**UNIVERSITY OF EL SALVADOR**

**SCHOOL OF SCIENCES AND ARTS**

**FOREIGN LANGUAGES DEPARTMENT**

# "EVALUATION AND TESTING IN THE FOREIGN LANGUAGE DEPARTMENT OF THE UES".

***PRESENTED BY:***

**MAIRA YESENIA NAJARRO BARAHONA**

**CARLOS ROBERTO HERNÁNDEZ REYES**

**JOSÉ ROBERTO SALMERÓN CASCO**

**UNDERGRADUATE WORK PRESENTED TO OBTAIN THE DEGREE OF:**

**LICENCIATURA EN IDIOMA INGLÉS OPCIÓN: ENSEÑANZA**

**L10411**

***ACADEMIC ADVISOR:***

**LIC. ISRAEL OLIVA**

**Main Campus, March 12th 2009.**

**AUTHORITIES OF THE UNIVERSITY OF EL SALVADOR**

Ing. Rufino Antonio Quezada Sánchez

RECTOR

Arq. Miguel Ángel Pérez Ramos

ACADEMIC VICE RECTOR

MsC. Oscar Noé Navarrete Romero

ADMINISTRATIVE VICE RECTOR

Lic. Douglas Vladimir Alfaro Chávez

SECRETARY GENERAL

**AUTHORITIES OF THE SCHOOL OF ARTS AND SCIENCES**

Lic. José Raymundo Calderón Morán

DEAN OF THE SCHOOL OF ARTS AND SCIENCES

Dr. Carlos Roberto Paz Manzano

VICE DEAN OF THE SCHOOL OF ARTS AND SCIENCES

MsC. Julio César Grande Rivera

SECRETARY OF THE SCHOOL OF ARTS AND SCIENCES

**AUTHORITIES OF THE FOREIGN LANGUAGE DEPARTMENT**

MtI. Edgar Nicolás Ayala

HEAD OF THE FOREIGN LANGUAGE DEPARTMENT

MsD. Rodolfo Alexander Sibrián

GENERAL COORDINATOR OF THE GRADUATION PROCESS AND ADVISOR

## ACKNOWLEDGMENTS

This team first and foremost wants to hold in estimation God's blessing, wisdom and guidance for the achievement of this project. We are also grateful to our Academic Advisor, Professor José Israel Oliva who always believed in our capacity and supported us in the development of this project. Indeed, his valuable support, time, friendship and advice contributed enormously to the success of this paper.

We emphatically appreciate the cooperation offered by professors of the Department of the Foreign Languages such as Rhina Franco, Peter Salazar, Fidel Navidad, Alexander Sibrián, Dora Evelyn Portillo and some other professors who kindly collaborated with this team providing their sample tests and providing us information during the interviews. Thanks a lot for the help provided to this searching project. We hopefully think this investigation will contribute to the quality of the Department of Foreign Languages.

To our esteemed friends and those who anonymously collaborated in different ways and gave us support to fulfill this graduation paper. We will always be grateful for your help provided.

And finally, this job is dedicated to our lovely family who always stood by us. Your love and support all these years inspired us to conclude this major.

**INDEX**

***Page***

**Introduction** …….....…….…….……..……..… vi

**CHAPTER I RESEARCH TOPIC**

**I. Objectives** ……...…….…….…….….....…… 7

1.1 General Objective….…….…...…..…...…….....…7

1.2 Specific Objective….…….…….…....…….....… 7

1.3 Statement of the problem……...…..… 8

**CHAPTER II THEORETICAL FRAMEWORK**

2.1 The need for tests ……..……...…….…….……...…...… 10

2.2 What is evaluation? ……..…...…….……..…….....….. 12

2.3 Summative and formative evaluation …..……..…….……... 15

2.4 The importance of testing and test construction …..... 17

2.5 Types of tests ……...……..……..……...…… 20

2.6. Test Items ….…….……...……...…..... 22

2.6.1 About Selection items …….....……...……..... 22

2.6.2 About Supply items …….....……...……..... 24

**CHAPTER III METHODOLOGY**

3.1 Documentation.. ……...……..……..…….… 26

3.2 Searching techniques …….……...…….……...…… … 26

3.3 Difficulties ……....…….…...…...…..…..……. 27

**CHAPTER IV RESULTS**

4.1 Data Analysis ……...…….……..……...…...……..….....… 29

4.1.1 Analysis of the sample tests ……...…….…….....… 30

4.2 Findings ...…….……..……...…...……...……...… 35

4.2.1 Students' knowledge and competence is measured with more frequency than students's performance ……..…..…. 35

4.2.2 Measuring students' writing skills ……..... 37

4.2.3 Measuring students' speaking skills ……....…. 39

4.3 Analysis of the in-depth interviews ……..……...…...…….....… 42

**CHAPTER V CONCLUSIONS AND RECOMMENDATIONS.**

5.1 Conclusions …….....……..…….……...……... 48

5.1 Recommendations …….....……..…….……..…….. 49

**APPENDIXES.**

Appendix 1 ……...…….……..……..…...……..…..… 57

Appendix 2 ……...…….……..…….....…...……..… 59

**GLOSARY** …...……..……...…….……..... 72

**BIBLIOGRAPHY** ……....……...…..…….…….… 76

# INTRODUCTION

For generations before, tests have been used to measure and grade the abilities of learners everywhere. Testing has so far evolved into a thoughtful, well organized process to more accurately grade and assess the possession, comprehension and application of knowledge from learners.

Through the study of teaching methodologies, concepts like “evaluation”, “assessment” and “grading” have refined and gained meaningful value to help educators become more focused in the more needed areas. Thanks to the accurate analysis of test results an educator can now more efficiently help learners with different approaches in the concerning areas of learning.

This research project has been carried out with the help and cooperation from the Teaching Staff of the Department of Foreign Languages of the University of El Salvador who were kind enough to provide time and resources in the field of test-design, and test implementation. Hopefully, this report’s conclusions and recommendations will be deemed as useful towards their future work, as their contributions and help were to this project’s elaboration.

# CHAPTER I. RESEARCH TOPIC

## 1. OBJECTIVES.

### 1.1 GENERAL OBJECTIVE:

To carry out an analysis upon evaluation instruments applied in the Foreign Language Department of the University of El Salvador in subjects such as English Composition, Readings and Conversation, English Grammar and subjects of the area of abilities so that such analysis can be considered in the future to improve evaluation criteria in teachers.

### 1.2 SPECIFIC OBJECTIVES:

- To find out the form, design and implementation of the tests applied in the Foreign Language Department.

- To reflect upon the quality of these evaluation instruments based on the validity, reliability, practicality and beneficial effects on the process of teaching - learning.

## 1.3 STATEMENT OF THE PROBLEM

The Department of Foreign Languages of the University of El Salvador is an academic institution in charge of preparing students to become professionals in languages. This charge is mostly responsibility of the staff of teachers of this Department, since they are the individuals who carry out different processes to improve students' learning and prepare them to deal with issues concerning teaching and other areas in which languages are essential. An important process teachers are to carry out is the application of evaluation instruments to measure students' proficiency and command of the language in different subjects. This implies teachers are to judge students' competence and performance in the different language abilities which each subject suggests.

An analysis upon evaluation instruments applied by the teachers of the Foreign Language Department of the University of El Salvador in subjects of the area of abilities, English Composition, Readings and Conversation and English Grammar is one way to view the application of such instruments based on the different criteria and beliefs teachers include to judge students' competence and performance in the acquisition of English as a second language. In order to analyze these evaluation instruments, it is necessary to focus on these subjects only, for they deal with the development of the abilities and skills necessary to acquire a language in its four macro skills (reading, listening, speaking, and writing). This analysis is based on tests which take forms of quizzes, mid-term exams, oral quizzes, final exams, since they gain more emphasis at the moment of testing students' proficiency in the acquisition of English as a second language.

The implementation of evaluation instruments is an essential part teachers of the Department of Foreign Languages (DFL) develop in the teaching methodology and plans every semester. The application of tests in second language classrooms enables teachers to see whether or not a student is prepared to deal with more complex tasks in other subjects during several stages of the career, and it's a good opportunity for every student to demonstrate his / her skills and abilities required by the course objectives. Therefore, it is important to focus on an analysis about how these evaluation instruments are applied based on the validity, reliability, practicality and beneficial effects on the process of teaching - learning.

# CHAPTER II. THEORETICAL FRAMEWORK.

## 2.1 THE NEED FOR TESTS.

The application of evaluation instruments is implemented as an important method to measure an individual's knowledge, skill or ability in a specific area. The application of these instruments in learning English as a second or foreign language is widely important and necessary, for the facilitator or teacher has the chance to gain a concrete and formal evidence of the individual's improvement in the acquisition of the language. The Foreign Language Department of the University of El Salvador does not have an overall evaluation system in which all the teachers test the students' proficiency in the different subjects they study. Therefore, teachers have the chance to carry out the evaluation processes they consider correct in order to judge the students' competence and performance in the acquisition of English as a second language.

An important point to take into account at the moment of applying evaluation instruments is that language abilities and skills are not as simple to measure in an individual compared to pure sciences. Observers consider that "students' true abilities are not always reflected in the tests scores they obtain" (Hughes 1989; 2). This statement does not imply, of course, that any kind of testing of language abilities can not be reliable at all. Tests can be proven to be reliable if they are accurately linked to the instruction taught and the activities carried out during classes. If a test is not connected to these features, it will fail to measure accurately an individual's language abilities.

According to John Clark (1972) at the moment of evaluating students' competence and performance in the acquisition of a second language, teachers are to include two important factors: "First and foremost, teachers are to design tests or a different instrument of evaluation, according to students' needs". Since students are focused on acquiring a second language, they will need to improve listening, reading, writing and speaking skills. However, one specific skill will be more required than the other ones according to students' needs and according to the objectives of the subject. The second important factor teachers have to take into account, according to Clark, is "to carry out an evaluation according to the activities taught during classes". It's impossible to carry out a valid evaluation instrument isolating the topics, activities and instruction studied during a certain period of time.

The four language skills (reading, writing, speaking and listening) are fundamentally important when teachers judge students' performance and competence in the language, especially in the subjects of the area of abilities, since teachers have the duty to test out more than a language skill in a single test. However, an important point is to evaluate students using the language in more meaningful situations; that is, judge students' performance in real – life situations though this does not mean that the grammar aspects will be totally isolated from an evaluation activity.

## 2.2 WHAT IS EVALUATION?

It's very important to understand, before continuing to describe evaluation processes, the definition of evaluation itself. There are three more terms which are confusing and somehow related to evaluation itself: *testing, assessment* and *measurement*. Therefore, it's quite useful to make distinctions of these terms. First and foremost, evaluate means to express a judgement towards a specific behaviour. For instance, a student is paying attention to a professor lecturing a speech in a conference. At the end of the speech, this student might judge the content of the speech in different forms according to his personal opinion or beliefs. Was the speech boring or interesting? Did the audience agree with the ideas expressed in the speech? Are the ideas in the speech really worthy of being carried out? Etc.

