skills and macro-skills in speaking which can be adopted as the criteria for an oral assessment (Brown 2004 p. 142). Micro-skills refer to "producing the smaller chunks of language" while the macro-skills "imply the speaker to focus on larger elements: fluency, discourse," and so forth.

In order to obtain a higher level of response in students, Underhill (1987 p. 47) considers some elicitation techniques for oral assessment. He argues that it exist at least 60 different techniques and the choice of one of these techniques corresponds to the needs of each teacher or course (Underhill 1987 p. 47) Some of those techniques are: *discussion/conversation, oral report, learner-learner joint discussion/decision making, role-play, interview, Learner-learner description and re-creation, Form-filling, Making appropriate responses, Question and answer, Reading blank dialogue, Using a picture or picture story..* (Underhill N. 1987. *Testing Spoken Language*. Cambridge: Cambridge University Press.)



*Elements involved in oral assessment*

According to Brown (2004 p. 140) there are some aspects to be considered when an oral test is carried out. Oral assessment practices involved an aural interaction that can compromise the validity and reliability of the test since speaking is "a productive skill that can be directly and empirically observed". (Brown 2004 p. 140). Brown also argues that oral tests should be designed in order to elicit the aspects of language a teacher wants to assess (Brown 2004 p. 140). Since communicative tests intend to develop real-world tasks it becomes difficult to focus on only certain aspects of language or special patterns. As Heaton (1990 p. 67), Brown (2004 p. 140) also considers that awarding grades or scores becomes difficult because of the nature of the exam.

*Fluency*

According to Hedge (2000) this term is closely related to speech. The author considers that fluency

... is the ability to link units of speech together with facility and without strain or inappropriate slowness, or undue hesitation. (Hedge, T. 2000. *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.)

*Accuracy*

Accuracy is used to refer "to the ability to produce grammatically correct sentences that are comprehended" (ESL

Glossary: Definitions of common ESL/EFL terms from  
<http://bogglesworld.com/glossary/accuracy.htm>)

*Formats used to assess oral production*

Some approaches that are connected with the idea of developing communicative competence are according to Richards and Rodgers (1986) The Natural Approach, Cooperative Language Learning, Content-Based Instruction and Task-Based Language Teaching. It is important to consider how teachers work in terms of oral assessment within these approaches and discover if some criterion is explicit through a grid to assess oral abilities.

Some grids that can be found during this research are placed on the Appendix section. Some of them have been devoted to assess students' performance in an oral practice which is not exactly an exam. However their content is useful for the purposes of this research.

One important grid found during this research was the grid proposed by the European Council, in which many categories are described according to the expected level of the student.

III. CHAPTER 3: METHOD

*Design of the study*

The present research was undertaken based on the ideas set forward by Brown (2004), who states the main steps in designing tests, implementing marking and scores and who provides in an adequate way the features of effective tests. This author's ideas are relevant since he considers all the elements involved in foreign language assessment and also the characteristics of oral assessment, two main elements in this paper.

This research was intended to compare the performance of the students when a grid is used for assessing oral production.

This was a descriptive study since it did not involve the manipulation of treatments (Schafer, 2001). Moreover the design of the study can be placed as a causal comparative study because this kind of study:

...compares the likeness and difference among phenomena to determine if certain factors or circumstances tend to accompany certain events, conditions, or processes. (Key James P. 1997.)

*Procedure*

This research was carried out in the Language Center of the Linguistics and Languages Department at the University of Nariño. This Language Center was created in 1961 and its goal is to offer foreign language learning programs to

people in Pasto. This center is found in the campus located in Panamericana Avenue. Since this academic unit fosters the search of pedagogical advancements through research and methodological experimentation, it is the best place to apply an innovative grid and to verify its results.

Students had to perform two oral tests that had been designed according to the parameters established in this paper and to the contents studied during this period of time.

Both tests contained 40 different questions. Both were recorded and graded in a different way. The first one was graded without using any kind of grid and no previous criteria were arranged with students. The second one consisted of the same number of questions but it was graded according to the grid designed in this paper. The criteria for this test were specified in advance with the students.

Students had the opportunity to prepare the questions a week before the test took place and the test was administered individually by the teacher.

Once students took the tests, a survey was used in order to know their opinions and perceptions about the use of the grid. Students' scores are included in this paper in order to compare their performances.

### *Subjects*

The students who were part of this research were 20 students of first level of English at the Language Center in the University of Nariño. They are teenagers and their ages ranged from 14 to 17 years old. There were 15 women and 5 men in this group. They attended classes from Monday to Thursday from 4:00 p.m. to 6:00 p.m. during a period of sixteen weeks.

### *Materials*

Some of the materials used in order to accomplish the objectives of this study were: a tape recorder, a computer, bibliography and formats used in communicative approaches for assessing oral production, internet resources, cassettes, the grid designed by the teacher, eighty different questions whose purpose is to involved students into an informal conversation and the review of some topics studied in the classroom, copies of the exams and of the grid for each student, .

### *Instruments*

The instruments were the two tests planned for an oral assessment practice and the grid developed during this research.

### *Variables*

The two variables in the study are:

*Independent variable:* The grid

*Dependent Variable:* The performance of students in oral assessment.

*Data Gathering and Analysis*

For the purposes of this research data were collected through the use of two oral exams. This kind of practice is recognized as an interview. According to Key (1997) an interview represents a face to face interaction, whose purpose is to obtain verbal responses from some "respondents".

The type of interview expected in this research is an unstructured one, since the idea is to make students express freely based on a series of questions that should be considered as a plan of some topics to be prepared. Some of the benefits proposed by Key (1997) when using this kind of data gathering tool are

- They have few restrictions.
- If preplanned questions are asked, they are altered to suit the situation and subjects.
- Subjects are encouraged to express their thoughts freely.
- Only a few questions are asked to direct their answers.
- In some instances, the information is obtained in such a casual manner that the respondents are not aware they are being interviewed. (Key, P.J. 1997. Other Data Gathering Tools for a Research Investigation. *Research Design in Occupational Education*. Oklahoma State University. Retrieved October 25<sup>th</sup>, 2005 from <http://www.okstate.edu/ag/agedcm4h/academic/aged5980a/5980/newpage17.htm> )

Two exams were used; both oral tests were recorded in order to have reliable grades that can be certified at any moment by any person who can be interested in the results. Once both tests were applied, students were asked to answer a questionnaire. This questionnaire was used since Key (1997) considers that questionnaires are useful tools to know about deep "feelings, beliefs, experiences, perceptions, or attitudes" of people about a given situation.

Questionnaires can be structured or unstructured. For the purposes of this research, a structured questionnaire was used. According to Key (1997) these kinds of questionnaires are:

... a very concise, preplanned set of questions designed to yield specific information to meet a particular need for research information about a pertinent topic. (Key, P.J. 1997. Other Data Gathering Tools for a Research Investigation. *Research Design in Occupational Education*. Oklahoma State University. Retrieved October 25<sup>th</sup>, 2005 from <http://www.okstate.edu/ag/agedcm4h/academic/aged5980a/5980/newpage17.htm>)

The questionnaire comprehended eight questions with Yes/No answers. The questionnaire was written in English, but if students needed some extra explanations about the content of the questionnaire, the information was administered in Spanish.

Once the results of both tests were prepared, the analysis was done as follows. As this was a casual

comparative study, because its objective was to determine the influence of a variable or factor - in this case the grid - in other variables, the results of both tests were compared. Tests were graded in a range from Zero to Five. Both tests were analyzed individually, and then the comparison was carried out. What was compared was the performance of the whole group and not the performance of each student when using or not the grid. Moreover, when using the grid, the results provided the teacher with reliable information about the level of performance of his students. An analysis was carried out for each major category in the oral exam (fluency and accuracy) and each intermediate description of the performance in these categories.

Once this comparison was done, the results obtained from the questionnaire were described and analyzed. All this process was done using descriptive statistics analysis. The analysis proposed in the Results Section was based on the ideas suggested by Brown. (Brown J.D. 1983. *Understanding Research in Second Language Learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press)



#### IV. CHAPTER 4: RESULTS

##### *Results*

During the research some major aspects, which were intended to be measured at the beginning of this paper, were found. Since the aims of this paper focused on comparing students' performance in two oral tests whose marking system differed in the use of certain criteria to award marks, it is convenient to describe the facts found out in both tests.

Before considering the results it must be kept in mind the following aspects:

- Both tests consisted of 40 different questions whose purpose was to offer different topics to talk about. However, those questions were based on major aspects of language that were studied during the course.
- Both tests were handed in to the students a week before the exam was placed.
- Both tests were recorded, and in this way, they became reliable proves of the existence of the tests and the performance of students.
- Students' performance was graded once the exams finished.
- Exams were carried out using an interview.
- The criteria for the exam were explained a week before the test was carried out.

- Both tests were designed taking into account the parameters registered in this paper about the design and the marking of oral tests.

The results that were found in the first oral exam demonstrated that students' performance was not at the expected level they should be.

As a matter of fact, any student failed the exam; however their results and their performance appeared to be at an unexpected level. Only the 38% of the students awarded a score above four, but the best grade obtained by a student was four point three (4.3). Three students awarded this grade. The worst grade was three point zero (3.0). Only one student passed the exam with this mark.

In order to analyze the results obtained by students, some major categories were established as the framework to analyze the results. Any student obtained a grade above zero (0.0) and less than one point zero (1.0). The same occurred with the next category: students' grades were not in a range of more than one point one (1.1) and less than two (2.0). Eight percent of the students' population that took the exam were in the next category presented as the range between grades higher than two point one (2.1) and three (3.0).

Surprisingly, the highest percentages of students' grades were in the next range. Those students whose

performance awarded a mark above three point one (3.1) and four (4.0) represented the 54% of the whole population. As it was said before, only the 38% of the students ranged above four point one (4.1) and five (5.0) which is the best grade a student could obtain.

