

**UNIVERSITY OF EL SALVADOR
SCHOOL OF ARTS AND SOCIAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES**



**THE METHODOLOGY AND INSTRUMENTS USED BY TEACHERS TO ASSESS
SPEAKING SKILLS OF STUDENTS IN THE INTERMEDIATE ENGLISH I
COURSES OF THE FOREIGN LANGUAGES DEPARTMENT AT THE
UNIVERSITY OF EL SALVADOR DURING SEMESTER II-2014**

PRESENTED BY:

MENA ORELLANA, ANDREA YAMILETH
MONGE NAVAS, JESSICA DENISSE

MO07012
MN07005

FINAL RESEARCH REPORT IN ORDER TO OBTAIN THE DEGREE OF
BACHELOR OF ARTS IN ENGLISH WITH EMPHASIS IN TEACHING

MIGUEL ANGEL CARRANZA CAMPOS
RESEARCH ADVISOR

MANUEL ALEXANDER LANDAVERDE
COORDINATOR OF GRADUATION PROCESSES

NOVEMBER 27TH / 2015

MAIN UNIVERSITY CAMPUS, SAN SALVADOR, EL SALVADOR

AUTHORITIES OF THE UNIVERSITY OF EL SALVADOR

Mario Nieto Lovo, Engineer.
PRESIDENT

Ana María Glower de Alvarado, M.A.
ACADEMIC VICE-PRESIDENT

Oscar Noé Navarrete; M.A.
ADMINISTRATIVE VICE-PRESIDENT

Francisco Cruz Letona, B.A.
ATTORNEY GENERAL

Ana Leticia Zavaleta de Amaya, Doctor.
SECRETARY GENERAL

AUTHORITIES OF THE SCHOOL OF ARTS AND SOCIAL SCIENCES

José Raymundo Calderón Moran, B.A.
DEAN

Norma Cecilia Blandón de Castro, M.A.
VICE-DEAN

Alfonso Mejia Rosales, M.A.
SECRETARY

AUTHORITIES OF THE DEPARTMENT OF FOREIGN LANGUAGES

José Ricardo Gamero Ortiz, M.A.
HEAD OF DEPARTMENT

Manuel Alexander Landaverde
COORDINATOR OF GRADUATION PROCESSES

Miguel Ángel Carranza Campos, M.A.
RESEARCH ADVISOR

EVALUATING COMMITTEE

Miguel Ángel Carranza Campos, M.A.
Ana Grace Gómez Alegría, M.A.
Mauricio Salvador Contreras

TABLE OF CONTENTS

Table of contents

Introduction	1
Chapter I	3
1. Statement of the problem	3
1.1 Description of the problem	3
1.2 Delimitation of the problem	4
1.3 Justification	4
1.4 Objectives	5
1.4.1 General objective	5
1.4.2 Specific objectives	5
1.5 Research questions	6
1.5.1 Research general question	6
1.5.2 Research specific questions	6
Chapter II	7
2. Historical framework	7
Chapter III	8
3. Theoretical framework	8
3.1 Assessment of oral performance	8
3.2 Assessment of oral performance	10
3.3 Kinds of oral assessment	13
Chapter IV	17

4. Hypothesis	17
4.1 Research hypotheses	17
4.2 Conceptual definition of variables	17
4.3 Operational definition of variables	18
Chapter V	20
5. Research design	20
Chapter VI	21
6. Population and sample	21
6.1 Population	21
6.2 Sample	21
Chapter VII	22
7. Analysis of data	22
7.1 Data gathering process	22
7.2 Data analysis	24
7.2.1 Analysis of survey results	24
7.2.2 Analysis of classroom observation results	33
7.2.3 Analysis of teachers' interviews	37
7.2.4 Triangulation of instruments results	44
Chapter VIII	48
8. Conclusions, recommendations and limitations	48
8.1 Conclusions	48
8.2 Recommendations	51
8.3 Limitations	53