Nevertheless, evaluate, subjected to a point of view considering the teaching and learning process, means to express a value to the skills, abilities or knowledge the student has developed in a required proof. In addition, evaluate, according to Shrock and Coscarelly (1999; 8), "is the process of making judgments regarding the appropriateness of some person, program, process or product for a specific purpose". Evaluation, thus, can be considered as an overall process which might include of course testing, measurement and assessment depending on the purpose of the evaluation process as well as based on quantitative and qualitative information.

"Testing", as it is defined by specialists, "is an activity whose main purpose is to convey (usually to the tester) how well the testee knows or can do something" (Ur, Penny

2002; 33). In this situation, the test taker has the chance to demonstrate a skill, ability or knowledge considering the nature of the test. He might succeed or fail depending on the degree of mastery of the contents the student has. At the end of the test, the facilitator will be able to determine the right answers and the wrong ones through a proper grading process, and he will have gained a formal and concrete evidence of the student's performance or competence. The student, on the other hand, is able to confirm whether or not he has the capacity to fulfil the requirements established in the objectives of a specific subject.

"Measurement", according to Shrock and Coscarelly (1999), "is the collection of quantitative data to determine the degree of whatever is being measured". When a teacher applies measurement at the moment of testing language abilities, he is able to determine the correspondent degree for the test taker. With this result, the teacher or facilitator realizes whether or not the individual has the required skills, abilities or knowledge in a specific area; or in a different stage, whether or not the student can be promoted to a higher level. In other words, it is a basic primary skill required for Grading but it can be developed into a strong tool for objectively judging a performance or another skill.

*Assessment*, on the other hand, usually involves qualitative information based not only in students' performance and competence, but also in his behaviour. This kind of process is surely more difficult to carry out in a second language classroom, for it demands to measure a student's profile based on behaviours such as participation, level of enthusiasm for the acquisition of a second language, etc. According to works on the

subject, "*Assessment is systematic information collection without reference to making judgements of worth*" (Heaton. 1990; 81) As a matter of fact, assessment is a qualitative process systematically carried out throughout observations based not only in a student's specific skill, but also in his behaviour itself. It is considered difficult due to there aren't any immediate results based on students' behaviour; the teacher or facilitator accumulates certain information related to the student's behaviours; and then he creates a belief which can be considered as a profile as well.

Evaluation is an important part of language teaching, for teachers and educators have the opportunity to test the language abilities and competences a student has in different areas of the language. In short, evaluations are helpful in the teaching-learning process, for they offer the teacher concrete evidences whether students are learning or not. When teachers test students' competence and proficiency in the learning of a second language, they usually develop an evaluation system either during the whole term of the course or during a specific instruction to be taught; an evaluation system which must be related to the objectives and activities taught during classes. During all these activities developed during classes, teachers usually consider important to employ two types of evaluation: Summative Evaluation and Formative evaluation.

## 2.3 SUMMATIVE AND FORMATIVE EVALUATION.

The application of summative and formative evaluation is quite common in second language classrooms. The former, or commonly considered as quantitative evaluation, estimates the achievements students have on the instruction they received during classes. It has been considered that "assessments are useful to evaluate, or sum up, the outcomes of instruction and are exemplified by end-of-chapter tests, projects, term papers, and final exams" (Daves 1989; Pg. 4). It's evident that these types of assessments usually judge students' outcomes by giving them a grade, placement or promotion according to the abilities or knowledge students had applied in the tests during a certain period of time. The great advantage of carrying out this type of evaluation is that the teacher has formal views of students' learning throughout evident results. In summative evaluation, teachers measure students' learning by giving them a number (probably from 1 to 10) a letter (A, B, C, D, E, F, etc.) or any other kind of mark. These results, however, do not include other important students' qualities such as the level of participation, enthusiasm and motivation they have towards an instruction taught.

On the other hand, there is a different type of evaluation considered for many experts as Qualitative or Formative evaluation. Such evaluation is almost never included in the outcomes and achievements students obtain at the end of the course, since it is considered quite difficult to measure. It takes place when the process of instruction is still going on. It suggests a kind of evaluation throughout observations teachers carry out in the classroom based on students' behaviour and motivation towards the language learning.

Summative evaluation is mostly employed to test students' proficiency in learning a new language, for it enables teachers testing and grading students in a more systematic form, and it's more convenient since teachers are supposed to follow the requirements school's authorities suggest to evaluate students. Therefore, each teacher is supposed to prepare certain instruments to test students in the different areas of language abilities. At the end of the course, semester or unit these results will be presented in the form of marks which let teachers and students themselves know how they did their performance and competence in a subject or specific area of the language.

While Formative Assessment might become in an useful tool for teachers in the application of tasks, it is not in the nature of education facilities to leave the final judgment of the students' proficiency to what could be considered a subjective opinion, one that could be based by the very personalities either of students or of teachers or even of both. Hence, all education facilities request the use of Summative / "official" assessment, considered by many specialists as Formal Evaluation, consisting mostly of formal exams, projects or even term papers. Specialists considered that "despite many educators' opinions and views on official assessments, learners, their parents, the authorities and the public in general affirm the importance of these, and take them quite seriously" (Daves 1989; Pg. 4). The nature of the importance of Formal assessment is that it is focused not really in the educating process itself but in its results; this formal education is meant to constitute proof of the actual performance of the student. In short, it doesn't ask "is this the best way to teach this or that?", but asks "Did the students *learn?* How can we/they actually *prove* it? ".

That is why formal assessment takes place usually at the completion of the instruction of a unit or chapter.

## 2.4 THE IMPORTANCE OF TESTING AND TEST CONSTRUCTION.

Teachers' responsibilities and tasks are both varied and challenging. Among these tasks, measuring students' proficiency in the acquisition of a second language is perhaps one of the most difficult and also one of the most discussed among education specialists. To assess a student's proficiency is to make sure that he or she has become able to perform or display a skill, an ability, or knowledge on a required subject. Nevertheless, testing students' proficiency in the acquisition of English as a second or foreign language must also fulfill the objectives by deciding important matters when actually testing students: what should be tested, the items to use, how long it should take and whether a textbook assessment or a teacher-made test should be used.

Tests are important to obtain information about students' learning. Grant and Hill (1992; 6) implies that "The first important decision when preparing to assess pupil achievement is to identify the information processes, and skills that will be tested. A valid achievement test is one that provides students a fair opportunity to show what they have learned from instruction". Grant and Hill state that the most important content which a test should be based on is to focus the attention in the objectives and the actual instruction that took place during classes, so that students have a fair chance to demonstrate what they have learned when they are taking tests.

A test should provide a meaningful indication of how well students have learned what they were taught. In other words, tests involve to measure students' skills and knowledge, for the test is to include the activities, indications and exercises which were taught during instruction. Appropriate tests contribute to a positive scenario for both students and teachers. They help teachers answer the question: "Have I been effective in my teaching?" "Have students really learned from the instruction taught?" Test, thus, are important to diagnose teachers' own efforts as well as those from students.

According to Harold S. Madsen, a Second-language teaching specialist, the importance of well-made tests in ESL (English as a Second Language) and EFL (English as a Foreign Language) can help students in two situations: "First, well-made tests can help create positive attitudes towards the class; In the interest of motivation and efficient instruction, teachers almost universally aim at providing positive classroom experience for the students". And the other is: "A second way that English tests can benefit students is by helping them master the language. They are helped, of course, when they study for exams and again when exams are returned and discussed" (1993). On the first part, Madsen implies that tests enable students to have positive attitudes for the instruction received, since they are bound to have the best success in their learning. The second way is focused on the importance of mastering the language, for tests motivate students to study more and commit themselves with their learning.

After considering the importance of the application of tests, the construction of the test itself must be an important step in measuring student's proficiency in the acquisition of

the language. This has to consist of the amount of contents to be included as well as many of the observations on the students' stamina and attention spans. The development of formal and valid tests requires the implementation of questions or items which have to fit students' necessities and be linked to the instruction and topics presented in classes previously. If these aspects are not included in an evaluation instrument, it might not be valid nor reliable, for students will not have the chance to demonstrate their abilities skills or knowledge fairly.

Other important factor is how long the test should last. It is advised that a test takes no longer than the regular length of the class period. This implies to take into consideration that little extra time that it takes to distribute the tests, the description of the items and the mechanics of the test, and still leave out some time off the class period so as to not rush students nor over-stress them.

Last, and only if the bibliography allows it and the teacher considers it as possibility, a teacher should decide whether to design the test himself or use a textbook format. It is advised that this is decided only after a teacher identifies the objectives that he or she has taught and now wants to assess, after all, the more that the instruction on the classroom deviates from the textbook objectives or from the lesson plan, the less valid a textbook test will be. "The main consideration in judging the adequacy of a textbook test is the match between its questions and what pupils were actually taught in class" Weir (1990). Nevertheless, sometimes the textbook test is not linked at all to the instruction and lessons taught by a specific teacher during a period of time. Thus, a textbook not always

provides students the chance to demonstrate what they really learned, for this type of test is frequently isolated from teaching itself.

## 2.5 TYPES OF TESTS.

The information obtained from tests concerning the process of learning a language will probably vary from situation to situation. Therefore, it is important to have test categorization to select a suitable test for a particular purpose. There are four types of tests which are appropriate according to the students' necessities: Proficiency tests, Achievement tests, Diagnosis tests, and Placement tests.

Proficiency tests are designed to measure students' abilities in the language according to the experience they may have had in that language. These types of tests are not based on the objectives or contents of a specific course, but on the previous experience the student has obtained. In addition, these tests are useful for a particular purpose. For instance, a proficiency test might be useful to determine whether or not a student has the capacity to follow a course of study at an American university. TOEFL tests are very common proficiency tests.

Achievement tests, on the other hand, are the most common in second language classrooms, for they measure students' progress in their own learning. In contrast to proficiency tests, achievement tests are *directly* related to language courses; the specific purpose is to determine how successful students have been in achieving the objectives of the course. Achievement tests can be helpful divided in two types: Final Achievement

tests and Progress achievement tests. The former is administered at the end of the course, which usually include all the topics or chapters taught during a period of time. Progress tests, on the other hand, measure the progress students have had at the end of a unit or topic taught. They usually determine how successful students have been in the mastering of specific abilities taught during a unit or a topic in general.

*Diagnostics tests* are usually useful to identify students' strength and weaknesses. They are important at the beginning of courses to realize about students' needs in order to create profiles about their abilities and knowledge. They might be extremely helpful when teachers want to develop an excellent long-term plan of instruction and a more clear content in the objectives. Nevertheless, diagnostic tests are almost never applied at the beginning of courses, since they are difficult to design.

And finally, *placement tests,* as their name suggests, provide information to place students in a stage or level according to their current knowledge and abilities. They are useful to indicate students the current capacity they have towards the language. An important point in placement tests is that it can be presented as either as an oral test or as a written test, depending on what it is measured in the student. In general, there is a combination of the two tests, oral and written, to find out what is the correct level of the student to place.