The combined average of the students was three point seventy-three (3.73) which is apparently adequate to their level.

The second oral exam presented the following results. Only one student failed the exam. Sixty percent of the students obtained a grade ranging from three point one (3.1) and four (4.0). This is expected since these tests were designed in order to conserve the characteristics mentioned before in this paper and proposed by Brown about the reliability of a test. Thirty percent of the students obtained a grade above four point one (4.1). Remarkably, one student obtained a grade of four point nine (4.9) represented the highest grade among students in both the first and the second test.

The average of students in the second examination was three point sixty-two (3.62) which is closer to that obtained in the first exam.

In terms of the descriptions used in the grid, number three was mostly used for both fluency and accuracy aspects.

Sixty percent of the students were ranged in these categories. Ten percent of students were in a lower category - category number two - in both aspects; twenty percent of students ranged in the fourth category.

The hypothesis established at the beginning of this research was proved. The use of a grid in an oral exam had neither positive nor negative effects in the performance of students during such a test. However, it is important to take into account some aspects that are going to be considered broadly in the discussion section of this paper.

The questionnaire presented the following results.

For question number one that interrogated about the usefulness of knowing the criteria of an oral exam before the exam is placed, all students answered that it is useful to know these criteria.

In question number two, whose purpose was to ask students if they considered that when they know the criteria of an oral exam they would improve their grades, all of them answered positively.

The third question offered the following results: 63.6% of the students considered that the criteria given to them helped them to obtain better grades. On the other hand, 36.4% of the students considered that the criteria did not help them to obtain better scores.

In question number four, 90.9% of the students agreed that the criteria served them to focus on specific aspects of language, while only nine point one percent argued that the criteria used did not make them focus on certain aspects of language.

Question number five, as well as question number six, was answered positively by all students. The fifth question asked students about the possibility of correcting their mistakes in an effectively way when using the criteria. In the sixth question, students' perception about the "appearance" of the test in terms of organization and planning was demanded.

Question number seven was answered positively by 90.9% of the students. Only nine point one of them considered that they preferred not to know their grades immediately the test is finished.

Question number eight produced the following results: 18.2% of the students considered that no matter the use of the criteria, oral tests were still considered threatening; while the 81.8% of them observed oral tests as less threatening and more acceptable tests when using the criteria.

The combined average of students concerning their assumptions about the use of the grid resulted in this:

90.9% of the students had a positive view of the use of the grid during the oral test; nine point one per cent of students had a negative view of the use of a grid during an oral test.

V. CHAPTER 5: DISCUSSION AND CONCLUSIONS

*Discussion*

The hypothesis proposed at the beginning of the study was proved. Students did not improve or decrease their performance in an oral test no matter the use of a grid.

This discovery can lead us to interpret the results focusing on the process carried out during this research. It is meaningfully important for teachers to adopt certain kind of criteria when they have to provide their students with an oral assessment practice. As it was established by the students in the questionnaire, the use of a grid is a good way to increase the validity of a test since it tells them what the test will deal with and they can realize that the questions matched the suggested criteria. Moreover, the use of a grid before, during and after an oral assessment practice is, according to the students' answers, a powerful tool to improve their performance in a test and in this way, obtain better grades.

If both exams are considered, students obtained similar grades. This means that the tests used to assess students communicative competence - in the case that this competence can be considered as a basic one - possessed a high level of reliability. Brown (2004 p. 21-22) considers this as an important factor to be taken into account in language

teaching: the way a test is built should not affect students negatively or positively. As it was shown in this paper, having clear criteria in mind considering the content of the course and the level of proficiency of students contributes to creating reliable tests. Meanwhile some factors as the validity, the authenticity, the practicality as well as the washback can be considered in a different stage or measured in a different way; the reliability of a test is measured by the performance of the students in two different moments of their learning process (Brown 2004 p. 21-22). Both tests presented in this research had a high level of reliability; they were authentic and practical. Moreover, both measured what they intended to determine (validity).

Both tests were designed and applied following the directions of major authors whose works are remarkably important in the English learning and teaching field. The significance of using such important works, made this research a consistent one that can offer different options for further research and that can also serve as an overview for those interested in evaluation, assessment and testing.

Some relevant aspects of the use of the grid are closely related to what Underhill called a "protocol" (1987 p. 95). Basically, the grid serves to students and teachers to carry out an assessment practice based on some aspects of



language; this increases the validity of tests since the criteria selected - in this case - tends to make students focus on oral aspects of the language.

During this research, the review of some current literature made the researcher find that a variety of important institutions around the world had created their own grids to determine the level of their students. The European Council for instance, has a well-structured recognized grid whose purpose is to determine a common framework for language examinations. On the other hand, some teachers like Sarah Rapnouil-Dunn have created their own grids, following major directions in the teaching and learning of foreign languages, but arranging them to the students' context. It is compulsory for the University Of Nariño, and for their languages teachers to implement specific criteria in every assessment practice and to make criteria available, known and applied by all teachers in order to improve the teaching process. Moreover, it is important for the university to create some kind of grid that can help to assess students' performance during their career, and in this way improve the teaching of a foreign language. This means that teachers can focus on main aspects of language they consider a student of the university at the Linguistics and Languages Department has to perform at the

end of his/her learning process at the university. These criteria permit teachers to work in order to develop - throughout a common methodology, maybe a communicative one - and assess specific aspects of language in different periods of time.

Another important feature during this research was the quality of utterances produced by students during the second exam. They intended to do what the teacher said would be assessed. Students improved their performance in terms of the length of interventions and free speaking, which is considerably important in the learning of a foreign language as it is expressed in grids such as that of the Common European Framework. Taking into account their level, students now know how to face oral exams, to produce language which is appropriate in quality and in quantity, and in which way they can improve their grades.

An important element when using the grid is the possibility for teachers to give feedback to students immediately after the exam is finished. This possibility appeared when the teacher marked students' performances and gave them the reasons why they obtained certain grades based on the grid. As Heaton (1990 p.67) proposed, marks were not done during the exam but at the end. This produced that students focus on the exam as well as the teacher and

prevented students from getting nervous. Sometimes instead of paying attention to the students, teachers spend the whole exam writing notes about the performance of each student which may be confusing. A grid which is marked easily is a useful tool to avoid this situation.

A grid is also a good element for diagnosing the students' needs and to make a plan to help learners overcome their difficulties. This was also made evident by the use of the grid created for this research. Since the final results indicated that sixty percent of the students ranged in the third category proposed in the grid, it becomes a teacher's work to move students from this category to the next one. It means that in terms of accuracy, students need to improve their grammar so they can create more accurate utterances. Concerning fluency, although it was said before that the length of free speaking among students was increased comparing the first exam, students need to produce longer utterances and maintain the precision of the message they intended to convey.

No major limitations were found during this research. However, students experience in taking oral exams can be considered as a problem since they did not know how to behave during an interview, or what was expected from them for instance, during the first exam. Another problem is the

absence of few students during both tests who did not take the tests because they did not feel confident. Nevertheless, their absence was not a problem for obtaining a good idea of the group attitude and performance.

Further research would be desirable in order for the university to create their own system for awarding grades to students, select candidates and evaluate programs and the curricula at the languages and linguistics department. It is hoped that this first step will serve as the basis for teachers and students to go deeper in the study of foreign language assessment.

#### *Conclusions*

Some important conclusions have already been presented in the discussion section. However, there are some key aspects that need to be kept in mind in the long run. Here are some of them:

The use of a grid does not mean that it will automatically and almost "magically" improve students' grades during an oral exam. However, a grid does contribute to improve the attitude students have towards oral exams and to reduce anxiety and negative feelings, which is very relevant if we consider that anxiety may have a huge influence in the performance of a student.

Grids are useful tools that serve teachers and students to focus on specific aspects to be measured during an oral exam and it helps learners study or concentrate on certain key elements of language that otherwise might have been overlooked by them when studying for a test or when practicing in class.

Teachers and students need to be aware of some kind of criteria to assess oral and any other skills, in order to have a less subjective assessment activity. As Brown (2004) explains the rater of a test, in this case the teacher, may be affected by different factors that lead to grading in a subjective way. A grid can reduce this risk by giving teacher some focal elements that may contribute to being fair to all students and to make students feel that they are not being assessed as persons, rather it is their competence to do something what is being assessed.

A successful assessment activity requires the design of appropriate instruments that conserve the characteristics proposed by major authors as Brown, Heaton, Weir, Underhill and so on. These characteristics are important in that they help teachers create tests which are close to real life, which are not time consuming in the grading part, which make students feel like they can actually pass a test no matter the topic and mostly, which truly show that assessment and

tests play a crucial part in improving the learning process; they should not just be seen as a way to *punish* students.

Students have a good opinion about the use of a grid during an oral test not because they can memorize exactly what to say to please the teacher, but because they realize that their teacher knows more than a few sentences in English. A grid shows learners that the actions teachers take in the classroom have a theoretical foundation and that the decisions we make have the aim of helping them learn English, which is the ultimate goal.

The tests used during this research were appropriate and consistent. This is remarkably important since students' performances were not manipulated in order to prove the hypothesis presented in this research.

Students at the Language Center require to be assessed orally more frequently. This is a factor that influences students' performance in L2, and ultimately, the development of their communicative competence. Despite the fact that creating and administering oral tests is time consuming, the washback they provide is very valuable for teachers, since this is the basis for reinforcing those aspects students have trouble with. However, oral assessment is not done with a recipe and teachers need to implement some specific criteria according to the requirements of each level and the

requirements and goals proposed by the Language Center and by the Department.

The grid is a tool that must be considered as an instrument whose purpose is to organize any kind of tests and as a potential source to provide students with appropriate feedback.

The University of Nariño requires the implementation of some "protocol" that serves as a tool to select, assess and decide whether or not a candidate or a student has the required level of English to be part of the university or to pass to a different semester. Through the use of some criteria, the English level at the university would be increased to become one of the highest in our country.