Appendixes	54
Appendix A: students' survey	
Appendix B: observation sheet	
Appendix C: teachers' interview	
Bibliography	64

INTRODUCTION

Assessment is an important part of any course at the Foreign Languages Department of the University of El Salvador. Assessment is defined as the process of collecting information about learners' performance in order to make judgments about their learning (Spratt, Pulverness, & Williams, 2006). That is why students are subjected to different evaluation processes in the different areas and skills like writing, reading, listening, or speaking. Since speaking is one of the two skills in which production of the language is expected, it requires meticulous and careful assessment, therefore, a variety of activities are carried out and a different number of instruments are used in order to accurately measure a student's speaking skills. These activities should be appropriate to the level and they should meet different criteria. And not only the activities are important, but also the way they are assessed: are they assessed fairly and reliably? Are these activities being objectively scored? These are just only some of the questions that arise when thinking about students' speaking abilities and how exactly they are being assessed.

The general purpose of this research is to investigate what the methodology and instruments used at the time of assessing students' speaking skills are. In order to do this, a descriptive study will be carried out. A descriptive study is defined as "a study that searches to specify important properties of people, groups, communities or any phenomenon that is submitted to analysis" (Dankhe, 1986). Since this project aims at specifying important properties of the assessment of speaking abilities a descriptive study is best suited.

The population of this research will be the students of the Intermediate English I courses of the Foreign Languages Department at the University of El Salvador in the semester II-2014. Both quantitative and qualitative data will be obtained through this research by using three different techniques. These techniques are: the survey, the interview, and the observation. The survey will be administered to students of the previously mentioned courses, the interviews will be conducted with some of the teachers of said courses, and finally the observation checklist will be used when observing classes of the same courses.

CHAPTER I

1. STATEMENT OF THE PROBLEM

1.1 DESCRIPTION OF THE PROBLEM

In the Intermediate English I courses at the Foreign Languages Department, the students' speaking level of the language is expected to be developed in areas such as: good pronunciation, good use of structures, proficiency and accuracy appropriate to the level, among others. That is why students are subjected to different assessment processes in order to check their progress regarding the speaking performance required and at the end of a course being promoted to the next corresponding level.

Thanks to the experience of the researchers as both students and teacher's assistants in the English courses at this Department, it has been perceived at different times how students in advanced courses have difficulties or problems in their production of the English language that should have been minimized or corrected in previous courses. Basic problems of pronunciation and grammar are some examples of said difficulties. The question of how were students' being promoted to the following level when they presented these basic difficulties arose and with it other questions like how were speaking abilities assessed and what aspects were involved in it appeared too, and this situation was selected as a research topic.

At that moment, it was understood that a variety of activities need to be carried out in order to measure accurately a student's speaking expertise, and that there are different elements concerning the assessment of speaking than just the activities used.

1.2 DELIMITATION OF THE PROBLEM

The topic of this research is the methodology and instruments used by teachers to assess speaking skills of students in the Intermediate English I courses of the Foreign Languages Department at the University of El Salvador during semester II-2014. The subjects of the study will be the students of the Intermediate English I courses at the Foreign Languages Department from the University of El Salvador. The research will be carried out during the second semester of 2014, which covers the months of August to November of the same year.

1.3 JUSTIFICATION

This research was made in order to find out what the exact activities and instruments used for evaluating the speaking skills of students in the Intermediate English I courses are. This research is considered valuable since oral production is an important part of the English classroom since it is the tool of communication among students and between students and teachers, and it will be just one of the abilities that will give students great opportunities in the work field, particularly in teaching, in the future.

The main purpose is that at the end of this research, after finding out what could be the exact issues with the evaluation of oral performance, some suggestions and ideas can be made that might help improve this situation. It is expected that the possible suggestions can benefit all those involved: teachers and of course students, whose oral production is the main focus of the research.

1.4 OBJECTIVES

1.4.1 General objective

- To research how students' speaking skills in the Intermediate English I courses is assessed by the teachers of the Foreign Languages Department.