## 2.6 TEST ITEMS.

As it was discussed before, formal tests comprise a series of short evaluating instruments called questions or items. Each item must be brief to an extent, and set a clear problem for the student to solve. In written tests, there are two main types of test questions: selection items and supply items. As their name implies, selection items make the students choose the one and only correct answer from any number of choices, while *supply items* require the student to supply or construct an answer (or, in high-level content-objectives, justify it, as well). The most common examples of selection items are: true and false, multiple choice, and item-matching. Some examples of supply items are: short-answer and completion (also known as fill-in-the-blanks items) and essay items.

### 2.6.1 ABOUT SELECTION ITEMS.

The general aspect of selection items is that the answer or answers have been already provided and it's up to the student to *select* the correct answer. In the case of multiple-choice items, a "stem" presents the problem or question and then an array of options/ choices presents the answer; this is one of the most used kinds of items and a popular format for all kinds of tests of all levels, particularly when evaluating the application of Grammar and Semantics. With the appropriate introductory material, Multiple-choice items can be extremely effective at evaluating comprehensive reading. Another vantage aspect about multiple-choice items is that a test can have a large number of items to be answered in a short period, and of course, scoring is usually quick and objective. On the other hand, multiple-choice items, as most selection-type items, don't allow pupils to construct, organize, nor present their own answers, and is, to an extent,

susceptible to guessing; also, the time to construct these items counts against them since it is often hard to find adequate options. Teachers must also keep in mind that the reading ability can influence pupil performance, and must therefore familiarize students with the proper vocabulary and question format.

Another selection-type item is popularly referred to as True-and-False items and they require to be classified into a category out of two: true or false; yes or no; correct or incorrect. To construct these items the teacher must find and verify that the statements are irrefutably true or false and this has a tendency to stress memorization behavior, something that most high-education teachers want to avoid, since higher education seeks to develop high-level behavior like analysis, summarization and evaluation. While seemingly simple and susceptible to guessing (there are many jokes regarding this kind of item and these being answered by flipping coins), if the items are constructed adequately, true-and-false items can not only assess knowledge and comprehension but assess higher-level thinking and behavior but also, like multiple-choice, a large number of items can be given in a short time and the scoring process is facilitated.

The last selection-type item, matching items, consists of a column of "premises", a column of "responses" and directions for matching the two. Like multiple choice items, the answer to each stem or premise is among a number of possible choices, but the same array of possible choices is made available to all the other stems and premises. The main problem with matching items, according to some authorities on the subject, is that this format is limited to assessing low-level behavior. Moreover, it might lead students to "guessing" and

provide an erroneous idea that students really have the knowledge or ability in a subject or area of knowledge tested.

### 2.6.2 ABOUT SUPPLY ITEMS.

The main feature and what makes supply items differ from selection items is how these no longer provide with an answer but rather demand the pupil to provide (supply) one. Supply questions are most useful for assessing student's knowledge application and problem-solving skills. We can group supply items in two main types; the first is the short answer/completion items, both short-answer and completion item types present a student with a blank to fill or to be supplied with the appropriate response. The short answer is basically a question-answer item, while completion items present an incomplete sentence. Both often tend to assess mainly factual knowledge and comprehension but they reflect a bit more how well can students construct their answers. Also, another aspect a teacher might want to keep in mind is that it is relatively easy to write and design these items and that abroad range of knowledge can be assessed , however, scoring can be a bit time consuming.

The second main type of supply items are called essay items, which request the students to compose and construct their answers, thus being the most useful for assessing higher-level and thinking processes like analyzing, synthesizing, and evaluating. Essay items are also the primary way teachers can assess pupils' ability to organize, express and defend ideas. The biggest problem with essay items is the amount of time they consume to answer and score, but also the fact that students with high writing ability are favored while

those less proficient are underscored. The main advantages for a teacher to construct these test items are that they directly assess complex higher level behavior, it takes less time to construct than the other item types and that they assess integrative, holistic outcomes. The main disadvantages of essay items are the difficulty and length of time they take to score.

An evaluation plan takes all the important things above into consideration. Providing good instruction, identifying important objectives, and selecting appropriate question formats help accomplish the objectives of tests. Nevertheless, these important preparatory steps can be a failure if the test itself is confusing, misleading or constructed poorly, for students will not have the chance to show what they have learned.

# CHAPTER III. METHODOLOGY

## 3.1 DOCUMENTATION.

At the beginning of the searching project, several works related to the subject were read in order to become acquainted with concepts, definitions and theory regarding evaluation. After that, the most essential parts of these data were selected and then studied to create a theoretical frame and thus conceive ideas for the path of this research. The next step was to decide what searching techniques were necessary to acquire information about the implementation of tests and then analyze the results based on these techniques.

## 3.2 RESEARCHING TECHNIQUES.

An important technique implemented in this research was *data recollection.* This consisted on the recollection of sample tests considered as evaluation instruments applied by teachers. All these samples had already been implemented by teachers during the years 2004, 2005 and 2006. The subjects of concern were specifically related to the area of abilities, Grammar, as well as Composition and Readings and Conversation. The purpose of the recollection of these evaluation instruments was to analyze what type of items were applied in the test as well as how the tests were designed.

Another important technique applied in this research was *in-depth interview*. Some teachers of the FLD were chosen randomly from among the ones who had taught the subjects involved in the research. Then, a questionnaire was designed to find out teachers' opinions, difficulties and criteria they have at the moment they apply such evaluation

instruments in their respective courses. The purpose this interview was also to find out how teachers measure important skills such as writing and speaking in students.

The implementation of the interviews took a form of conversation so that teachers felt more confident and comfortable to describe the questions asked. They were asked about the aspects regarding the tests they usually develop such as 'how often they apply tests' and 'what kind of tests they apply in a specific course' as well as how they justified their validity, reliability and beneficial effects to students. Also, they were asked how they associate the objectives of the course and the instruction they teach to the tests they usually apply in their courses. Another important point teachers were asked during this interview was the reasons why students don't accomplish the objectives of the course and their expectations as teachers when their proficiency and competence is measured in these tests. Teachers provided their opinion according to their personal experiences in the subjects taught as well as the beliefs they have towards evaluation itself. These and other points were discussed during the interview.

### 3.3 DIFFICULTIES.

The most constant problem was to request teachers of the FLD to provide tests they had already implemented during the previous three years. Some teachers were glad to provide the tests they had already implemented in the courses they were in charge previously; but some others simply denied collaborating with this research project. The apparent reason teachers denied to collaborate providing these sample tests was that they did not have the tests at hand. Another significant difficulty was the classification of the

tests, for every test had to be classified according to the specific subject, semester and year (2004, 2005 and 2006).

The problems emerged during the classification was that some tests did not have the date they were implemented, so some teachers were asked in what semester and year these tests were implemented in order to classify them correctly. Another difficulty was that some tests did not specify whether it was implemented in the career of *Profesorado* or *Licenciatura*, so that some teachers were also asked in what specific career the tests were carried out to avoid confusions.

At the moment the theoretical frame, the tests samples and the interviews with teachers were ready. The next step was the analysis of the results. The results were related to the sample tests and the interviews carried out with teachers who taught the subjects involved in the research. The specific purpose of this analysis was to find out the form, design and implementation of the tests applied in the DFL to measure students' skills. When the analysis was completed, some reflections were made based on the quality of the evaluation instruments related to the validity, reliability, practicality and beneficial effects towards the process of teaching and learning English as a second language in the DFL..

# CHAPTER IV. RESULTS.

## 4.1 DATA ANALYSIS.

The analysis of this research was based on the results acquired from the application of searching techniques as Data Recollection and In-depth interviews. During the Data Recollection, the sample tests were classified based on the type of items. Then, in the same classification, all these evaluation instruments were split by year. In this way, it was possible to facilitate the analysis of these samples tests according to the type of items and the design itself.

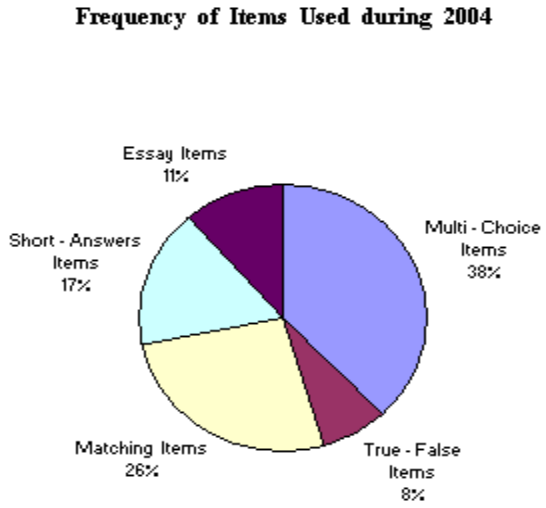
After the classification of these sample tests and its respective analysis, a random number of teachers who had taught the subjects during the selected time period involved in the research were subjected to in-depth interviews. The purpose of these interviews was to get insights of the tests development process and the reasons to use the evaluation items in language tests. Moreover, another important purpose for these interviews was to get information about the approaches and experiences teachers have towards evaluation itself. After that, the analysis of all these results was carried out as well.

### 4.1.1 ANALYSIS OF THE SAMPLE TESTS.

The following figures present the frequency each item is implemented in tests in the subjects of the area of abilities, Grammar I, Grammar II, English Composition I and English Composition II during 2004, 2005 and 2006.

***Figure 1.0***

| Type of Item | Frequency |
|---|---|
| Multi - Choice Ítems | 20 |
| Trae - False Ítems | 4 |
| Matching Ítems | 14 |
| Short - Answers Items | 9 |
| Essay Ítems | 6 |

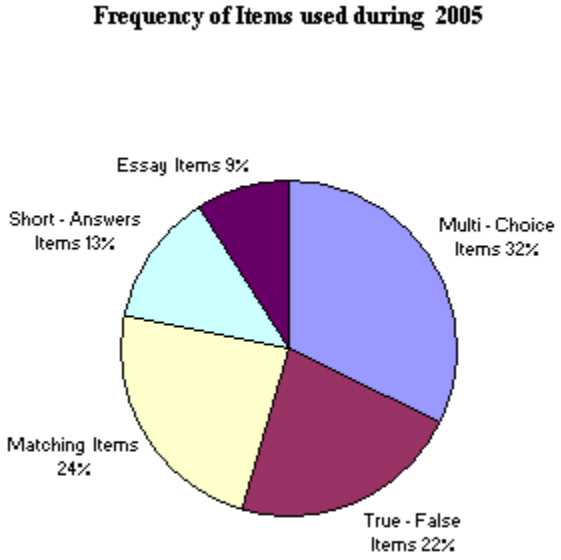



In the figure 1.0 it is possible to see that from 21 test samples collected, it was in 9 occasions short – answer items were implemented, whereas essay items present 6 occasions. In selection items, on the other hand, it is possible to see that true-and-false items present 4 occasions, matching items with 14, and multi – choice presents 20 occasions. This figure also shows the occasions transformed into percentages. The group of selection items (multi-choice items 38%, true-and-false items 8%, and matching items 26%) makes a total of 72%. On the other hand, the group of supply items (essay items 11%

and short-answers items 17%) makes a total of 28%. It is important to highlight, therefore, that the design of the test is mainly based on selection items, giving supply items a little level of frequency.