Moreover, it is noticeably important that teachers share the way in which students' performances are measured at the university, through the use of a common criteria by all of them. This can reduce subjectivity in tests and offer students as well as teachers a more consistent method for obtaining and awarding marks.

REFERENCES

- A Resource for the development and/or description of Assessment Tools, Retrieved October 25<sup>th</sup>, 2005 from [http://www.ecml.at/projects/voll/our\\_resources/graz\\_2001/selection/assessment/menu\\_assessment.htm](http://www.ecml.at/projects/voll/our_resources/graz_2001/selection/assessment/menu_assessment.htm)
- Brown, H.D. 2004. *Language Assessment. Principles and Classroom Practices*. White Palms: Pearson Education
- Brown J.D. 1983. *Understanding Research in Second Language Learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press
- Downloads (pdf) for teachers. *I. Aids for self-assessment of language skills*. Retrieved October 8<sup>th</sup>, 2005 from [http://sprachenportafolio.ch/esp\\_e/esp15plus/u\\_downloads.htm#eins#eins](http://sprachenportafolio.ch/esp_e/esp15plus/u_downloads.htm#eins#eins)
- Downloads (pdf) for teachers. *II: Forms to certify experiences*. Retrieved October 8<sup>th</sup>, 2005 from [http://sprachenportafolio.ch/esp\\_e/esp15plus/u\\_downloads.htm#zwei#zwei](http://sprachenportafolio.ch/esp_e/esp15plus/u_downloads.htm#zwei#zwei)
- Downloads (pdf) for teachers. *III. Aids to describe certificates and diplomas and to assign them to levels*. Retrieved October 8<sup>th</sup>, 2005 from [http://sprachenportafolio.ch/esp\\_e/esp15plus/u\\_downloads.htm#drei#drei](http://sprachenportafolio.ch/esp_e/esp15plus/u_downloads.htm#drei#drei)
- Harrison, A. 1983 (First published). *A Language Testing*



- Handbook*. Modern English publications: Hong Kong.
- Heaton J.B. 1990. *Classroom Testing*. Longman: New York.
- Hedge, T. 2000. *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.
- Key, P.J. 1997. Other Data Gathering Tools for a Research Investigation. *Research Design in Occupational Education*. Oklahoma State University. Retrieved October 25<sup>th</sup>, 2005 from <http://www.okstate.edu/ag/agedcm4h/academic/aged5980a/5980/newpage17.htm>
- Lazaraton A. (2001). Teaching Oral Skills. In M. Celce-Murcia (Ed) *Teaching English as a Second or Foreign Language*. (pp. 111-113) Boston: Heinle & Heinle.
- McKay, S.L & Hornberger N.H (Eds) (1998) *Sociolinguistics and Language Teaching*. Cambridge. Cambridge University Press.
- McNamara, T. 2000. *Language Testing*. Hong Kong: Oxford University Press.
- Nunan, D. (1992). *Action research in language education*. 13p. Retrieved 20 August 2002 from <http://www.les.aston.ac.uk/lsu/research/tdtr92/tdtrdn.html>
- Nunan, D. 1995. *Research Methods in Language Teaching*. New York: Cambridge University Press.

Richards, J.C. 1994. *The Language Teaching Matrix*. New York: Cambridge University Press.

Richards, J.C. & Rodgers, T. 1986. *Approaches and Methods in Second Language Acquisition*. New York: Cambridge University Press.

Schafer, William D. (2001). *Replication: a design principle for field research*. *Practical Assessment, Research & Evaluation*, 7(15). Retrieved October 8, 2005 from <http://PAREonline.net/getvn.asp?v=7&n=15>

Seville-Troike (1996). *The ethnography of communication*. In Hornberger H. & McKay S. (Ed) *Sociolinguistics and Language Teaching*. New York. NY: Cambridge University Press

The Basics of Experimental Design [A Quick and Non-Technical Guide] Retrieved September 9<sup>th</sup> 2005, from <http://www.sytsma.com/phad530/expdesig.html>

The E-Teach Website (2005). *Oral assessment grids samples*. Retrieved October 25<sup>th</sup>, 2005 from <http://teachers.domainepublic.net/shared/Oral/Assessment%20Grids/>

Underhill, N. 1987 (First published). *Testing spoken Language: A handbook of oral testing techniques*. Cambridge University Press.

University of Minnesota. (2003) *American Psychological Association format (APA) (5th Edition)*. Retrieved October 8<sup>th</sup>, 2005 from [http://: www.crk.umn.edu/](http://www.crk.umn.edu/)

Weir, C.J. 1993. *Understanding and Developing Language Tests*. Prentice Hall International.

Wikipedia the free Encyclopedia. *Fluency*. Retrieved October 25<sup>th</sup>, 2005 from <http://en.wikipedia.org/wiki/Fluency>