1.4.2 Specific objectives

1. To identify the methodology used for assessing students' speaking skills in the Intermediate English I courses at the Foreign Languages Department of the University of El Salvador.
2. To identify the instruments used for assessing students' speaking skills in the Intermediate English I courses at the Foreign Languages Department of the University of El Salvador.
3. To verify the application of the principles of language assessment at the time of assessing students' speaking skills in the Intermediate English I courses at the Foreign Languages Department of the University of El Salvador.

1.5 RESEARCH QUESTIONS

1.5.1 Research General Question:

- How are students' speaking skills being assessed by teachers in the Intermediate English I courses of the Foreign Languages Department?

1.5.2 Research Specific Questions:

1. What is the methodology used by teachers of the Foreign Languages Department of the University of El Salvador to assess students' speaking skills in the Intermediate English I courses?
2. What are the instruments used by teachers of the Foreign Languages Department of the University of El Salvador to assess students' speaking skills in the Intermediate English I courses?
3. Which principles of language assessment are applied by teachers of the Foreign Languages Department of the University of El at the time of assessing students' speaking skills in the Intermediate English I courses?

CHAPTER II

2. HISTORICAL FRAMEWORK

The process of assessing students' oral performance at the Foreign Languages Department has changed during the course of the last fifteen years. The activities and tools that were used in the past are not the same used these days.

A set of interviews were conducted to some teachers of the Foreign Languages Department to ask them about the changes that have occurred in this very important process. One of those interviewed teachers was Ricardo Cabrera Martinez. He explains that in the very early days of the Foreign Languages Department, the interview was the most used format to assess a student's oral performance. At around 1995, activities like oral tasks and oral presentations began to be implemented at the Department. While at that time there were some tools that were used to assess students, they were not as advanced and complete as they are now. The use of technology and a different amount of workshops have helped to improve the process.

These days, a variety of activities are used to evaluate students' oral performance, including the now so famous interview. And more advanced tools like very detailed rubrics are now used at the time of evaluating oral performance at the department.

CHAPTER III

3. THEORETICAL FRAMEWORK

3.1 ASSESSMENT OF ORAL PERFORMANCE

Assessment is defined as the process of collecting information about learners' performance in order to make judgments about their learning (Spratt, Pulverness, & Williams, 2006). During a course, learners are assessed for different purposes, and different techniques and methods can be used to do so.

There are two ways in which learners can be assessed: informally and formally. Among informal assessment procedures we find in-class observation, monitoring of activities, and homework assignments, and for formal assessment we find examinations and tests (Spratt et al., 2005). The use of informal or formal assessment procedures will depend on the teacher and the aspects that need to be assessed.

In order to assess learners' abilities and skills, whether they are reading, writing, listening, or speaking, there are five important principles that need to be considered in all kinds of language assessment, whether it is formal or informal (Brown & Abeywickrama, 2010):

- **Validity**

First, we need to define what validity is. According to Bailey (1998) validity is the criterion that tells us that a test is exactly measuring what it should be measuring. For

example, if we are assessing a basic level learner, we cannot expect the learner to perform at an intermediate or advanced level because our assessment would not have validity. Validity is described as one of the most complex and important criterion in order to prove if a test is effective or not (Brown & Abeywickrama, 2010).

- **Reliability**

Reliability is the criterion that indicates that a learners' performance should be measured in a consistent way and that the test should also be consistent and dependable (Brown, 2001). If the same test is administered on different days and to different students, the test should offer similar results. Reliability involves not only the test itself but also other factors like the student, the test administration, and the scoring process (Brown & Abeywickrama, 2010).

- **Practicality**

Practicality involves the different aspects related to creating, administering and scoring a test (Brown & Abeywickrama, 2010). These aspects include the cost, the time it takes to create and administer a test, how easy it is to score it and how easy it is to report the results. If a test is practical, it will have an appropriate amount of time for students to complete, instructions will be easy to explain to students even if they are a lot and there is only one

- **Authenticity**

The concept of authenticity is very difficult to define (Brown & Abeywickrama, 2010), but in essence authenticity involves the use of "real world" elements in testing. These real world elements can include real-life contexts or situations that a student can recognize and

relate to, natural language, meaningful and interesting topics, task items that have a relationship and are not isolated, and real-world tasks.