***Figure 2.0***

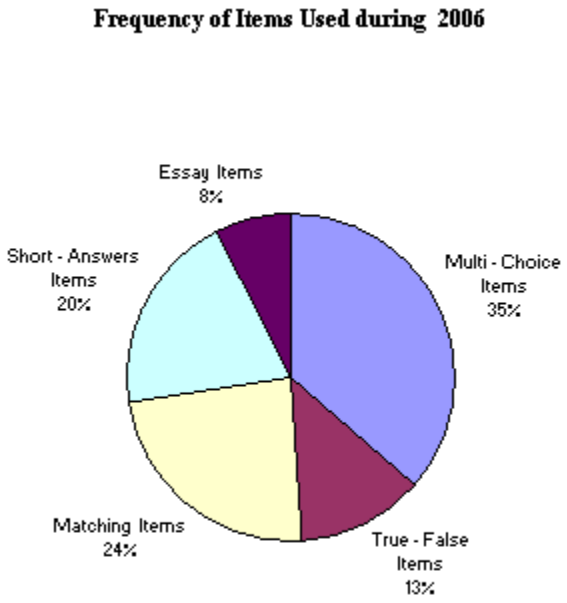
| Type of Item | Frequency |
|---|---|
| Multi - Choice Items | 22 |
| Trae - False Items | 15 |
| Matching ítems | 16 |
| Short – Answers Items | 9 |
| Essay ítems | 6 |




The figure 2.0 represents the frequency of items considered in the design of tests during 2005. It is evident, again, that teachers have a special consideration towards selection items. From 26 samples gathered, it was concluded that in 22 occasions multi – choice items were applied in these tests; true-and-false items present 22 occasions; and matching items present 16 occasions. The total of frequency in selection items is, thus, 78% (if multi-choice items with 32%, true-and-false items with 22%, and matching items with 24% are summed up together) On the other hand, supply items have a total of 22% (9% in Essay items and 13% in Short-answers items).

***Figure 3.0***

| Type of Item | Frequency |
|---|---|
| Multi - Choice Items | 29 |
| Trae - False Items | 10 |
| Matching Items | 19 |
| Short - Answers Items | 16 |
| Essay Ítems | 6 |

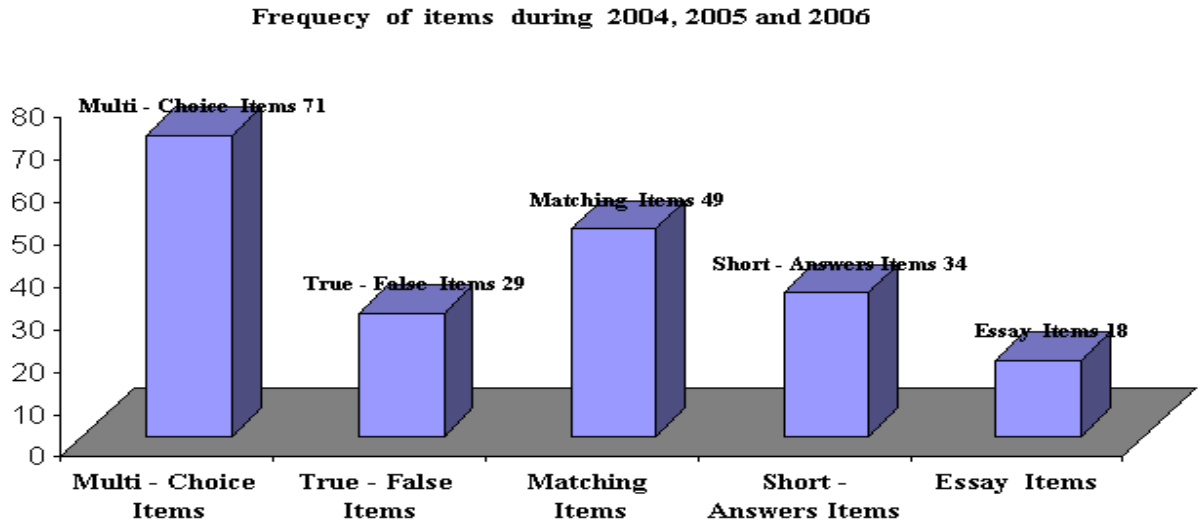


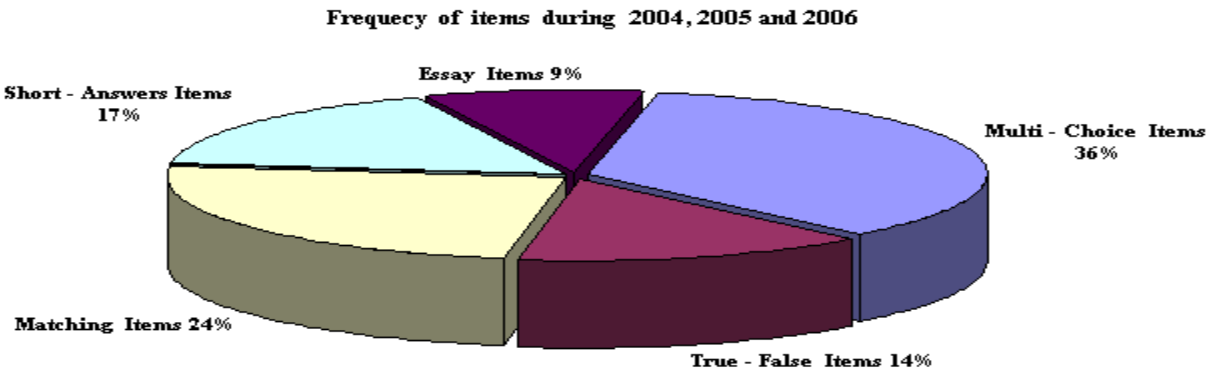

The figure 3.0 represents the frequency of items implemented during 2006. From 35 samples collected, Selection Items, as it was during 2004 and 2005, have the highest percentage. In this instance, the percentage of occasions is 72%. Short – answers items as well as essay items display a total of 28%. Essay items and short-answers items, though, are mostly considered in the design of mid-term exams and final exams as it was claimed above. Also, these types of items are commonly considered in subjects such as English Composition II and Readings and Conversations which generally measure students' proficiency of the language in more practical situations, and demand the students display more complex skills and abilities. Nevertheless, according to these figures and according

to the evidence of the previous collection of sample tests, it is possible to hold that selection items are employed with more frequency than supply items at the moment teachers design quizzes and end-of-chapter tests. Designing tests based mostly on selection items, thus, inhibit students display high performance skills such as analysis, evaluation or synthesis. This might lead teachers measure students' skill and abilities from a subjective point of view, since they are measuring students' competence of the language only.

***Figure 4.0***

| Type of Item | Frequency |
|---|---|
| Multi - Choice Items | 71 |
| True - False Items | 29 |
| Matching Items | 49 |
| Short - Answers Items | 34 |
| Essay Items | 18 |






The figure 4.0 represents the total of frequency of items during the whole three years (2004, 2005, and 2006) It can be observed that supply items (short-answers items and essay items) represent 26% of the total of frequency during the three years. Therefore, it is possible to identify, in this figure, the short frequency supply items are implemented in tests at the moment to measure students' proficiency of the language. Supply items enable students to exceed their skills since they have the chance to apply analysis, synthesis, descriptions and other important skills which might broaden students' master of the language. However, it can be observed that selection items have the highest percentage of the frequency during the three years (true and false items 14%, multi choice items 36% and matching items with 24%) making a total of 74%. At the moment of giving more concern to selection items it is possible to state that the students' knowledge of the language is measured with more frequency than students' performance of the language in more practical situations.

## 4.2 FINDINGS

### 4.2.1 STUDENTS' KNOWLEDGE AND COMPETENCE IS MEASURED WITH. MORE FREQUENCY THAN STUDENTS' PERFORMANCE.

The main focus of evaluation in the DFL is mostly based on measuring students' knowledge and competence of the language, isolating sometimes students' performance of the language in communicative purposes. What is the reason, then, to give more emphasis to the knowledge and competence rather than the performance itself? The fact of designing tests based on selection items somehow facilitates the construction of the test, for it is possible to include more items, and the student has more variety as well as more possibilities to succeed, since sometimes it requires little analysis or the student finds the correct answer throughout common sense. Furthermore, having a test based on selection items, the teacher has the chance to grade such tests very easily because of the nature of the items. However, does the application of selection items in the design of tests really benefit students in their proficiency and master of the language? A test based on selection items is probably successful measuring students' knowledge of the structure and form of the language, but how could the application of such knowledge be implemented in practical situations? Students really need to be measured in more meaningful situations beyond the structure and form of the language.

Measuring students' knowledge and performance is of course meaningful, for students have the chance how much they have learned from the structure and form of the

language. However, such knowledge would be useless if it is not applied to more practical activities. Therefore, it is important to design and implement evaluation instruments that enable students to display both their competence and performance. If teachers decide to implement tests which measure both students' competence and performance, they will be able to obtain a better evidence of students' profile.

It is possible to state that subjects such as English Grammar strictly suggest to measure students' knowledge and competence of the language at the moment to implement tests, for it is a subject that is supposed to deal with tense structures and specific rules of the language. Hence, is it possible to give room to measure students' performance in a subject strictly about competence such as English Grammar? The traditional form to measure students' skills in Grammar is always related with knowledge of the structures as well as rules. Nevertheless, such knowledge could be applied in practical activities such as presentations. Throughout presentations, students could have the chance to explain the form and use of either the structures or rules they are studying. In this way, students will have the chance to acquire a better idea of what they are studying, for they would have the chance to acquire the knowledge (at the moment they are getting prepared for such presentations) as well as apply the performance (at the moment they carry out the correspondents presentation)

Even though a test based in selection items is relatively easy to design as well as easy to score by teachers, it is important to highlight that it is subjected to several disadvantages: First, a student with this type of test has the chance to "guess" and, with

certain fortune, succeed in the test. This situation might lead to obtain erroneous results from students' real skills and abilities. Another disadvantage, as established above, is that the student does not have the opportunity to display high performance skills such as analysis, evaluation or synthesis. These skills are often measured during mid - term exams and final exams in the subjects of the area of abilities; nevertheless, since these types of tests are implemented in the middle and at the end of the semester, this means that students have the chance to display these skills perhaps twice or three times every semester. Hence, students do not have the chance to express opinions, present solutions to problems or describe events with more frequency during tests which give students the chance to be measured in more meaningful situations.

### 4.2.2 MEASURING STUDENTS' WRITING SKILLS.

It was possible, during the collection of sample tests, to notice that essay items are usually implemented in subjects as Composition I and Composition II. It is of course common to find these kinds of items in these evaluation instruments since such subjects are required to measure students' writing skills. Nevertheless, it was noticed that, in some evaluation instruments of these subjects, teachers do not have specific criteria to evaluate students' tasks. It was possible to view this particular factor in journals, essays, and other papers. It was observed that teachers just correct students' grammar and spelling mistakes in their works and then give a mark (from 1 to 10).