Figure 1. Students' Performance First Oral Exam

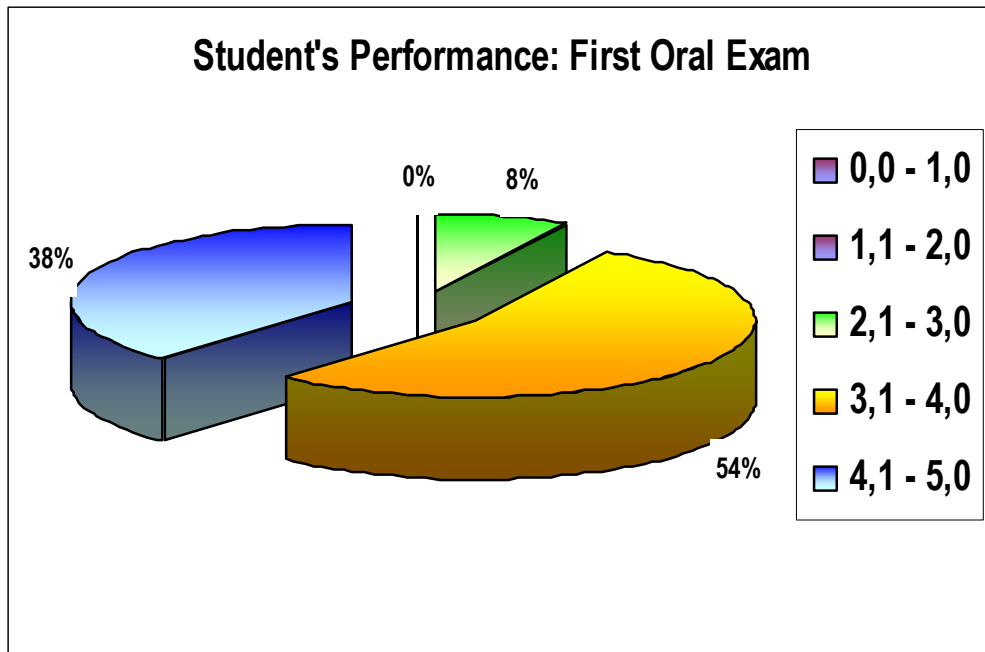


Figure 2. Students' Performance Second Oral Exam

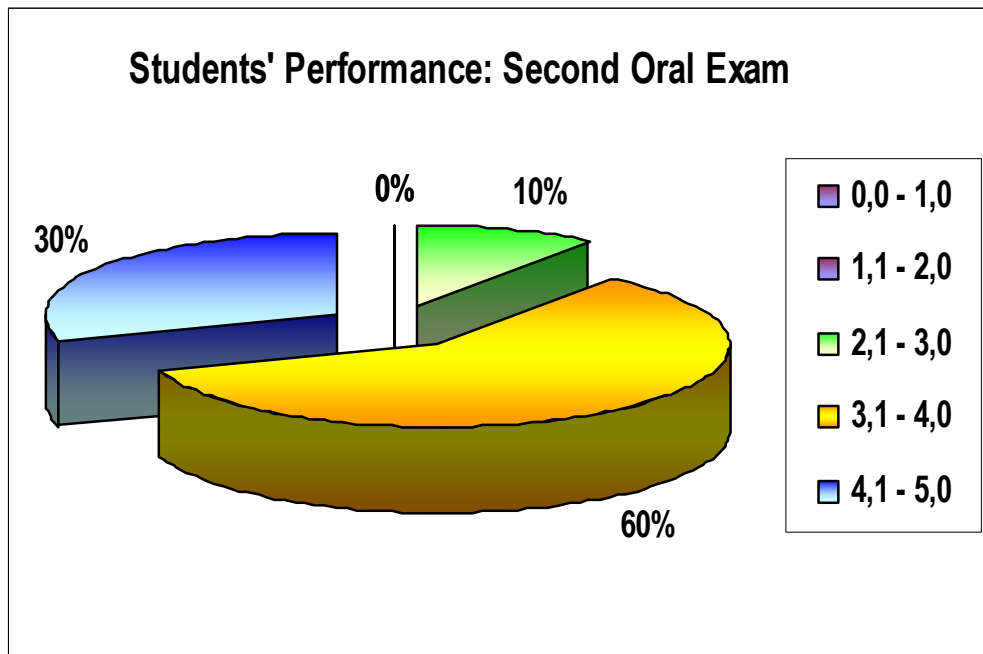


Figure 3. Students' Performances' Comparison

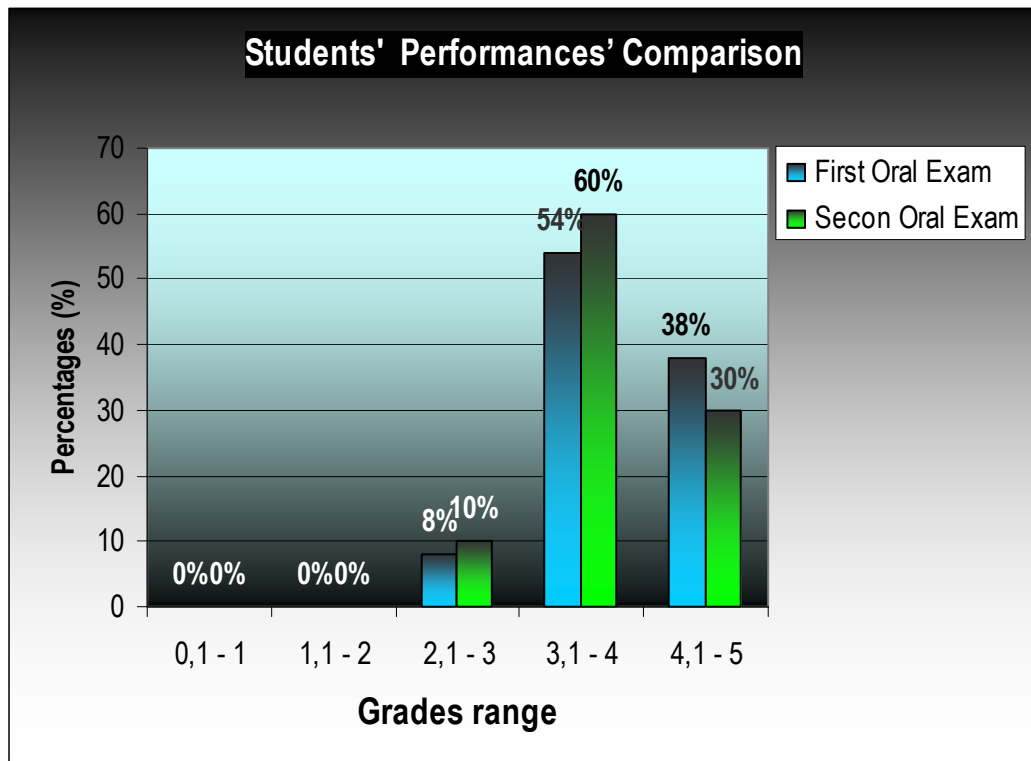
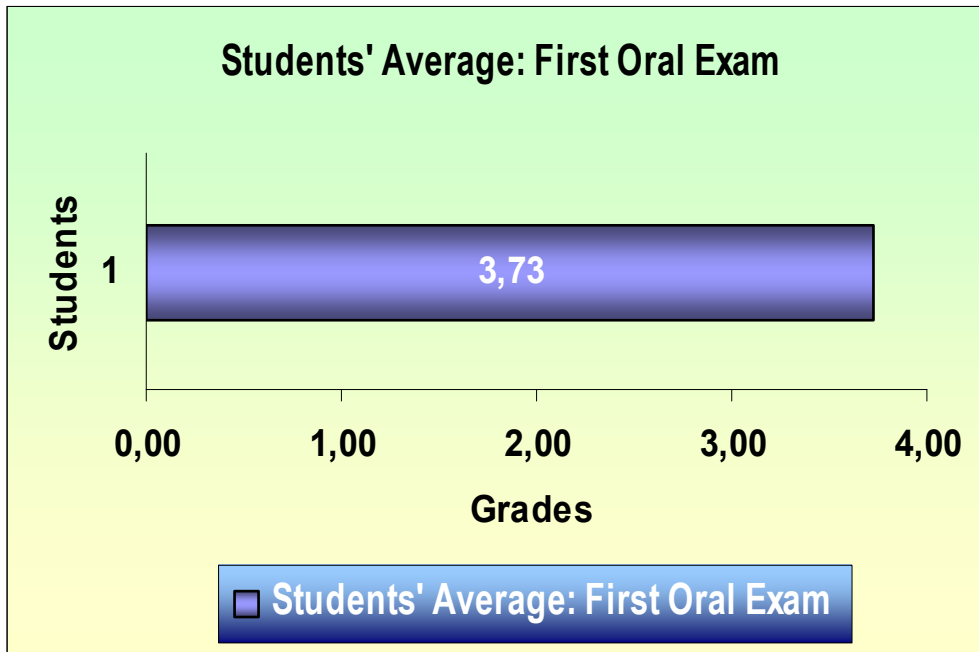
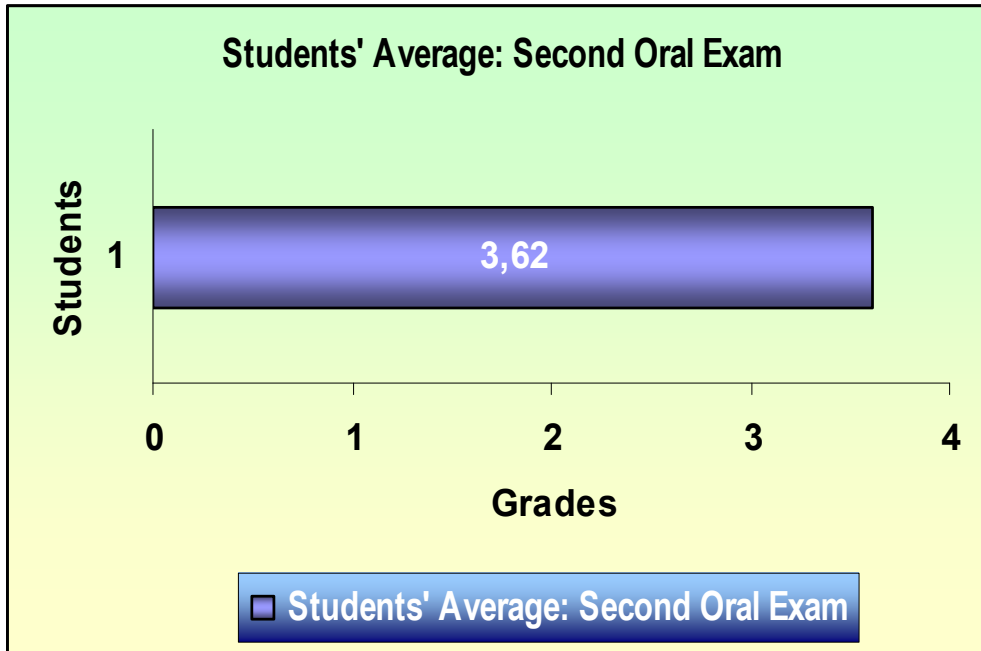