- **Washback**

According to Brown & Abeywickrama (2010), washback is “The effect of testing in teaching and learning”. It is the positive or negative effect that testing can have in a student. For example, if a student’s performance is appropriately assessed, whether the student did good or bad, he or she will feel encouraged to continue working well or to improve. When the washback is beneficial, it will positively influence how teachers teach and how students learn. Giving feedback, making comments and not just giving a grade, praising students when necessary and even making suggestions are part of washback.

3.2 ASSESSING ORAL PERFORMANCE

We need to start by saying that assessing oral performance is a very subjective process (Harmer, 2001). This kind of assessment depends on the perceptions of the tester, so what might seem reasonable for one tester can be unreasonable to another. This is where we have to bring up again the criteria of validity and reliability, but now focused on the assessment of oral abilities.

As a solution to the problem of subjectivity in the assessment of oral performance, Harmer (2001) presents three ideas that can help to minimize it:

- **Training those who will evaluate students**

Before carrying out an oral evaluation, testers whether they have experience or not, should prepare themselves by knowing what they have to do: how to assess using rubrics, how to score oral performance, knowing the contents that will be evaluated, knowing how to carry out the types of tasks they will use, and in the case of interviews, knowing how to interact with learners and how to make the interviews flow naturally.

- **Having more than one tester**

For more reliable results, having another tester present at the time of evaluating oral performance is an idea that needs to be considered. For example, it is difficult conducting an interview and keeping track of a learner's performance at the same time, and also, having only one person testing could be not as reliable as it is desired (Hughes, 2002), because attention can be focused too much either on scoring or on the performance itself. Students might feel uneasy if they are not paid enough attention at the time of speaking and they might feel the score is unfair. By having two testers, tasks can be divided: one tester focuses on interviewing and the other one on scoring, in that way, assessment can be more reliable.

- **Using appropriate assessment tools**

The use of rubrics or assessment scales can make the scoring of oral performance more valid and reliable because teachers will know beforehand what they will score and evaluate, and students will know what is expected of them. These rubrics should include a

list of different criteria and a detailed explanation of each of them. According to Hughes (2002), the criteria levels of performance that could be included in the assessment of oral performance are:

- **Accuracy:** It refers to “the use of correct forms of grammar, vocabulary, and pronunciation” (Spratt et al., 2005, p. 34).

- **Appropriacy:** It refers to the correct use of language according to context, content, and level.

- **Range:** It refers to the wide variety of languages used to express ideas.

- **Flexibility:** it refers to the ability of adapting to changes such as changes of topic and changes of direction in a conversation.

- **Size of contributions:** It refers to both lengths (time) and quality of language used. It should be considered that a contribution can be long, but can lack accuracy; and on the other hand, a contribution can be short but there is accuracy and complex language is used.

3.3 KINDS OF ORAL ASSESSMENT

According to Lazaraton (as cited in Celce-Murcia, 2001), there are two ways in which oral performance can be assessed. The first one is the evaluation of classroom performance, and the second one is the use of large scale oral examinations.

- **CLASSROOM PERFORMANCE**

It is important to mention that assessment of speaking should be a constant process, this according to the reliability principle described before, and having only two formal oral assessments per semester for example, could mean having a limited view of a student's oral abilities. That is why students' participation in class activities should be included in the assessment of speaking. Different authors suggest a different number of activities that could be used to develop speaking inside the classrooms.