When teachers check and then give back students' composition papers, students just notice on their spelling and grammar mistakes they had in the papers. What could be the possible disadvantages of giving emphasis only to spelling and grammar mistakes in students' writing tasks? At the moment teachers are checking such papers, they usually isolate other important patterns such as measuring punctuation, vocabulary, organization and coherence of ideas students are expressing. The point is not only giving emphasis to the structure and order, but the focus should also be emphasized on the content of such ideas as well as the ideas students really mean. The application of such criteria and patterns should be specifically carried out in subjects as Composition and the subjects of the area of abilities, for such subjects give teachers the chance to measure students' high performance skills throughout the application of analysis, description , synthesis, etc.

At the moment of grading students' tasks related to writing skills concentrating the attention on spelling and grammar mistakes only, it is possible to isolate students' content of the ideas they really are expressing in these papers. Therefore, it is also important to focus on the ideas and opinions students present during such tasks, and give the content of the ideas they express value and notoriety, so that students notice on these important components at the moment they write. Hence, if teachers focus their attention on both spelling and grammar mistakes as well as the content of the ideas they present in these papers, it will be possible to encourage students to be more critical, creative as well as better users of the language at the moment they practice their writing skills in these tasks.

### 4.2.3 MEASURING STUDENTS' SPEAKING SKILLS.

Testing students' speaking skills is an important part in students' acquisition of the language. Measuring this skill is quite meaningful for teachers, since students provide a quick and practical evidence of their command of the language. The activities teachers carry out to measure this skill in the F.L.D. take forms of interviews, presentations, discussions etc. Nevertheless, the level of frequency these activities are implemented in the subjects of the area of abilities is very poor, for they are mostly considered at the end or at the middle of the course, usually at the same time mid-term exams are implemented. Therefore, students do not have the chance to be measured and judged displaying the real usage of the language through their speaking skills.

Students probably have the chance to practice their speaking skills during classes, but it is also important to test this skill with more frequency for evaluation purposes, for students usually apply more interest and effort when they know the activity they are going to perform represents a grade. The low frequency of testing students' skills in subjects of the area of abilities might lead to several disadvantages for students. A possible disadvantage could be that students do not show a good performance of the language in subjects such as Reading and Conversation, which demand students have an advanced level of their speaking skills. Another disadvantage could be reflected in students themselves as users of the languages: students are getting prepared to be language teachers; therefore, they are supposed to be excellent users of the language since they will probably be models for other learners in the future.

It is important to emphasize, however, that there are some factors that make it difficult to evaluate students' performance of the language in more communicative situations as well as make them display high performance skills during tests. The most obvious factor is that the number of students per course is very large. In spite the direction of the DFL demands the split of two or even three groups for every subject, especially those subjects of the area of abilities, there is an average from fifteen to twenty-five students in every group.

It is quite difficult, therefore, to evaluate students' performance of the language with more frequency, especially at the moment to measure students' speaking skills; for it takes a lot of time to evaluate every student at a time in an interview based on a specific topic of interest or a different practical activity. This situation might be evident during presentations. A group of students are carrying out a presentation, whereas the rest is usually doing nothing during such presentations, or some others do not attend classes at the moment the presentations are carried out. The development of these activities, thus, might lead to delays in the accomplishment of the course programs due to the large number of students to be evaluated.

Testing and evaluation of language skills are a very important part of language teaching. Teachers' duties are both challenging and demanding at the moment of testing students' proficiency in the acquisition of English as a second or a foreign language. Testing provides significant information about students' real level as well as an opportunity for themselves to commit with their own learning, for taking a test demands to study and

get prepared for facing such responsibility. Therefore, the design of a test is vital to achieve this information and behaviour from students. A test or another evaluation instrument has to be connected to the real instruction studied and practiced during classes, for it gives students a fair chance to prove the real capacities they have towards the learning of the language.

During this analysis, it was noticed that, at the moment teachers of the FLD design tests and measure students' proficiency of the language, the main emphasis is most of the time based on testing students' knowledge of the structures of the language: e.g. the difference of simple present and present progressive, or how it is the structure of present perfect in questions, affirmative and negative sentences, etc. Testing students' knowledge of the structures of the language is vital; nevertheless, sometimes it is isolated the fact of testing students during the real usage of these and other structures in real life situations or in communicative activities. It is possible that sometimes these activities are somehow implemented during classes, but they also need to be implemented for evaluation purposes, for students have the chance to prepare a better performance considering that such activities are essential to have a good grade.

The knowledge of grammar is completely essential in the learning of a second language, for it is the basis to learn other important features of the language. Nevertheless, the knowledge of grammar might be useless if it is not connected to more meaningful situations such as essays, conversations, oral presentations, oral and written descriptions etc. These procedures, though, are time consuming and, of course, more demanding on

teachers, since the DFL host a large number of students, especially in the subjects of the area of abilities. Therefore, it is fundamental to carry out new evaluation methods that measure both students' knowledge and performance at the same time, so that students have the chance to broaden their skills and be more prepared at the moment they become professionals.

## 4.3 ANALYSIS OF THE IN – DEPTH INTERVIEWS.

For the implementation of these in-depth interviews, first and foremost, a questionnaire was designed in order to recollect information and data from teachers of the DFL (See Appendix 1). The purpose of these interviews was to find out teachers' opinions, difficulties and criteria they have at the moment they design and apply evaluation instruments in their respective courses. At the moment the questionnaire was ready, a group of teachers were chosen randomly among the ones who had taught the subjects involved in the research. Then, this group of teachers was informed about the purpose of the interview and then the interview itself was scheduled with each teacher according to the day and the time it was perfect for them.

An important factor taken into account during the interview was to make it as a form of conversation so that teachers could go further in the description of their opinions, difficulties and criteria they have at the moment they deal with issues related to testing and evaluation. In this form it was possible to obtain more specific information related to the criteria and opinions they have at the moment of designing tests. Another important factor of this interview was not only to gather information about the design and application of

tests in the subjects of concern, but also to get insights about why students sometimes do not fulfil the objectives of the course as well as how they connect the objectives of the course to the evaluation instruments they apply, etc.

The interview itself started asking teachers about their experience in the FLD so that they could feel comfortable at the moment of starting the interview. After that, the first question of interest was related about how they evaluate their students in the subjects of interests for this searching project. Most of them claimed they implemented the traditional forms of evaluation which include written quizzes, mid-term exams, and incorporate the assessment of the four macro skills or a specific skill, depending on the subject they teach. At the beginning of every semester teachers also incorporate the evaluation system they plan to carry out in the specific subject they are in charge.

An important point some teachers emphasized was that some of them usually work as a team to design tests, so they have a special organization for evaluating a specific subject. This is indeed a positive experience for them, for a teacher has the chance to learn from other teacher's evaluation criteria, and the test itself is subjected to different ideas discussed previously. Moreover, when teachers design these evaluation instruments as a team, it is possible to include more variety and innovation to the tests they design. In spite some teachers carry out these new forms of designing tests, they also stated that some other teachers implement text-book tests for quizzes or end-of-chapter tests so that they "don't complicate their life". These decisions are sometimes considered erroneous since the text-

book test is not totally connected to the real instruction taught during classes; therefore, it is not a fair chance for students to demonstrate what they really learned during classes.

Another question of interest during these interviews was whether or not teachers carry out oral exams during their courses and how they design these evaluation instruments to measure students' oral skills of the language. Most teachers answered these questions based on the subjects of the area of abilities or the subjects of the readings and conversation. Unfortunately, as it was claimed in the analysis of the sample tests, teachers usually carry out oral exams in the middle or at the end of the semester. An important point during the oral exams teachers implement during these periods of time is that some teachers do not have specific patterns to measure students' oral proficiency. Most of the time teachers just carry out either an interview or require students develop oral presentations. So, during these activities, some teachers usually write down the mistakes the student produces at the moment of speaking; and according to the number of mistakes, that is the final grade they have, isolating other important patterns such as vocabulary, pronunciation, coherence, etc. Nevertheless, some other teachers usually carry out these interviews as a team: two teachers interview a student, and at the end of this interview they open a discussion about how well the student carried out the interview, and the student's grade is subjected to discussion.

The next question teachers were inquired about was related to the reasons some students do not accomplish the objectives of the course and their expectations as teachers. The answers teachers described were quite mixed, and the reasons were both internal and

external. The external reasons were related to the strikes in the university which somehow delays the development of the course programs. In addition, some companies such as Dell, Sykes and Teleperformance have been establishing call centers in the city, and students have the chance to work part-time for these companies. Therefore, students do not have the chance to commit themselves completely with their learning process.

On the other hand, there is a larger variety of internal reasons why students do not fulfil the objectives of the courses which teachers have observed in the development of the courses they teach. One of these reasons is related to students' attitude towards their own learning. They do not take their learning seriously nor accept the different responsibilities required in a career. Another reason is that every course, especially the courses of the area of abilities, is very large. In this scenario, students' learning styles are different, and another point is that some students have had previous knowledge of the language and some others are beginners. So teachers usually have to deal with these issues not only during classes but also during the application of evaluation instruments.

The next question of interest was based on teachers' own experience in the different courses they have taught. The question was specifically related about how tests and other evaluation instruments affect students' proficiency of the language. The answers they expressed were quite similar: Tests can affect students positively or negatively, depending on students' own attitude towards their learning. According to some teachers, tests are useful, for they motivate students study to get prepared for their evaluations. They consider that some students usually study and commit themselves with their own learning when they

know they will be measured in a test or a different evaluation instrument; so students worry themselves.

Tests might affect students negatively at the same time. Since the outcomes of tests promote students to a higher level or enable them to pass or fail a subject, students usually study just to pass a subject or to obtain a good grade in a test. Furthermore, some teachers consider that when students fail an important test, they feel so frustrated that they don't want to continue either with the subject or the career itself. In other words, the result of a test affect some students negatively, since a negative outcome sometimes is a signal that they are not capable or they cannot deal with the work load they have in their career.

Finally, as a last enquiry, teachers were asked about what their personal opinion was about the application of tests to measure students' proficiency in the acquisition of English as a second language in the FLD. First of all, most teachers agree that tests are absolutely necessary to measure students' proficiency, for the outcomes give them concrete evidence of their learning. Therefore, they consider that the application of *summative* evaluation is the best, for the results show numbers, and they give them a concrete evidence of students' learning in a specific subject. Furthermore, teachers agree that the tests and other instruments of evaluation they apply on the different subjects have to be linked to the objectives of the course so that they accomplish their commitment with students' learning. Nevertheless, some teachers claim that the evaluation processes most teachers of the FLD apply are quite traditional, and there should be a change in the way teachers evaluate. Some teachers consider that there should be other measuring tools and base teacher's opinions

and criteria on the student's capacity. Moreover, some teachers agree that they have not been trained to evaluate or to design tests, and also they consider that there should be supervision in the different evaluation processes teachers carry out in the FLD.