Figure 4. Students' Average First Oral Exam

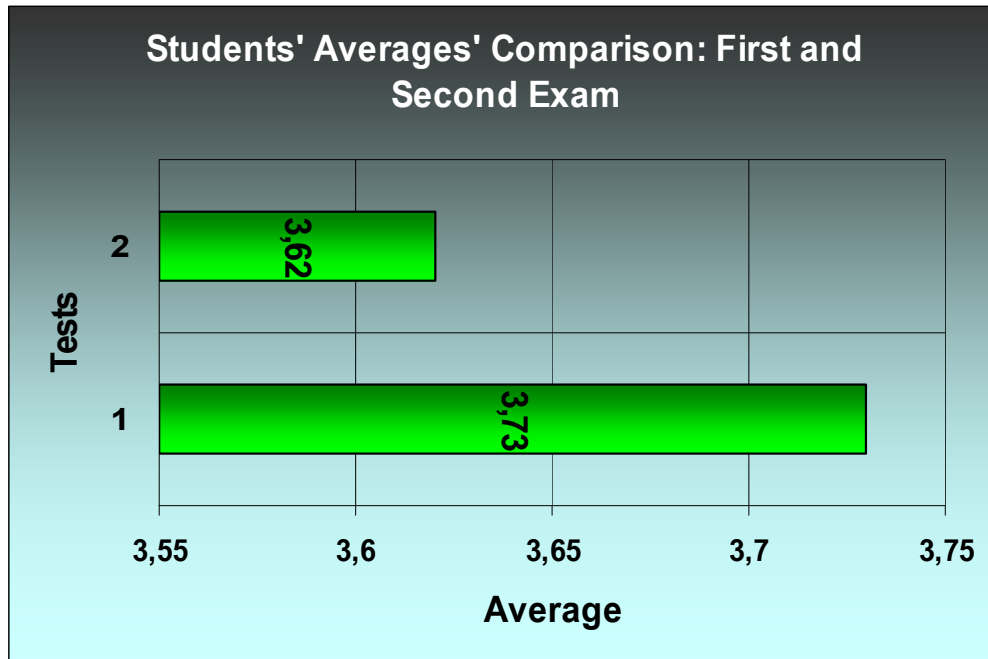


Students' Average Second Oral Exam

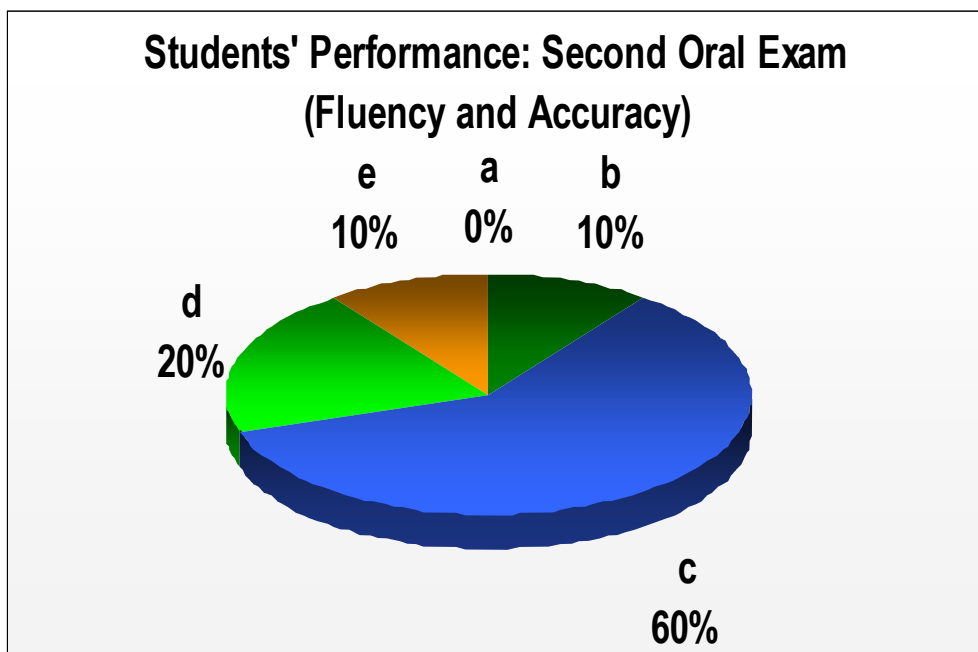




Students' Averages' Comparison

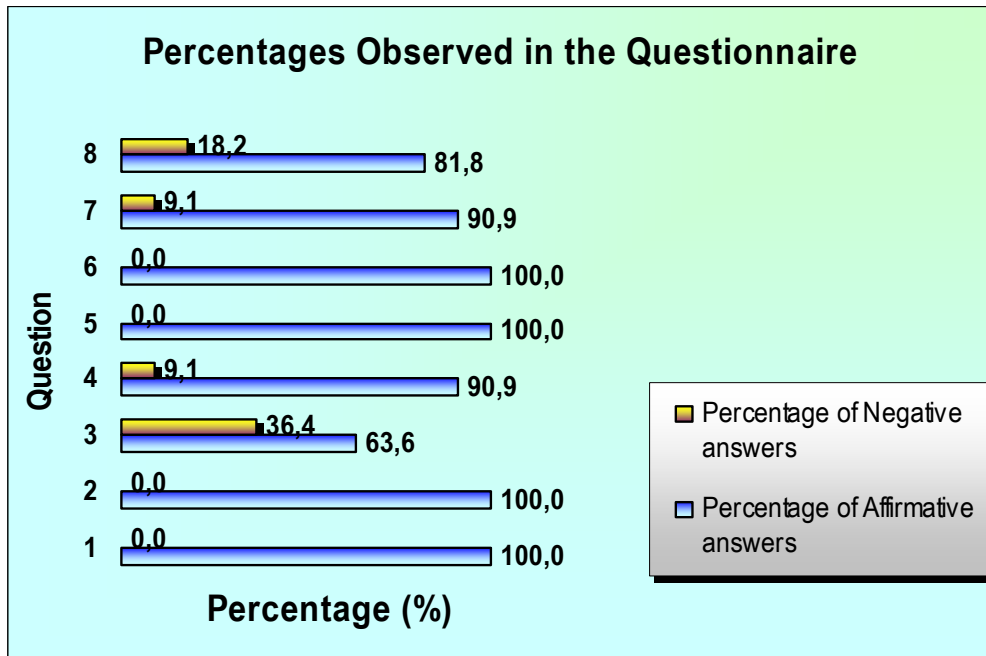


Students' Performances: Second Oral Exam (Categories)

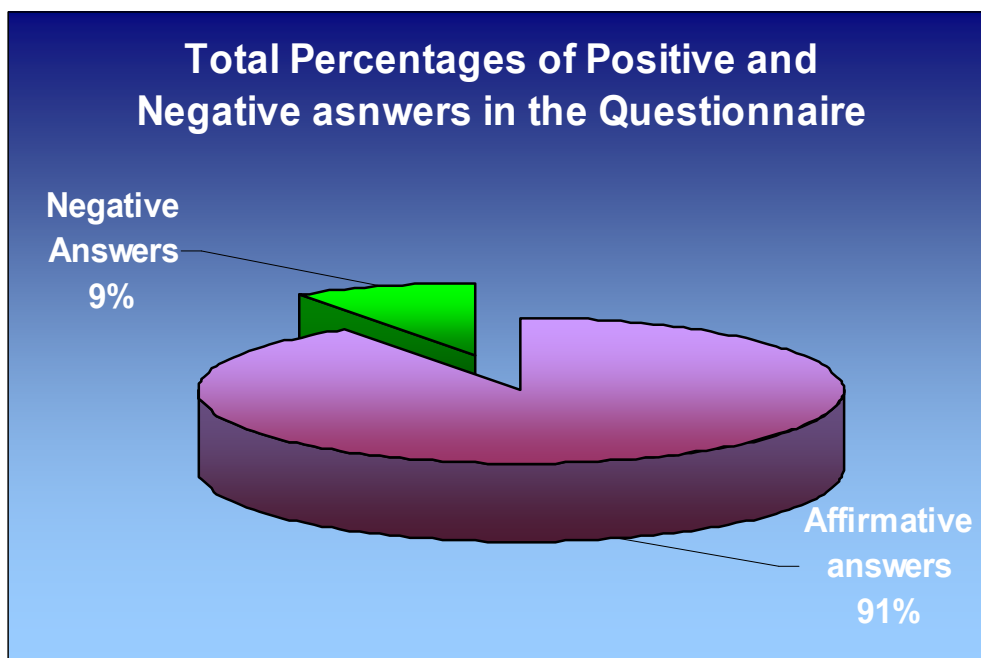


**Categories used on the oral assessment grid.** Each letter corresponds to a category - letter **a.** represents the lowest category and letter **e.** the highest. Notice that most of the students (60%) were ranged in the third category in terms of Accuracy and Fluency during the exam.

Students' Answers' Percentages in the Questionnaire



Students' Total Percentages in the Questionnaire



APPENDIX A. First Oral Exam Format

**University of Nariño**      **Language Center**      **Teacher:** Diego Fernando de la Portilla  
**Oral exam 1.**      **Level:** 1 **Group:** 6

You are going to be asked to answer ten of these questions during your oral exam. Be ready.

1. What's your full name?
2. How old are you?
3. Where do you live?
4. What's your phone number?
5. How many brothers and sisters do you have?
6. What do you do on weekends?
7. Do you have a pet?
8. Are you married?
9. Do you play any instrument?
10. Describe yourself/ your best friend/ your favorite actor, actress.
11. What's your mother's name?
12. What does your father do?
13. What time do you get up? Have breakfast? Have lunch? Go to the bed? Come to the university?
14. What do you do on your free time?
15. How does your mother/father/boyfriend/girlfriend/ looks like?
16. When is your birthday?
17. Who do you live with?
18. What's your favorite sport/movie/singer/actor/actress?
19. What do you want to study?
20. How many days a week do you study?
21. What do you usually do before an exam?
22. Do you always do the same things?
23. Describe your personality
24. How many hours a week do you work?
25. What job would you most like to do?
26. Why do you like living in the city?
27. What are some of the advantages of living in a city?
28. What are some of the disadvantages of living in a city?
29. Do you know the neighbors who live near you?
30. What's your favorite city? Why?
31. Do you live with your parents?
32. How many aunts and uncles do you have?
33. How many brothers and sisters do you have?
34. How often do you see your cousins?
35. How often do you see your grandparents?
36. What are your parents like?
37. Do you think your parents understand you? Why or why not?
38. What's your favorite color? Why?
39. How often do you go to movies?
40. How often do you rent videos?

APPENDIX B. Second Oral Exam Format

**University of Nariño**  
**Oral exam 2.**

**Language Center**  
**Level: 1**

**Teacher: Diego Fernando de la Portilla**  
**Group: 6**

You are going to be asked to answer ten of these questions during your oral exam. Be ready. Take care of the assessment sheet that accompanies this exam.

1. What kind of clothes do you buy most of the time?
2. Where do you like to spend your vacation?
3. How is the weather in this place?
4. What are the most dangerous natural disasters in your opinion? Why?
5. What is your opinion about the possibility of a natural disaster in our city?
6. What was, according to you, the most important event in the last years?
7. What was, according to you, the most important event in the history of our country?
8. What was, according to you, the most important event in the history of our planet?
9. What was the name of your first boyfriend/girlfriend/pet/friend?
10. Where did you grow up?
11. What did you do on your last birthday?
12. What things of your childhood do you miss?
13. Base on the map, give some directions to go from one place to another
14. What is a good idea for a first date?
15. Which places in the city do you visit most frequently? Why?
16. What do you like doing on your free time?
17. What are you doing on Friday, on weekend, on Saturday, next week, etc?
18. How did you like the English course at the Language Center?
19. Where is the Nariño Square (the Carnival square, Valle de Atriz shopping center) located?
20. What are your best memories of your childhood, your life in high school?
21. Who was the most important person when you were in the school?
22. What was the most important moment in your childhood?
23. What was your impression about your classmates at the beginning of the course?
24. Where did you go for your last vacation?
25. If you stayed in Pasto, what did you do?
26. When were you born?
27. Where were you born?
28. Where was your father/mother/sister/brother born?
29. What was the most important thing your parents taught you?
30. When did you start studying English?
31. How is the weather in Pasto? (Other cities)
32. How do you like the weather in Pasto?
33. Which city has the most terrific weather in Colombia? Why?
34. When you were a child, what did you want to be when you grew up?
35. When you were a child, what job did you want to have when you grew up?
36. Describe one of your best dreams
37. Why did you choose to study English at the University of Nariño?
38. What was the last movie you saw?
39. Why did you like it or why not?
40. What did you expect to learn during the English course?