Some of the activities suggested by Harmer (2001) include discussions, role plays, and conversations. *Discussions* are probably the most common type of activity used in the English classroom and they are not difficult to set up and evaluate. These can be done in pairs or groups, and it gives the teacher the chance to see how students interact with each other by using language. *Role plays* are another common activity inside the English classroom. It also involves interaction and the use of a variety of language. A benefit of this kind of activity is that it is fun and entertaining. *Conversations* are "the most fundamental form of oral communication" (Harmer, 2001, p. 108) and almost all ESL/ EFL students can benefit from it.

Some of the activities suggested by Brown (2001) are intensive, responsive, transactional, interpersonal, and extensive speaking performance activities. *Intensive* performance involves the use of any activities in which students practice certain forms of language whether they are grammatical or phonological. *Responsive* performance includes student-teacher interaction or questions and comments made by students. These replies are usually short and do not extend to dialogues. *Transactional* performance involves the use of meaningful conversations where participants exchange information. These conversations can be in pairs or groups. *Interpersonal* performance has to do with maintaining social relationships as opposed to the transactional performance in which the exchange of facts and information is the main purpose. These types of activities suggest the use of natural language. And extensive performance involves oral reports, summaries, and even short speeches.

And finally, Hughes (2002) suggests using activities like questions and requests for information, using pictures to elicit speaking, role plays, and discussions. In *questions and requests for information* students can share their points of view and opinions and do not involve the use of yes/no questions. The use of *pictures* can be used for eliciting descriptions and can involve narration too. And as mentioned before, *role plays* and *discussions* are also great activities to elicit speaking in the classrooms.

- **LARGE SCALE ORAL EXAMINATIONS**

They are important because students can be assessed in more detail and more information about their abilities can be obtained. Hughes (2002) lists three formats that can

be used for assessing oral performance: the interview, interaction with peers, and response to tape-recordings.

The *interview* is the most common format used to assess oral performance. This format has both its advantages and disadvantages. One advantage is that a lot of valuable information can be obtained and is suitable mostly for teachers because it is easier for them to score. One disadvantage is that students might feel intimidated by having to interact directly with a teacher; therefore results might not be what they expected.

Interaction with peers can be a discussion between two students or a debate. The advantage of this kind of activity is that students can feel more comfortable interacting with each other than interacting with a teacher, and it also has advantages for the teacher because two or more students can be evaluated simultaneously. In this case, more than one tester will be needed. The disadvantage is that one student's performance can be affected by the others because one of them might dominate the activity and not allow the other to demonstrate his or her abilities, especially if they are not at the same level.

Response to tape-recordings while useful they are inflexible because there is no way of following up a student's responses. Also, it lacks the interaction of the other two activities.

At the time of developing large scale oral examinations there are some elements that should be taken into consideration. First of all, enough time should be given to develop the activity and it should be as long as necessary (Hughes, 2002). Using less than 15 minutes for a large scale oral examination would not be enough to obtain the necessary information from

a student's performance, but providing around 30 minutes will give enough time to collect the necessary information to the tester. Second, the test should be planned carefully. In an interview for example, having only a general idea of the things that are going to be asked would be a mistake. The tester should have a pattern to follow. This could be preparing questions for the interview beforehand. Third, the use of a second tester is very helpful. As mentioned before having only one tester can cause difficulties at the time of assessing a student's performance and having two can make the task easier and more reliable.

CHAPTER IV

4. HYPOTHESES

4.1 RESEARCH HYPOTHESES

- **Hypothesis 1:**

Students' speaking in the Intermediate English I courses at the Foreign Languages Department of the University of El Salvador is deficiently assessed.

- **Hypothesis 2:**

There is a lack of variety in the activities used to assess the oral performance of Intermediate English I students of the Foreign Languages Department.

- **Hypothesis 3:**

Students' speaking might improve by receiving the appropriate feedback to correct current mistakes.

4.2 CONCEPTUAL DEFINITION OF VARIABLES

Hypothesis 1:

- Variable: Deficient assessment
 - *Deficient:* Lacking some element or characteristic.
 - *Assessment:* The process of collecting information about learners' performance in order to make judgments about their learning

Hypothesis 2:

- Variable: Lack of variety
 - *Lack*: Deficiency or absence of something needed, desirable, or customary.
 - *Variety*: A number of different types of things, especially ones in the same general category.