# CHAPTER V. CONCLUSIONS AND RECOMMENDATIONS.

## 5.1 CONCLUSIONS.

Testing and evaluation of language skills are a very important part of language teaching. Teachers' duties are both challenging and demanding at the moment of testing students' proficiency in the acquisition of English as a second or a foreign language. Testing provides significant information about students' real level as well as an opportunity for themselves to commit with their own learning, for taking a test demands to study and get prepared for facing such responsibility. Therefore, the design of a test is vital to achieve this information and behaviour from students. A test or another evaluation instrument has to be connected to the real instruction studied and practiced during classes, for it gives students a fair chance to prove the real capacities they have towards the learning of the language.

During this analysis, it was noticed that, at the moment teachers of the FLD design tests and measure students' proficiency of the language, the main emphasis is most of the time based on testing students' knowledge of the structures of the language: e.g. the difference of simple present and present progressive, or how it is the structure of present perfect in questions, affirmative and negative sentences, etc. Testing students' knowledge of the structures of the language is vital; nevertheless, sometimes it is isolated the fact of testing students during the real usage of these and other structures in real life situations or in communicative activities. It is possible that sometimes these activities are somehow implemented during classes, but they also need to be implemented for evaluation purposes,

for students have the chance to prepare a better performance considering that such activities are essential to have a good grade.

The knowledge of grammar is completely essential in the learning of a second language, for it is the basis to learn other important features of the language. Nevertheless, the knowledge of grammar might be useless if it is not connected to more meaningful situations such as essays, conversations, oral presentations, oral and written descriptions etc. These procedures, though, are time consuming and, of course, more demanding on teachers, since the DFL host a large number of students, especially in the subjects of the area of abilities. Therefore, it is fundamental to carry out new evaluation methods that measure both students' knowledge and performance at the same time, so that students have the chance to broaden their skills and be more prepared at the moment they become professionals.

## 5.2 RECOMMENDATIONS.

In this analysis the subjects of concern have been English Composition, English Grammar, Readings and Conversation and the subjects of the area of abilities. These were the subjects of concern because they deal with students' abilities and skills at the moment to produce the language in communicative purposes. However, during the analysis of the sample tests, it was evident that teachers' main concern is measuring students' knowledge and competence of the language. Therefore, it should be considered to measure not only students' competence but also their usage of the language in communicative purposes. These instruments should be connected to real situations, so that they enable students to

produce and use the language in more practical activities. Furthermore, the knowledge of the grammar structures should also be implemented in high-performance skills so that students develop important skills such as analysis, synthesis, evaluation and criticism.

The importance of measuring students' performance should be based on making students have a chance to give their best in practical activities and demonstrate the level of both competence and performance they have towards the language. In short, testing students' performance and competence gives teachers the possibility to obtain better information about students' real capacities in the master of the language. The order of structures, the use of tenses and grammatical rules, as well as other important aspects as the cohesion, vocabulary and coherence should be taken into account not only in written and grammar aspects, but also apply this knowledge in more communicative purposes.

Measuring students' performance of the language, thus, give teachers a real view of students' profile as well as a fair chance for students to demonstrate their real capacities and skills in the use of language. Therefore, it is recommendable the implementation of oral exams with more frequency, especially during the courses of the area of abilities. The design of oral exams is an excellent evaluation instrument to measure students' performance of the language. In addition, the implementation of oral exams which take forms of presentations, interviews, and conversations are usually connected to real life situations and practical communicative activities. However, before the implementation of an oral exam, it is also important to consider the patterns to be measured in students'

performance. That is, design an evaluation instrument with the aspects to evaluate in students' speaking skills.

During the application of the searching techniques which include the analysis of the sample tests and the in-depth interviews, it was possible to view the low level of frequency oral exams are applied during courses, especially in the subjects of the area of abilities. There are factors which make it difficult to carry out oral exams during courses such as the large number of students per course. Moreover, teachers sometimes do not have specific patterns to measure students' use of the language in speaking activities, since they have never been trained to design such patterns. This team, therefore, recommends the use of the following sheet which might enable teachers to carry out oral exams and apply fair patterns to measure their speaking skills.

University of El Salvador
Department of Foreign Languages
Oral Evaluation Sheet

Teacher's Name: _____
Level: _____
Date: _____

Student's name: _____ Grade: _____

| **ASPECTS** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| COMPREHENSION | | | | | | | | | | |
| GRAMMAR | | | | | | | | | | |
| VOCABULARY | | | | | | | | | | |
| FLUENCY | | | | | | | | | | |
| COHERENCE | | | | | | | | | | |
| PRONUNCIATION | | | | | | | | | | |

OBSERVATIONS: _____

_____

_____

University of El Salvador
Department of Foreign Languages
Oral Evaluation Sheet

Teacher's Name: _____
Level: _____
Date: _____

Student's name: _____ Grade: _____

| **ASPECTS** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| COMPREHENSION | | | | | | | | | | |
| GRAMMAR | | | | | | | | | | |
| VOCABULARY | | | | | | | | | | |
| FLUENCY | | | | | | | | | | |
| COHERENCE | | | | | | | | | | |
| PRONUNCIATION | | | | | | | | | | |

OBSERVATIONS: _____

_____

_____

An important advantage of testing students' speaking skills using these patterns is that teachers have the chance to measure students as they are speaking. So it is possible to measure not only students' grammar at the moment of speaking, but also measure other important aspects such as vocabulary, fluency and pronunciation which are essential for teachers notice students' real speaking skills. The versatility of this evaluation instrument is an important advantage the moment teachers are measuring students' speaking skills, for it is possible to use the same sheet at the moment of a presentation, an interview or a conversation itself. This sheet also provides a space for observations so that teachers have the chance to write down on such space, suggestions for improvement or students' mistakes at the moment they are speaking. Students, thus, will have the chance to notice what they did well and wrong when the results are given back to them.

Another important advantage of the application of this evaluation instrument to measure students' speaking skills is that it is certainly easy to carry out by teachers. During an oral exam, applying this evaluation instrument, teachers will notice students' speaking capacities just throughout observation and listening to them at the moment they are producing the language, and then decide what number (from 1 to 10) they can give to every aspect of consideration. It is possible to carry out this process with two students at the same time when they possibly carry out a conversation, and teachers will be able to reduce the time in spite the group of students is very large. Measuring two students at the same time will also give teachers the chance to test their interaction and, of course, apply this activity into real – life situations, and have a chance to evaluate students' performance in a more interactive form.

The evaluation instruments which measure students' written skills are another element which should be taken into account with more frequency, especially in the courses of the area of abilities. It was stated above that supply items are commonly included in the design of mid term exams and final exams, for students are required to write descriptions or short essays in these tests and they usually represent a higher percentage in evaluation systems. Nevertheless, these kinds of items should be included during quizzes and end-of-chapter tests as well. Measuring students with tests containing more essay items will give them the chance to apply their knowledge of the structures in more practical situations and help to organize their ideas with more order and coherence. Essay items should also be applied with more frequency in subjects of the area of abilities so that students have a basis when they take Composition I and Composition II.

Establishing patterns to measure students' writing skills is another important fact which should be improved in both the criteria and the evaluation system teachers apply in courses. It was observed during the recollection of tests, and as it was stated in the analysis of the results, that at the moment some teachers check journals, descriptions, essays and some other papers, they just correct grammar and spelling mistakes and then provide a grade, which is usually from 1 to 10. This was observed not only in evaluation instruments of subjects as Composition, but also in other subjects of the area of abilities. In these situations, teachers should also have specific patterns to measure students' writing skills in such essays, journals and descriptions. When teachers judge these papers, they not only should focus their attention on grammar mistakes and spelling, but also in some other

important aspects as organization of ideas, vocabulary and punctuation. So at the moment teachers give a mark to such papers, they should also present what is the mark for their vocabulary, the spelling, organization of ideas and grammar as well. In this form the student will have the chance to realize what his weaknesses and strengths are in their writing skills and of course try to improve their weaknesses.

The DFL does not have financial resources to carry out special trainings for teachers in the improvement of the design and implementation of language tests and evaluation in general. It is possible, however, to carry out meetings and colloquiums among teachers themselves to discuss and analyze the evaluation programs they carry out in the courses. Teachers can present the evaluation methodology they want to implement before the courses or the semester start, or even before the implementation of examinations. In the development of these meetings and colloquiums, they will obtain opinions and constructive criticism from other teachers in order to learn from the opinion of others and thus develop a better evaluation program.

The implementation of new evaluation instruments which test students' performance and competence in the acquisition of the language will give teachers a clear view of students' real profile in the command of the language. It, of course, will be quite demanding, for it requires from teachers more effort due to the construction and design of such instruments might be time consuming and stressful. However, the functioning of such evaluation instruments in teachers' evaluation plans will certainly lead to obtain better outcomes from students, since they will get prepared not only to be measured according to

their competence and knowledge of the language, but also at the moment they use the language in more practical activities.

Measuring both students' competence and performance of the language will provide teachers better information about students' real abilities and skills. The main objective to evaluate students in more communicative purposes and in a more practical usage of the language should also be based in providing them positive experiences they will implement in the future as language teachers. Amateur teachers usually teach the same way they were taught the moment they start to have their first experiences in the area of teaching. They, of course, might implement these kinds of evaluation activities as well, for it was the form they were evaluated as students. It can be a reality if new methods of testing are implemented giving a special concern to evaluating both students' knowledge and performance of the language.

# APPENDIXES

## Appendix 1. In-depth Interview format

### Q U E S T I O N N A I R E

**Description:**

The following questionnaire is designed to carry out interviews to teachers of the FLD in order to find out their opinions, difficulties and criteria they have at the moment they design and apply evaluation instruments in their respective courses.

1. How do you evaluate to your students?

2. What type of tests do apply to your students?

3. How do you design the tests you apply to your students?

4. In your opinion, what is the difference between testing and evaluating?

5. How do you associate the objectives of the course and the instruction you teach to the tests you apply in your courses?

6. How do you measure students' oral proficiency in your courses?

7. How do you measure students' writing proficiency in your courses?

8. How do you make students apply in tests high - performance abilities such as analysis, synthesis, evaluation, criticism, etc.?

9. Can you explain the way students accomplish the objectives of the course as well as your expectations as a teacher?

10. What are the reasons why students don't accomplish the objectives of the course and your expectations as a teacher?

11. How tests might affect the students' proficiency in the acquisition of a second language?

12. What is your opinion about the application of tests to measure students' proficiency in the acquisition of a second language?

**Appendix 2. In-depth Interview transcripts**

## INTERVIEW WITH LIC. JUAN CARLOS CRUZ

**Date: Wednesday, April 25th 2007**

1. **As a first question, tell us about your experience in the Foreign Language at University of El Salvador.** I have been here for three years and a half, at the Foreign Language Department, in the University of El Salvador.

2. **What subjects have you taught in these last three years?** I have taught composition for 3 times, Basic and Intermediate English, also Grammar and Pronunciation.

3. **How do you evaluate your students?** I have my own filosofy in evaluation, I use theoretical things. Whenever I evaluate my students, they really have to demonstrate what they are really learning. Tests go with that purpose. Some teachers just take tests from the book, which is not that good.