APPENDIX C. Questionnaire Format

**Oral Assessment Criteria questionnaire**

**Please answer these questions according to your personal opinion. Mark with an x Yes or No according to your point of view.**

1. *Do you think it is useful to know the criteria used in an oral exam in advance?*

Yes \_\_\_\_\_ No \_\_\_\_\_

2. *Do you think that when you know the criteria of an oral exam, your scores can be improved?*

Yes \_\_\_\_\_ No \_\_\_\_\_

3. *Do you think the criteria help you get better grades?*

Yes \_\_\_\_\_ No \_\_\_\_\_

4. *Do you think that to use the criteria in advance permits you to focus on certain specific aspects of language that you should study?*

Yes \_\_\_\_\_ No \_\_\_\_\_

5. *Do you think your mistakes can be corrected more effectively when some criteria is given to you?*

Yes \_\_\_\_\_ No \_\_\_\_\_

6. *Do you think the criteria make an oral test a better planned/organized one?*

Yes \_\_\_\_\_ No \_\_\_\_\_

7. *In an oral test, do you prefer to know your grades immediately after the exam is finished?*

Yes \_\_\_\_\_ No \_\_\_\_\_

8. *Do you think oral tests are more acceptable and less threatening if they include some kind of criteria?*

Yes \_\_\_\_\_ No \_\_\_\_\_

APPENDIX D. Grid Format

University Of Nariño - Languages Center		
Oral Assessment Criteria grid		
Oral Assessment Criteria		
Subject:	<input style="width: 90%;" type="text"/>	Date: <input style="width: 90%;" type="text"/>
Student:	<input style="width: 95%;" type="text"/>	
<p><b>For the teacher:</b> If some other aspects you consider important, appeared during the exam, it is necessary to write them in the "Observations" column. Remember to read these criteria <b>BEFORE</b> the exam is carried out and to mark it immediately when the exam has finished. You can also use intermediate grades according to the students' performance. (e.g. 3,5 - 1,5 etc.)</p>		
Oral Assessment Criteria		
Grade description in terms of...		
Grade	Fluency	Accuracy
1	<input type="checkbox"/> One-word answers. No communication is established.	<input type="checkbox"/> Many mistakes in terms of grammar that interfere with the message.
2	<input type="checkbox"/> Short utterances that do not produce acceptable communication	<input type="checkbox"/> Many mistakes in grammar but the message can be inferred.
3	<input type="checkbox"/> Very limited utterances with a clear message	<input type="checkbox"/> Small amount of mistakes that do not interfere with the message.
4	<input type="checkbox"/> Acceptable interventions and longer periods of free speaking	<input type="checkbox"/> Low number of mistakes that do not interfere with the message. Clear answers.
5	<input type="checkbox"/> Expected level of speaking according to the level of the student. Long periods of free speaking and clear message.	<input type="checkbox"/> Few errors in grammar. Clear and coherent sentences in speech. The message is intelligible.
Average	<input style="width: 80%;" type="text"/>	<input style="width: 80%;" type="text"/>
Final Grade	<input style="width: 80%;" type="text"/>	
OBSERVATIONS		
<input style="width: 95%;" type="text"/>		<b>SIGNATURE</b>



APPENDIX E. Grids used in oral assessment practices

APPENDIX E1. Oral Presentation Checklist

Oral Presentation Checklist

Content	Resources	Organization	Presentation Aids	Delivery
I used a strong attention-getting device.	I used credible library resources.	I organized ideas in a meaningful way.	Presentation aids were used during the speech.	I maintained eye-contact most of the time.
I used words that the audience could understand.	I used credible Internet resources.	I stated my topic clearly in the introduction.	Presentation aids were relevant.	I spoke to the entire audience, not just one or two people.
My vocabulary was strong and unambiguous.	I used interviews with others as a resource.	I included necessary background information.	Presentation aids enhanced the speech or reinforced main points.	My pronunciation was clear and easy to understand.
I used facts and logical appeals where appropriate.	I cited my resources using the appropriate format.	A clear introduction was present.	Presentation aids did not distract the audience.	My rate of speech was not too fast or too slow.
I used opinions or emotional appeals where appropriate.	I gave credit to the resources in my speech.	The introduction included a statement of the main points.	Presentation aids were creative.	My volume was not too loud or too soft.
I used supportive details.	A bibliography was available.	The audience could predict the presentation's basic structure from the introduction alone.	Visual aids were easy to read or see.	My body language was not too relaxed or too tense.
The information I gave was valuable.	I put things into my own words.	The body of the presentation contained support for the main points.	Audio aids were easy to hear.	My voice varied in pitch; it was not monotone.
The information and arguments were easy to follow.	I was able to answer questions from the audience.	I used helpful transitions between main points (e.g.: "First of all ..." or "Similarly ..." etc).	Presentation aids contained no spelling or grammar errors.	I used meaningful gestures.
I stayed focused; I did not stray off my topic.		Idea flowed logically from one point to the next.		I used notes sparingly; I did not read from them.
I was well informed on my topic.		A strong conclusion was present.		I used standard grammar.
I presented information that others didn't know.		The conclusion had a clear call to action or belief.		I didn't hesitate or lose my place.
I was able to answer questions from the audience.		The conclusion was a creative summary of my topic.		I didn't use filler words (uhm, uh, ah, mm, like).
		The audience could distinguish the introduction, body, and conclusion.		I didn't call attention to errors by apologizing.
				I didn't fidget, rock back and forth, or pace.
				I maintained good posture.

APPENDIX E2. Oral Presentation Assessment Matrix

ORAL PRESENTATION ASSESSMENT MATRIX

Name \_\_\_\_\_ Grade \_\_\_\_\_  
 Teacher's name \_\_\_\_\_ Course \_\_\_\_\_  
 Assignment \_\_\_\_\_ Level \_\_\_\_\_

Assessment criteria are listed under each category. Though sufficiently broad and flexible, the criteria may not pertain to every presentation. Please use the following scale when using this assessment tool.  
 5=Excellent 4=Very Good 3=Good 2=Fair 1=Needs Improvement

Categories, Explanations and Criteria		Rating Score 5-1
<b>1. Organization: the way the parts of a presentation fit together.</b>		
<ul style="list-style-type: none"> <li>• clear introduction</li> <li>• clearly stated central idea</li> <li>• main points supporting central idea</li> <li>• obvious structure</li> </ul>	<ul style="list-style-type: none"> <li>• smooth transitions</li> <li>• effective conclusion</li> <li>• appropriate length</li> </ul>	
<b>2. Content: the topics included in the presentation; key concepts are important.</b>		
<ul style="list-style-type: none"> <li>• content fulfills project requirement</li> <li>• clear information</li> <li>• significant information</li> <li>• up-to-date information</li> </ul>	<ul style="list-style-type: none"> <li>• original thinking evident</li> <li>• sources varied</li> <li>• sources cited properly</li> </ul>	
<b>3. Adaptation: the visual aids and level of interest created by the presenter.</b>		
<p><b>To content</b></p> <ul style="list-style-type: none"> <li>• visuals appropriate</li> <li>• technology enhances presentation</li> </ul>	<p><b>To audience</b></p> <ul style="list-style-type: none"> <li>• speech adapted for audience</li> <li>• awareness of audience interest evident</li> <li>• audience interest and enthusiasm maintained</li> </ul>	
<b>4. Language: the presenter's use of oral language throughout the report.</b>		
<ul style="list-style-type: none"> <li>• clear and precise</li> <li>• suited to subject</li> <li>• rich vocabulary</li> </ul>	<ul style="list-style-type: none"> <li>• avoids clichés, archaic forms, redundancies</li> <li>• standard language forms and patterns used</li> </ul>	
<b>5. Delivery: the connection between speaker and audience.</b>		
<p><b>Voice</b></p> <ul style="list-style-type: none"> <li>• understandable, audible, articulate</li> <li>• appropriate, varied pitch</li> <li>• stress and volume appropriate</li> <li>• appropriate rate</li> <li>• pauses used for emphasis</li> </ul>	<p><b>Eye Contact</b></p> <ul style="list-style-type: none"> <li>• maintained through most of presentation</li> <li>• spread throughout audience</li> </ul> <p><b>Body Movement/Gestures</b></p> <ul style="list-style-type: none"> <li>• appears relaxed, confident, poised</li> <li>• hands and body used appropriately</li> </ul>	

Comments: \_\_\_\_\_

APPENDIX E3. Oral communication Assessment Grid

**Oral Communications Assessment Grid**

Subject: \_\_\_\_\_ Date: \_\_\_\_\_  
 Team/Student: \_\_\_\_\_ Reviewer: \_\_\_\_\_

Topic (Weight)	Unacceptable (0)	Marginal (1)	Acceptable (2)	Exceptional (3)	Points
<b>Organization &amp; Structure</b> (1)	θ Not possible to understand presentation due to absence of structure.	θ Difficult to follow presentation due to erratic topical shifts and jumps.	θ Most information is presented in logical order which is easy to follow.	θ All information is presented in a logical, interesting and novel sequence, which is easily followed.	
<b>Content &amp; Knowledge</b> (3)	θ No grasp of information. Unable to answer questions about subject.	θ Uncomfortable with information. Capable only of answering rudimentary questions.	θ At ease with content and able to elaborate and explain to some degree.	θ Demonstration of full knowledge of the subject with explanations and elaboration.	
<b>Visual Aids &amp; Neatness</b> (2)	θ No visual aids.	θ Occasional use of visual aids, however they barely support text or presentation.  Several misspellings and/or grammatical errors on slides.	θ Visual aids are related to text and presentation.  Minor misspellings and/or grammatical errors.	θ Text and presentation are reinforced by the use of visual aids.  Negligible misspellings and/or grammatical errors.	
<b>Delivery &amp; Speaking Skills</b> (2)	θ Significant mumbling and incorrect pronunciation of terms. Voice level too low or too high.  Monotonous, no eye contact, rate of speech too fast or too slow	θ Occasional mispronunciation of terms.  Little eye contact, uneven rate, only little expression	θ Voice is clear and at a proper level. Most words pronounced correctly.  Some eye contact, steady rate, excessively rehearsed	θ Clear voice and correct, precise pronunciation of terms.  Good eye contact, steady rate, enthusiasm, confidence	
<b>Presentation Length</b> (1)	θ Too long or too short.  +/- 10 minutes	θ +/- 6 minutes	θ +/- 4 minutes	θ +/- 2 minutes	
<b>OVERALL PERFORMANCE</b>	θ Unacceptable	θ Marginal	θ Acceptable	θ Exceptional	<b>TOTAL</b>
<b>POINTS REQUIRED</b>	<b>0-6</b>	<b>7-13</b>	<b>14-20</b>	<b>21-27</b>	