Hypothesis 3:

- Variable: Providing feedback
 - *Provide*: To give something wanted or needed to (someone or something) : to supply (someone or something) *with* something.
 - *Feedback*: The return of a portion of the output of a process or system to the input, especially when used to maintain performance or to control a system or process.

4.3 OPERATIONAL DEFINITION OF VARIABLES

Hypothesis 1:

- Variable: Deficient assessment
- Indicator: Low oral production
- Technique or instrument to gather data:
 - ✓ Survey: questionnaire
 - ✓ Interview: list of prepared questions
 - ✓ Observation: observation checklists

Hypothesis 2:

- Variable: Lack of variety
- Indicator: Limited use of activities
- Technique or instrument to gather data:
 - ✓ Survey: questionnaire
 - ✓ Interview: list of prepared questions
 - ✓ Observation: observation checklists

Hypothesis 3:

- Variable: Providing feedback
- Indicator: Students' speaking errors
- Technique or instrument to gather data:
 - ✓ Surveys: questionnaire
 - ✓ Interview: list of prepared questions
 - ✓ Observation: observation checklists

CHAPTER V

5. RESEARCH DESIGN

The design of this research project is non-experimental. This is because variables are not going to be manipulated and since the situations already exist, there is nothing to be provoked or created. The situation and the subjects are going to be studied in their natural context, and there will not be manipulation of either.

In this non-experimental research the design to be used is a transactional descriptive study, since the information needed will be collected at a specific moment and not throughout a long period of time. This type of study is the best suited for this research since the purpose is to study a situation (the assessment of oral performance) and get ideas of the state of said situation and describe it.

CHAPTER VI

6. POPULATION AND SAMPLE

6.1 POPULATION

For gathering the data and results, the target population of this research will be the students of the Foreign Languages Department of the University of El Salvador in the semester II-2014 from both the B.A. in Teaching English and the B.A. in Modern Languages. More specifically, the students of the Intermediate English I courses.

6.2 SAMPLE

The sampling method to be used is probabilistic sampling. With probabilistic sample, all the members of the population have the same probability of being chosen.

The researchers have chosen the students of the Intermediate English I courses as their sample. With eight different groups of these courses and with a capacity for thirty-five students in each group according to the Academic Administration's registration information, fifty percent of these courses will be the final sample meaning: four groups of Intermediate English I courses. The groups will be chosen at random and all students in each group will participate as the final sample.

CHAPTER VII

7. ANALYSIS OF DATA

7.1 DATA GATHERING PROCESS

The data were gathered during the final months of semester II-2014. After the sample was selected and after the appropriate permissions from the teachers of the different groups were received, the data collection process started. First, the techniques to be used were selected based on the specific information that wanted to be collected. It was decided that the survey, the interview, and the observation were going to be the techniques to be used, and that the objective of using these three techniques was to triangulate the results and obtain more reliable and valid information at the end of the research and also to collect the information taking into account the perspective of those involved in the problematic situation: the teachers and the students, and also the perspective of the research team to finally compare all perspectives. Then the instruments to be used were selected, created, and tested. These instruments were a questionnaire, an observation checklist, and a list of prepared questions,

First, the observation was carried out during different dates in the month of October 2014. Four classes of the four different selected groups were observed meaning that a total of thirty-two hours of observation were carried out, eight hours in each of the four groups to be more exact. For this process, an observation checklist was used as an instrument. For two hours in four different classes the research team observed the activities developed in the

classroom and the way they were assessed. Not only the observation checklists were used at that time but also notes were taken at that time on the different things being observed.

Second, in the month of November 2014, the survey was administered to students of the Intermediate English I courses. It was decided to administer the survey almost at the end of the semester because by that time, students had done different large scale oral evaluations and therefore could answer the questions concerning large scale oral examinations presented in the questionnaire, which was the instrument used for this purpose.