4. **How do you evaluate oral skills?** Concepts go more for competence. But in my case, I go more for performance, not for competence. Students have to demonstrate what they really can do with the language.

5. **How do you associate the objectives of the course and the instruction you teach to the tests you apply in your courses?** The tests I apply to my students are based on the way I have taught or pushed them during the course, I usually sit and take my time to prepare the tests to be applied to my students, because sometimes, we just take tests form the textbooks and that is not related to what we have taught in class. I evaluate them and give feedback. The objectives are directly related with the instruction.

6. **How do you measure your students?** First take into account the progress. Check objectives in the program, intense of performance. In writing, definitely take form skills students should develop. Lack proper use of grammar, etc.

7. **Why do you think students do not accomplish the objectives of the course?** Unfortunately some teachers do not push their students. There is also a lack of

supervision, teachers do whatever they want, for example, in composition, it is not that easy to grade, and teachers do not complicate themselves to evaluate students. Also, strikes interfere with the program. Some students do not care. Some students do not accomplish because of their attitudes, they do not take learning seriously, that is why most of them fail, so teachers have to push them.

8. **How do you make students apply high performance abilities, such as analysis, synthesis, evaluation and criticism?** I make that happen, by teaching classes not easy, challenging activities, etc.

9. **Do you think tests are necessary?** Yes, they are. Because if you do not give a number, students are not going to be promoted. Some teachers do not even return the papers with the appropriate feedback, which is necessary, so students can improve their learning, but this is because we have not been trained in how to evaluate our students.

10. **Don you think tests can affect students?** Tests can affect students when they are not designed with the objectives of the course, because that can make students fail, and they might feel frustrated.

11. **What is your opinion about the application of test to measure students proficiency in the acquisition of a second language in the Foreign Language Department?** I think, they are not appropriate, we have not been trained to evaluate or to design test. This has to be improved, because teachers do whatever they want. There should be supervision.

## INTERVIEW WITH LIC. RICARDO GARAY

**Date: Wednesday, April 25th 2007**

1. **Can you tell us about your experience here in the University of El Salvador, at the foreign language Department?** I have been here since 1990. Most of the courses I have taught are Readings and Conversation, Grammar, Composition I and II.

2. **What are the subjects you have taught in the last three years?** Mostly, in the area of abilities, and Composition I and II.

3. **How do you evaluate your students?** I evaluate them almost every day in class, I make summative evaluation. I evaluate them every day, to make sure they are learning.

4. **How do you design tests?** Some people say "do not complicate" just take book tests … some teachers just modify them. They do not have an idea how to design tests. You cannot analyze the results. We sometimes do not pay attention to what we are doing.

5. **How do you associate the objectives of the course and the instruction you teach to the tests you apply in your courses?** Students do not know about the topic, I give them like six topics, check it if it has topic sentence, what is the purpose of the author. In speaking, I design tests according to what they are supposed to do at the proper level. Check sentences structures, etc. they also work as a team and they are given the same exam to evaluate themselves. In Composition, check writing abilities, where they have to write a paragraph and at the end, they have to write an essay. Taking into account aspects like vocabulary, structure, mechanics, etc.

6. **What do you do when students do not accomplish the objectives of the course?** Sometimes students misbehave in classes, so teachers have to push tem to reach the objectives, we should be getting information so you can preview things and act to improve those situations. As a teacher, you have the obligation to know at least their names. Give them some activities that make students to participate. Students sometimes are not pushed as the way they have to be pushed; some students do not bring the material they are supposed to use in class.

7. **How do you make students apply high performance abilities, such as analysis, synthesis, evaluation, and criticism?** Talking about internet, for example. Students work in pairs, I give them topics for each couple or group, so they have to talk about those topics. I make them to analyze the ideas and put them together, so then, they share with other couples or groups.

8. **Do you think tests affect students' proficiency?** Sometimes it affects them positively, and sometimes it affects them negatively. It is like a motivation, sometimes, depending on the students. Traditional way is not bad, depending on what teachers want to get.

9. **What is your opinion about the application of tests to measure students' proficiency in the acquisition of a second language here in the Foreign Language Department?** Tests give you information about what you have proposed at the beginning of a course. Objectives always have to be connected, some teachers do not link the objectives of the course to the tests they apply to their students.

## INTERVIEW WITH LIC. GRACE GOMEZ

**Date: Wednesday, April 25th 2007**

1. **Can you tell us about your experience here in the University of El Salvador, at the foreign Language Department?** I have been here since 1990, like about 17 years.

2. **What are the subjects you have taught in these last three years?** I have been teaching in profesorado. Sometimes, English, Reading and Vocabulary.

3. **How do you evaluate your students?** I give them written tests, quizzes, depending on different types of activities to achieve. Including different types of items. In intermediate or Advance English, with reading books, to evaluate aspects.

4. **How do you design the tests you apply to your students during the course?** In written tests, for example, we have many groups, we are about nine teachers and for the mid-term exams, each group decide different parts to be included in the test, like listening, writing, reading, etc. We have a meeting, where we decide what kind of items and how many items will be included in the test, but sometimes, teachers do not go to the meeting to decide how to design the tests. We all together prepare some functions based on what is taught in class.

5. **How do you make students to accomplish the objectives of the course?** Show instruments, so students know what to improve. It is better to design instruments. In basic level, I ask students to write in sentences about family, and then they bring sentences talking about the topic. I check the sentences, and give them back with the corrections. And at the end of the course, I ask them to bring all the papers with the corrections in a folder, to see how they have improved their learning.

6. **How do you associate the objectives of the course and the instruction you teach to the tests you apply in your courses?** That is very important. You have to take into account the objectives, so you cannot get lost in what you want to reach. It is also necessary, to have a clear idea about what students have to do. You have to check constantly the objectives for each skill.

7. **What do you think are the reasons why students do not accomplish the objectives of the course?** The reasons why students sometimes do not accomplish the objectives, is because sometimes the group is too big, and it is impossible to work with a big group, sometimes we have about forty five students in a group, and some students are beginners, and some others already have knowledge of the language, do that is why, it is difficult to work with this students. What I do is, I have an assistant that helps m e with the people that have more difficulties. Another reason why students do not accomplish the objectives, is because, sometimes they need to be pushed, if you do not ask them to do something, they do not do it, some students do not even have the material to be used in class, and they do not do anything if they do not have the material, or the book.

8. **Do you think tests affect students' proficiency?** It can affect positively or negatively, depending on the students, sometimes on the results. Some students only when they know they will be evaluated, they worry.

9. **What is your opinion about the application of tests to measure students' proficiency in the acquisition of a language in the Foreign Language Department?** I think, we have to change the way to evaluate. The things we carry out in the classroom, what we do in order to have students to learn. What to do to motivate students to keep students attention, in the class. Students progress to see if students progress or not, if not, what to do with that. My job is to have students to learn.

## INTERVIEW WITH LIC. MAURICIO CONTRERAS

**Date: Wednesday, April 28th 2007**

1. **Can you tell us about your experience in the Foreign Language Department?** I have been working here since august 2004. I worked in Santa Ana for about 15 years. Now here, I have 3 years. You can work with different kinds of students, you have to be with new methodologies for teachers, it has become easier because of the internet.

2. **What subjects have you taught in the last three years?** Teaching Practice I and II, Composition I and II, Intermediate English I, Oral Expression at profesorado, Compared Grammar English and Spanish.

3. **How do you evaluate your students?** Different criterias for evaluating students. We call the basic area. Special organization for evaluating. There is a group of meeting and we evaluate different criteria, like Grammar, comprehension, understanding students, whether or not, students understand what they are asked to do, and pronunciation, one of the most important things in oral tests. Grammar fact is different, because you take into account spelling, grammar, vocabulary, and some extent, some higher level, and coherence in writing.

4. **How do you design tests?** Talking about oral tests, we have actually the book we use right now in Basic English; we have a set of questions to ask students, using different pictures. We use some open questions. Then we evaluate pronunciation, fluency, comprehension. Writing test is different; it is divided in two sections, different from two years ago. It includes listening and reading. In the first part, we have students to listen to some materials they have to answer simple questions. On the second part, we have writing and grammar. In writing, they have to write some paragraphs and correct some sentences. In grammar part, we include all grammar clauses on syllables.

5. **What are the reasons why students do not accomplish the objectives of the course and the expectations as a teacher?** A lot of companies have come to El Salvador, a lot of students that I have in my class, work for Dell, Sykes, or Teleperformance. Some of my students work full time, and they don't come to the university, that is one of the most common reasons why students don't accomplish the objectives or my expectations as a teacher.

6. **How tests affect students?** When students fail, they do not want to continue studying, but it can also affect positively, because if they failed, they try to do their best for the next evaluation.

7. **Tell us your general opinion about the application of tests in the Foreign Language Department.** In my opinion, students always try to reach a higher level, and in order to promote students, we have to assign numbers. So, tests are necessary.

## INTERVIEW WITH LIC. ALEX LANDAVERDE

**Date: Wednesday, April 28th 2007**

**Experience in the language department**
3 years as a University teacher and worked for about 5 years in the department's CENIUES. During the Tenure, I've taught varied subjects: Basic English, Advanced Grammar, Composition, Grammar, Readings and Conversations, and some English courses.

**Evaluation**
It's based on the program. The program determines the type of evaluation. For example, English Courses include partial exams and a mid-term exam that includes the assessment of the 4 macro-skills. These are usually complemented with quizzes and tasks throughout the course. Usually, most of the tasks are the tests, the oral and the written exams.

Design & Type
Good question. When we work with teams (and this is not in all subjects) we assign an area of examination, one skill, base it on the contents and select the right material and level, and then design the test. On grammar, for example I'd find some material from internet or books like the students use, and base it from there. It's what we do, you get a nice hybrid exam that works well on teams on subjects like advanced English, making it a very positive experience and material for grading students.

Text book tests
The thing about them is you always have the power of decision; no teacher can escape from the task of doing a good test. In my experience, you get more rewards. I believe that one shouldn't "demonize" textbook tests; after all, they *were* made by PHD's . The tests reflect the parts of the book that they are testing. I do believe those tests could be the base and can be the reach of better tests.

The cons of textbook testing
The biggest problem of textbook tests is that they're too easy. Any institution acknowledges us to test at a college level. So these should be 3 times higher. Besides that, they're as much as we can do. I discussed once on how scientists make tests, since it's one of the hardest things to accomplish in education. Regardless of the evaluation system, on has to ask whether to base it on product or process. In Sweden, they have a very good system based not in formal evaluation but based just in the process. Students are never given any grades, but this is very telling.

But what really taxes on higher language acquisition teaching is that the thing that has to be evaluated the most is now the notional function accomplishment, what students do at the end of instruction. To complete the criteria, the active guidelines, the American council at TOEFL and the European Education framework has very good guidelines on what to evaluate. But when it comes to teaching, people tend to make it simple, and sometimes the guidelines slip. Some books, very new course books, don't come with tests; they come with

test-design devices. Asks for what elements you have and how to design it. It's still a challenge though, for instance, when we evaluate critical thinking or problem-solving strategies, which can be used in language learning, out o which we could make some critics to our own system

Differences between evaluation and testing.