APPENDIX E4. Evaluating Student Presentations

		Evaluating Student Presentations			
		Developed by Information Technology Evaluation Services, NC Department of Public Instruction			
		1	2	3	4
		Total			
Organization		Audience cannot understand presentation because there is no sequence of information	Audience has difficulty following presentation because student jumps around.	Student presents information in logical sequence which audience can follow.	Student presents information in logical, interesting sequence which audience can follow.
Subject Knowledge		Student does not have grasp of information; student cannot answer questions about subject	Student is uncomfortable with information and is able to answer only rudimentary questions.	Student is at ease with expected answers to all questions, but fails to elaborate.	Student demonstrates full knowledge (more than required) by answering all class questions with explanations and elaboration.
Graphics		Student uses superfluous graphics or no graphics	Student occasionally uses graphics that rarely supported text and presentation.	Student's graphics relate to text and presentation.	Student's graphics explain and reinforce screen text and presentation.
Mechanics		Student's presentation has four or more spelling errors and/or grammatical errors.	Presentation has three misspellings and/or grammatical errors.	Presentation has no more than two misspellings and/or grammatical errors.	Presentation has no misspellings or grammatical errors.
Eye Contact		Student reads all of report with no eye contact	Student occasionally uses eye contact, but still reads most of report.	Student maintains eye contact most of the time but frequently returns to notes.	Student maintains eye contact with audience, seldom returning to notes.
Elocution		Student mumbles, incorrectly pronounces terms, and speaks too quietly for students in the back of class to hear.	Student's voice is low. Student incorrectly pronounces terms. Audience members have difficulty hearing presentation.	Student's voice is clear. Student pronounces most words correctly. Most audience members can hear presentation.	Student uses a clear voice and correct, precise pronunciation of terms so that all audience members can hear presentation.
					Total Points:

APPENDIX E5. European Council: Common European Framework for  
Languages: Learning, Teaching, Assessment



December 2003

DGIV/EDU/LANG (2003) 13

**CEF Performance Samples:**

**For Relating Language Examinations to the Common European Framework of Reference for Languages:  
Learning, Teaching, Assessment**

**English**

**(Swiss Adult Learners)**

**Brian North (Eurocentres)  
Gareth Hughes (Migros Club Schools)**

THE COMMON REFERENCE LEVELS: Manual Table 2.1

Level		Global Descriptor (Table 1)	Salient Characteristics (CEF Section 3.5, simplified)
User	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.	It cannot be overemphasized that Level C2 is not intended to imply native-speaker competence or even near-native-speaker competence. Both the original research and a project using CEF descriptors to rate mother-tongue as well as foreign language competence (North 2002: CEF Case Studies volume) showed the existence of bilingual speakers well above the highest defined level (C2). Wilkins had identified a seventh level of "Ambilingual Proficiency" in his 1978 proposal for a European scale for uni-credit schemes.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.	<b>Level C2, labelled 'Mastery',</b> is intended to characterise the degree of precision, appropriateness and ease with the language which typifies the speech of those who have been highly successful learners. Descriptors calibrated here include: <i>convey finer shades of meaning precisely by using, with reasonable accuracy, a wide range of modification devices; has a good command of idiomatic expressions and colloquialisms with awareness of connotative level of meaning; backtrack and restructure around a difficulty so smoothly the interlocutor is hardly aware of it.</i>
Proficient			<b>Level C1</b> was labelled <b>Effective Operational Proficiency</b> . What seems to characterise this level is good access to a broad range of language, which allows fluent, spontaneous communication, as illustrated by the following examples. Can express him/herself fluently and spontaneously, almost effortlessly. Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions. There is little obvious searching for expressions or avoidance strategies; only a conceptually difficult subject can hinder a natural, smooth flow of language. The discourse skills characterising the previous band continue to be evident at Level C1, with an emphasis on more fluency, for example: select a suitable phrase from a <u>fluent repertoire of discourse functions</u> to preface his remarks in order to get the floor, or to gain time and keep it whilst thinking; produce clear, <u>smoothly-flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices</u> .
User	B2+		This band (B2+) represents a <b>strong Advantage</b> performance. The focus on argument, effective social discourse and on language awareness which appears at B2 continues. However, the focus on argument and social discourse can also be interpreted as a new focus on discourse skills. This new degree of discourse competence shows itself in conversational management (co-operating strategies): give feedback on and follow up statements and inferences by other speakers and so help the development of the discussion; relate own contribution skilfully to those of other speakers. It is also apparent in relation to coherenciation: use a variety of linking words efficiently to mark clearly the relationships between ideas; develop an argument systematically with appropriate highlighting of significant points, and relevant supporting detail. Finally, it is at this band that there is a concentration of terms on negotiating.
Independent	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	Descriptors calibrated at <b>Level B2</b> represent quite a break with the content so far. For example at the lower end of the band there is a focus on effective argument, account for and sustain his opinions in discussion by providing relevant explanations, arguments and comments; explain a viewpoint on a topical issue giving the advantages and disadvantages of various options; develop an argument giving reasons in support of or against a particular point of view; take an active part in informal discussion in familiar contexts, commenting, putting point of view clearly, evaluating alternative proposals and making and responding to hypotheses. Secondly, running right through the level there are two new focuses. The first is being able to more than hold your own in social discourse: e.g. understand in detail what is said to him/her in the standard spoken language even in a noisy environment; initiate discourse, take his turn when appropriate and end conversation when he/she needs to, though he/she may not always do this elegantly; interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without imposing strain on either party. The second new focus is a new degree of language awareness: correct mistakes if they have led to misunderstandings; make a note of 'favourite mistakes' and consciously monitor speech for them; generally correct slips and errors if he becomes conscious of them.

THE COMMON REFERENCE LEVELS: TABLE 2.1 (continued)

Level	Global Descriptor (Table 1)	Salient Characteristics (CEF Section 3.5, simplified)
Basic User	B1+	This band (B2+) seems to be a <b>strong Threshold</b> performance. The same two main features at B1 continue to be present, with the addition of a number of descriptors which focus on the exchange of <b>quantities</b> of information, for example: <i>provide concrete information required in an interview/consultation (e.g. describe symptoms to a doctor) but does so with limited precision; explain why something is a problem; summarise and give his or her opinion about a short story, article, talk, discussion interview, or documentary and answer further questions of detail; carry out a prepared interview, checking and confirming information, though he/she may occasionally has to ask for repetition if the other person's response is rapid or extended; describe how to do something, giving detailed instructions, exchange accumulated factual information on familiar routine and non-routine matters within his field with some confidence.</i>
	B1	<b>Level B1</b> reflects the <b>Threshold Level</b> specification and is perhaps most categorised by two features. The first feature is the ability to maintain interaction and get across what you want to, in a range of contexts, for example, generally follow the main points of extended discussion around him/her, provided speech is clearly articulated in standard dialect; <i>express the main point he/she wants to make comprehensibly, keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.</i> The second feature is the ability to cope flexibly with problems in everyday life, for example cope with less routine situations on public transport; <i>deal with most situations likely to arise when making travel arrangements through an agent or when actually travelling; enter unprepared into conversations on familiar topics.</i>
	A2+	This band (A2+) represents a <b>strong Waystage</b> (A2+) performance. What is noticeable here is more active participation in conversation given some assistance and certain limitations, for example: <i>understand enough to manage simple, routine exchanges without undue effort; make him/herself understood and exchange ideas and information on familiar topics in predictable everyday situations, provided the other person helps if necessary; deal with everyday situations with predictable content, though he/she will generally have to compromise the message and search for words; plus significantly more ability to sustain monologues, for example: express how he feels in simple terms; give an extended description of everyday aspects of his environment e.g. people, places, a job or study experience; describe past activities and personal experiences; describe habits and routines; describe plans and arrangements; explain what he/she likes or dislikes about something.</i>
	A2	<b>Level A2</b> appears to reflect the level referred to by the <b>Waystage</b> specification. It is at this level that the majority of descriptors stating social functions are to be found, like use simple everyday polite forms of greeting and address; <i>greet people, ask how they are and read to news; handle very short social exchanges; ask and answer questions about what they do at work and in free time; make and respond to invitations; discuss what to do, where to go and make arrangements to meet, make and accept offers.</i> Here too are to be found descriptors on getting out and about the simplified cut-down version of the full set of transactional specifications in "The Threshold Level" for adults living abroad, like: <i>make simple transactions in shops, post offices or banks; get simple information about travel; use public transport; buses, trains, and taxis; ask for basic information, ask and give directions, and buy tickets; ask for and provide everyday goods and services.</i>
A1	Can understand and use familiar everyday expressions (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.  Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.	<b>Level A1</b> is the lowest level of generative language use - the point at which the learner can interact in a simple way, ask and answer simple questions about themselves, where they live, people they know, and things they have, initiate and respond to simple statements in areas of immediate need or on very familiar topics, rather than relying purely on a very finite rehearsed, lexically organised repertoire of situation-specific phrases.