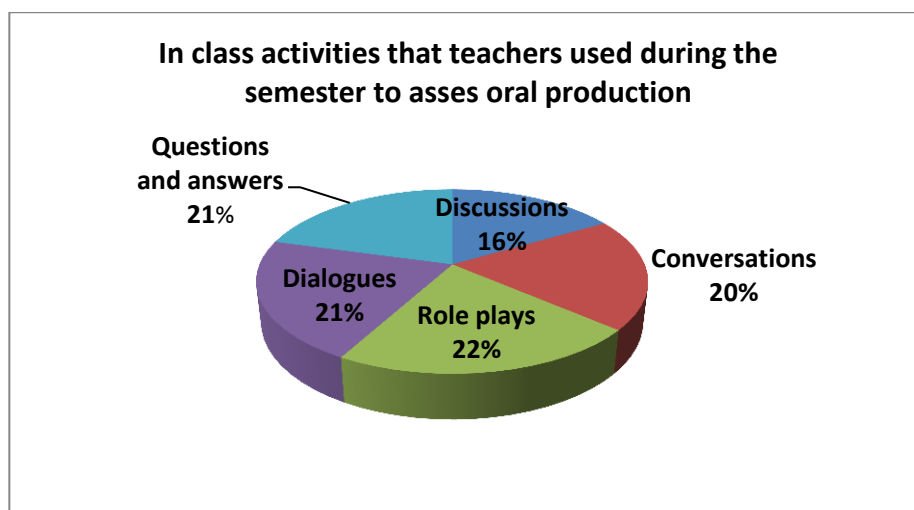
Finally, in the same month of November 2014, the interviews were carried out with teachers of the Intermediate English I courses. For this, a list of prepared questions was used as an instrument. The interviews were semi-structured interviews since teachers were given the opportunity to extend on their answers if necessary and not just answering questions providing short answers.

7.2 DATA ANALYSIS

7.2.1 ANALYSIS OF SURVEY RESULTS

SECTION I: IN CLASS ASSESSMENT

GRAPHIC 1:



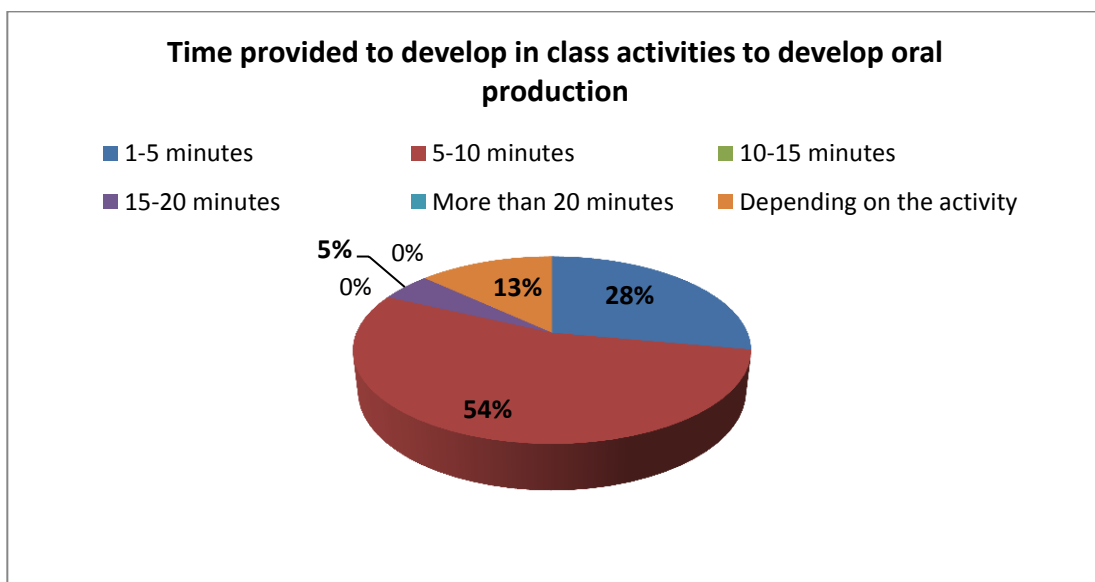
Graphic for survey question 2 from section 1