Testing is a circle inside assessment, it gives a grade; and assessment is inside Evaluation; both testing and assessment are smaller processes, evaluation is the most demanding and it involves complex procedures

Association content-evaluation/testing
As I said, it's based on the program. Like Profesorado students' English Composition and Oral Expressions, I take only one course for which I have communicative objectives: grammar to master, vocabulary and expressions, goals based and approached to course linguistics; the new approach, called Touchstone has a purpose to have the students sound natural, not bookish. We used to learn by the book or by memorizing dialogs, now it includes a conversational approach.
This is what my exam will include: Grammar, the one they've studied, at a bit higher level; Speaking, to use it naturally. Reading Comprehension, to read a paragraph and tell what it's about and give an opinion about it. Composition is a bit different, they have to write a coherent paragraph; there's no other evaluation other than this in this class, Because this is something no other subject asks: to write a paragraph; usually you're given time to do research and then they get to make a topic sentence, good development of ideas and a coherent conclusion.

On areas of abilities' oral proficiency evaluation and the latter's frequency
Cambridge is using "Teacher Knowledge Tests". On assessment, standardized tests are objective-oriented, while oral and composition assessments are subjective exams. This means that if 5 people are being tested orally, you'll get 5 different answers; oral exams can't be evaluated objectively. One can say pronunciation, clarity, fluency, vocabulary and grammar can be graded, though, but even if no 5 people will give the same response, all one must do is listen to them for about 5 minutes to place them on a level; the verbs, the expressions, the words one uses can allow one to place them on a level. But when we refer to specific instruments, such as an interview or proficiency test, the evaluators must ponder and discuss.

Students' competence versus students' performance
You teach as you were taught. Task-based learning is specialized on what you will be doing. I don't believe in one to one interviews; I prefer presentations and debates. People can display great level but still not sound natural. There are many variables like the number of students and teacher work load. We could do better.

On composition evaluation

We never really practice writing, even in Spanish. People want to write just like they speak or even Spanish-like it but it's difficult. Spanish writes big paragraphs but it's not the same in English; Spanish will use synonyms to richen the writing while English will keep using the same word so as to not confuse readers. We evaluate topic sentence, development, supporting sentences, conclusion, vocabulary, mechanics and punctuation. We as teachers have limitations, too. I'm no writer so I can't really tell my students that they're wrong.

High Abilities in composition
It's a challenge. And while it's easy to apply them, these higher-thinking skills are based on critical thinking, on the reaction to things, in being used to analyzing or fragmenting things, to summarize them and we're not analytical, we're not being even critical. We accept things. We're critical only in a passive-aggressive way. We encourage students to be critical.
Some Composition students wrote very good reactions on news articles and were asked to attach news printout as proof of reading it.

Opinion on not fulfilling the objectives
Most students aren't suited for this. There are cultural barriers. They want results without working for them. Sometimes it's about the character. Others were different, they had levels and skill and I cared very much for them. I don't really want to believe in the criteria, I don't really apply it; the criteria allows people to either pass or not, but students may have issues and can't really be put into the criteria; in other places, the thing is exactly the opposite. I just care for them to learn. But there are too many variables

How do tests affect?
We give them a lot of importance. We should get rid of "studying for tests" and start "studying for learning". And that's the disadvantage. We become immediatists. Readings and Conversation was a great challenge. I'm taking a test for a scholarship and it's of a very high level, but I'm enjoying the fact that at the time I prepare for it, I'm studying for learning.

Application on tests measuring proficiency
They're important, and we need to track and monitor them. But tests aren't fully reliable; in other words, don't get rid of tests but keep the opinion reflected on the results. One should have other measuring tools and base one's opinions on the student's capacity. Daniel Dock's "emotional intelligence" – related research portrayed great academically students ending up having disastrous personal lives, while others with social and leadership skills got top-executive positions. People graduate and the best among them can't be guarantied success.

## INTERVIEW WITH MSD. DORIS PORTILLO

**Date: Wednesday, April 29th 2007**

**What's your experience in the Foreign Language Department as a teacher?**

I've been working here since 1993. I have been teaching all the courses from the basic area, like Basic, Intermedia and Advanced. I've also taught subjects such as Didactis I Didactics II for both careers, Profesorado y Licenciatura.

**In the last 3 years (2004, 2005 and 2006) What subjects have you taught?**

I have taught Advanced I, Advanced II, Didactics of English I, II and III. Those have been basically the subjects I have taught

**In subjects concerning the area of abilities, such Basic, Intermedia and Advanced, how do you evaluate your students in these subjects?**

For example, now in Advanced we take into account what we call Tab based on Speaking activities and written activities. We have two for the written area and one for the spoken area. We also have another area, which is for quizzes. We have a written quiz every two units, and we have the mid term, oral and written exam, and the final exam too.

**How do you design these tests?**

First, we take into account the objectives stated in the program, that's what we do. There's one person in charge of designing for example quiz 1 and quiz 2, and the others. And this person has to take a look at the objectives stated on the program, and of course the content that has been decided. Based on those things, the quiz is designed.

**Do you usually work as a team when you design Mid-term Exams and Final Exams?**

For Mid – Term and Final exams yes, but not for the quizzes because every quiz is designed by one person or probably two. In this case, for example, this semester, I was in charge of designing quiz 1 and also Rolando Guzman worked with me. The second quiz was designed by Miguel Carranza. But we have the next evaluation, wich is the Mid-term, that will be designed by everybody . In my case, I have been assigned the written part.

**Now in your personal opinion as a teacher here in Foreign Language Department, what is the difference between evaluation and testing?**

Testing for me is about the instruments you use. In our case the written exams. And Evaluation is the process of judging, the quality of the product you get from your students

**In subjects such as Advanced, Intermediate, Basic, how do you measure students' oral proficiency?**

I can give you my opinion based on the advanced level because this is what I have been teaching most of the time. For the speaking skill, we have, as I said before, the oral presentation, which is part the Pad section, and then we have the oral mid term and the final oral exam. And the oral is basically based on the functions and the objectives stated on the program. So last year for example, we interviewed the students. There were two teachers interviewing every student in order to have a more objective discussion.

**Our group has gathered some samples of exams which have been carried out here in the Foreing Language Department. We have noticed that most of the teachers usually test students' competences rather than students' proficiency in the language or Performance. Do you think that teachers should apply more oral exams based specifically in students' performance in a real situation?**

I think it would be better if students were given more oral exams because we have the chance to have a complete scenario of how proficiency they are in the language, so I would agree if we decided to give students more orall exams. But not only oral exams but also written too. Through speaking and written pads the students have the chance to learn every piece of the language. We are not focused on specific pieces of language, or a specific grammar part. But if you have students write something, you'll be checking punctuation, the use of connectors, organization, coherence, and all those things. The same happen when you test speaking.

**How do you make students use in these tasks high performance abilities, such as analysis, evaluation, synthesis, critical thinking in these specific abilities, writing and speaking?**

For example, we have two pads we have to do with writing. For one, students have to write a book report. In that book report, students have to be critical. At every book report, there is a format to be pound but we also a component in which students write their own opinion or critic. In speaking activities, concerning the presentation, we have that performance too. Not easy topic but complex topics such as… ahh… let me remember some of them such as immigration causes and consequences. So in a way there's always a section in which students will provide his own opinion.

**Based on your opinion and your experience as a teacher, what are the reasons some students don't accomplish the objectives of the course and your expectations as a teacher?**

There might be many factors. The teacher might not be doing what he is expecting to do, the students themselves. The teacher has to see how much time students devote to learning. We need to see if students are investing in their own learning because for me I always believe that Teaching and Learning is a shared responsibility. Nobody can tell me if the student fail, it's the teacher's responsibility, or if the students succeed is the teacher's responsibility. It's half and a half. There are other external factors such as financial problems, economical problems, emotional problems

**How do tests affect students' proficiency in learning a second Language?**

In my case, I don't think it affects students' proficiency. Probably what I can say is that depending on the instruments you use, we are allowed to see objectively how much students know.

**We've interviewed other teachers that tell us that students study or get prepared to succeed in a test rather than learning actually**

But the instrument we use has something to do with the development of the students' proficiency. But that would be a different issue.

**Do you think it is sometimes necessary to push students to go farther than preparing students for a test?**

As teachers, we have the responsibility to encourage them not to push them. Because pushing my students can be something they don't want. If they are here at the university is because they decided to do so. If they don't do what they have to do as students, they will face the consequences. But, it would be fine if I invite my students to invest in their own learning outside the classroom.

**As a last question, what's your opinion about the application of tests to measure students' proficiency in the acquisition of a second language?**

Depending in how well the test is designed, we can limit what the student can produce. It doesn't matter how good your instrument is.

## GLOSARY.

**Assessment:**

A systematic collection of qualitative information used to cast a judgement.

**Diagnostic Test:**

A test implemented to find out what students know and don't know about a specific area of studies.

**DFL.**

Department of Foreign Languages.

**Evaluate:**

to express a judgement towards a specific behavior.

to express a value to the skills, abilities or knowledge the student has developed and can show proof of it.

**Formative Evaluation**

Qualitative information inducted from an observer regarding hard-to-measure qualities from the subject.

**Measurement:**

The collection of quantitative data to determine the degree of whatever is being measured.

**Placement Test:**

A test implemented to decide what level of class the learners should go into.

**Proficiency Tests:**

A test implemented to see how good somebody is at language. The contents of this test are not based on a course or syllabus that the learner has followed.

**Progress Test:**

A test implemented to find out how well learners have learned something.

**Selection items:**

items in which the answer or answers have been already provided and it's up to the subject to select the correct answer.

**Summative evaluation:**

Quantitative (measurable, and therefore, addable) information that compounds proof of level of mastery or non-mastery of evaluated contents.

**Supply items:**

Items in which a complete answer is expected from the subject.

**Test:**

A set of items used to generate data and provide quantitative proof that imply an evaluation judgement or grading.

**Testing:**

The process of developing and implementing a means through which one is able to evaluate, measure or grade a knowledge or behavior.

**Test items:**

short instruments used to develop an evaluation.

## BIBLIOGRAPHY


Clark, John L.D. (1972) *Foreign Language and Testing: Theory and Practice*. Philadelphia: Center for curriculum development.

Heaton, J.B. (1990) *Classroom Testing,* London: Longman

Hughes, A. (1989) *Testing for Language Teachers,* Cambridge: Cambridge University Press.

Madsen, Harold S. (1987); Techniques in Testing. London: Oxford University Press.

Sharon A. Shrock and William C.C. Coscarelli (1996) *Criterion – Referenced Test Development,* The United States: ISPI (International Society for Performance Improvement)

Ur, Penny. (1996) *A course in language teaching: Practice and Theory,* Cambridge (England); New York: Cambridge University Press

Weir, C. (1990) *Communicative Language Testing,* Hemel Hempstead: Prentice Hall International.