**GLOBAL ORAL ASSESSMENT SCALE: Manual Table 5.4**

<b>C2</b>	<p><b>Conveys finer shades of meaning precisely and naturally.</b></p> <p>Can express him/herself spontaneously and very fluently, interacting with ease and skill, and differentiating finer shades of meaning precisely. <small>Can produce clear, smoothly-flowing, well-structured descriptions.</small></p>
<b>C1</b>	<p><b>Shows fluent, spontaneous expression in clear, well-structured speech.</b></p> <p>Can express him/herself fluently and spontaneously, almost effortlessly, with a smooth flow of language. Can give clear, detailed descriptions of complex subjects. High degree of accuracy; errors are rare.</p>
<b>B2+</b>	
<b>B2</b>	<p><b>Expresses points of view without noticeable strain.</b></p> <p>Can interact on a wide range of topics and produce stretches of language with a fairly even tempo. Can give clear, detailed descriptions on a wide range of subjects related to his/her field of interest. Does not make errors which cause misunderstanding.</p>
<b>B1 +</b>	
<b>B1</b>	<p><b>Relates comprehensibly the main points he/she wants to make.</b></p> <p>Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair may be very evident. Can link discrete, simple elements into a connected, sequence to give straightforward descriptions on a variety of familiar subjects within his/her field of interest. Reasonably accurate use of main repertoire associated with more predictable situations.</p>
<b>A2+</b>	
<b>A2</b>	<p><b>Relates basic information on, e.g. work, family, free time etc.</b></p> <p>Can communicate in a simple and direct exchange of information on familiar matters. Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident. Can describe in simple terms family, living conditions, educational background, present or most recent job. Uses some simple structures correctly, but may systematically make basic mistakes.</p>
<b>A1</b>	<p><b>Makes simple statements on personal details and very familiar topics.</b></p> <p>Can make him/herself understood in a simple way, asking and answering questions about personal details, provided the other person talks slowly and clearly and is prepared to help. Can manage very short, isolated, mainly pre-packaged utterances. Much pausing to search for expressions, to articulate less familiar words.</p>
<b>Below A1</b>	Does not reach the standard for A1
<ul style="list-style-type: none"> <li>• Use this scale in the first 2-3 minutes of a speaking sample to decide approximately what level you think the speaker is.</li> <li>• Then change to Table 5.5 (CEF Table 3) and assess the performance in more detail in relation to the descriptors for that level.</li> </ul>	

**ORAL ASSESSMENT CRITERIA GRID: CEF Table 3; Manual Table 5.5**

	RANGE	ACCURACY	FLUENCY	INTERACTION	COHERENCE
<b>C2</b>	Shows great flexibility reformulating ideas in differing linguistic forms to convey finer shades of meaning precisely, to give emphasis, to differentiate and to eliminate ambiguity. Also has a good command of idiomatic expressions and colloquialisms.	Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).	Can express him/herself spontaneously at length with a natural colloquial flow, avoiding or backtracking around any difficulty so smoothly that the interlocutor is hardly aware of it.	Can interact with ease and skill, picking up and using non-verbal and intonational cues apparently effortlessly. Can interweave his/her contribution into the joint discourse with fully natural turntaking, referencing, allusion making etc.	Can create coherent and cohesive discourse making full and appropriate use of a variety of organisational patterns and a wide range of connectors and other cohesive devices.
<b>C1</b>	Has a good command of a broad range of language allowing him/her to select a formulation to express him/herself clearly in an appropriate style on a wide range of general, academic, professional or leisure topics without having to restrict what he/she wants to say.	Consistently maintains a high degree of grammatical accuracy; errors are rare, difficult to spot and generally corrected when they do occur.	Can express him/herself fluently and spontaneously, almost effortlessly. Only a conceptually difficult subject can hinder a natural, smooth flow of language.	Can select a suitable phrase from a readily available range of discourse functions to preface his remarks in order to get or to keep the floor and to relate his/her own contributions skilfully to those of other speakers.	Can produce clear, smoothly flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices.
<b>B2+</b>					
<b>B2</b>	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much conspicuous searching for words, using some complex sentence forms to do so.	Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he or she searches for patterns and expressions, there are few noticeably long pauses.	Can initiate discourse, take his/her turn when appropriate and end conversation when he/she needs to, though he/she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc.	Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long contribution.
<b>B1+</b>					
<b>B1</b>	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circumlocutions on topics such as family, hobbies and interests, work, travel, and current events.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns associated with more predictable situations.	Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.	Can initiate, maintain and close simple face-to-face conversation on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
<b>A2+</b>					
<b>A2</b>	Uses basic sentence patterns with memorised phrases, groups of a few words and formulae in order to communicate limited information in simple everyday situations.	Uses some simple structures correctly, but still systematically makes basic mistakes.	Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident.	Can ask and answer questions and respond to simple statements. Can indicate when he/she is following but is rarely able to understand enough to keep conversation going of his/her own accord.	Can link groups of words with simple connectors like "and", "but" and "because".
<b>A1</b>	Has a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.	Shows only limited control of a few simple grammatical structures and sentence patterns in a memorised repertoire.	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.	Can ask and answer questions about personal details. Can interact in a simple way but communication is totally dependent on repetition, rephrasing and repair.	Can link words or groups of words with very basic linear connectors like "and" or "then".

**SUPPLEMENTARY CRITERIA GRID: "Plus Levels"**

	RANGE	ACCURACY	FLUENCY	INTERACTION	COHERENCE
<b>C2</b>					
<b>C1</b>					
<b>B2+</b>	Can express him/herself clearly and without much sign of having to restrict what he/she wants to say.	Shows good grammatical control; occasional "slips" or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.	Can communicate spontaneously, often showing remarkable fluency and ease of expression in even longer complex stretches of speech. Can use circumlocution and paraphrase to cover gaps in vocabulary and structure.	165-166- 94/261 Can intervene appropriately in discussion, exploiting a variety of suitable language to do so, and relating his/her own contribution to those of other speakers.	162 Can use a variety of linking words efficiently to mark clearly the relationships between ideas.
<b>B2</b>					
<b>B1+</b>	Has a sufficient range of language to describe unpredictable situations, explain the main points in an idea or problem with reasonable precision and express thoughts on abstract or cultural topics such as music and films.	Communicates with reasonable accuracy in familiar contexts; generally good control though with noticeable mother tongue influences.	Can express him/herself with relative ease. Despite some problems with formulation resulting in pauses and "cul-de-sacs", he/she is able to keep going effectively without help.	94/111 & 130 Can exploit a basic repertoire of strategies to keep a conversation or discussion going. Can give brief comments on others views during discussion. Can intervene to check and confirm detailed information.	<i>No descriptor available</i>
<b>B1</b>					
<b>A2+</b>	Has sufficient vocabulary to conduct routine, everyday transactions involving familiar situations and topics, though he/she will generally have to compromise the message and search for words.	<i>No descriptor available</i>	Can adapt rehearsed memorised simple phrases to particular situations with sufficient ease to handle short routine exchanges without undue effort, despite very noticeable hesitation and false starts.	Can initiate, maintain and close simple, restricted face-to-face conversation, asking and answering questions on topics of interest, pastimes and past activities. Can interact with reasonable ease in structured situations, given some help, but participation in open discussion is fairly restricted.	Can use the most frequently occurring connectors to link simple sentences in order to tell a story or describe something as a simple list of points.
<b>A2</b>					
<b>A1</b>					

## ANALYTIC RATING FORM: Manual Form B2

*Eurocentres (North 1991/1993)/Swiss Project (Schneider and North 2000)*

**LEARNER'S NAME**

Ihr Name/Votre nom:

Niveaus/Niveaux: **B, W, W+, T, T+, V, V+, E, M**

1. Initial Impression

Einstufung mit der Globalskala  
Classement - échelle globale

2. Detailed Analysis with Grid / Beurteilung mit Raster / Estimation - grille

<b>RANGE</b> Spektrum Etendue	<b>ACCURACY</b> Korrektheit Correction	<b>FLUENCY</b> Flüssigkeit Aisance	<b>INTERACTION</b> Interaktion Interaction	<b>COHERENCE</b> Kohärenz Cohérence

3. Considered Judgement

Abschliessende Einstufung  
Classement final

APPENDIX E6. European Council: Global Scale. Calibration of  
Certificates and qualifications to the Common European  
Framework Levels

# Self-assessment Grid



	A1	A2	B1	B2	C1	C2
<b>Understanding</b>  Listening	I can understand familiar words and very basic phrases concerning myself, my family and immediate surroundings when people speak slowly and clearly.	I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local area, employment). I can catch the main point in short, clear, simple messages and announcements.	I can understand the main points of clear, standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point in many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.	I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.	I can understand both and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer texts such as technical or literary works.	I have no difficulty in understanding any kind of spoken language whether live or broadcast, even when I do not understand every word. I can understand the accent.
<b>Reading</b> 	I can understand familiar names, words and very simple sentences, for example in notices and posters or in catalogues.	I can read very short, simple texts. I can find specific predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters.	I can understand texts that contain mainly of high frequency everyday or job-related language. I can understand and describe in detail events, feelings and wishes in personal letters.	I can read articles and reports connected with contemporary problems in which the writers adopt particular attitudes or viewpoints. I can understand contemporary literary texts.	I can understand long and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer texts such as technical or literary works.	I can read with ease virtually all forms of the written language, including abstract, structured or linguistically complex texts such as manuals, specialised articles and literary works.
<b>Speaking</b>  Spoken interaction	I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and helps me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.	I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.	I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar or personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).	I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussions in familiar contexts, accounting for and sustaining my views.	I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and personal purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.	I can take part effectively in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and coherently. I have a good command of the major means of oral communication. I can back-track and restructure around the difficulty to smoothly re-orient people as they are aware of it.
<b>Spoken Production</b> 	I can use simple phrases and sentences to describe where I live and people I know.	I can use a series of phrases and sentences to describe in simple terms my family and other people living conditions, my educational background and my present or most recent job.	I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my habits.	I can present clear, detailed descriptions in a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and concluding with an appropriate conclusion.	I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and which helps the recipient to notice and remember significant points.
<b>Writing</b> 	I can write a short, simple postcard, for example sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.	I can write short, simple notes and messages. I can write a very simple personal letter, for example thanking someone for something.	I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions.	I can write clear, detailed text on a wider range of subjects related to my interests. I can write an essay or report, basing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.	I can express myself in clear, well-structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select a style appropriate to the reader in mind.	I can write clear, smoothly-flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.