ANALYSIS:

When students were asked about the type of activities that were used in classes to develop oral production and communication (*survey question 1 from section 1*), 22% of students responded that the most common activity used in classes were role plays, which is the majority of students. The next impacting results belong to the 21% of the population who agreed that dialogues were a commonly used activity in class assessment, as well as questions and answers. Then, we have 20% of the responses which belong to conversations, and finally the remaining 16% of the responses belong to discussions, which is the last category in this question. As a matter of fact, in class activities were always promoted in the classroom to develop speaking. The use of this kind of activities is really important since the theory explains that oral performance can be assessed not only through large scale oral

examinations but also through the use of speaking activities like the ones mentioned before. While these are useful activities, more variety can be added. If the same activities are used over and over, the class can feel repetitive and even tedious for students, even resulting in lack of participation.

GRAPHIC 2:



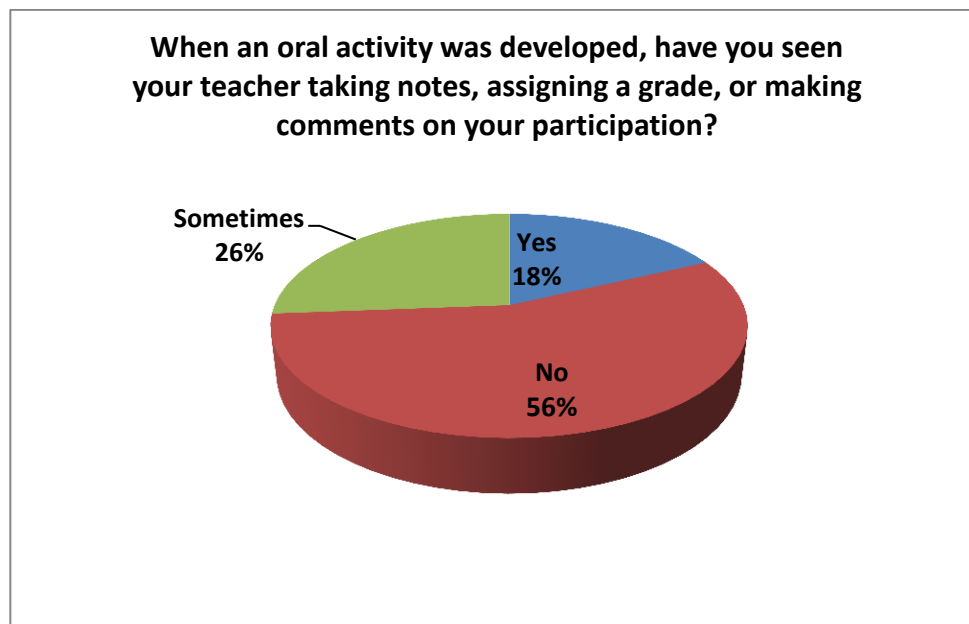
Graphic for survey question 4 from section 1

ANALYSIS:

Students were asked what was the estimated time provided to develop activities like dialogues, conversations, discussions, and role plays in a single class; their responses on one hand showed a 54% saying that the time provided was between five and ten minutes to practice. On the other hand, 13% of students' responses mentioned that it depended on the activity that was performed. 28% of the respondents answered "one to five minutes", and 5% answered that they were given fifteen to twenty minutes to develop these activities. Since "five to ten minutes" is the answer with the largest percentage it could be questioned if this

time is enough to listen to every student, particularly in large classes of around 35 to 40 students, because as author Lazaraton (as cited in Celce-Murcia, 2001) suggests, performance of students in classroom activities should be used to assess their speaking.

GRAPHIC 3:



Graphic for survey question 7 from section 1

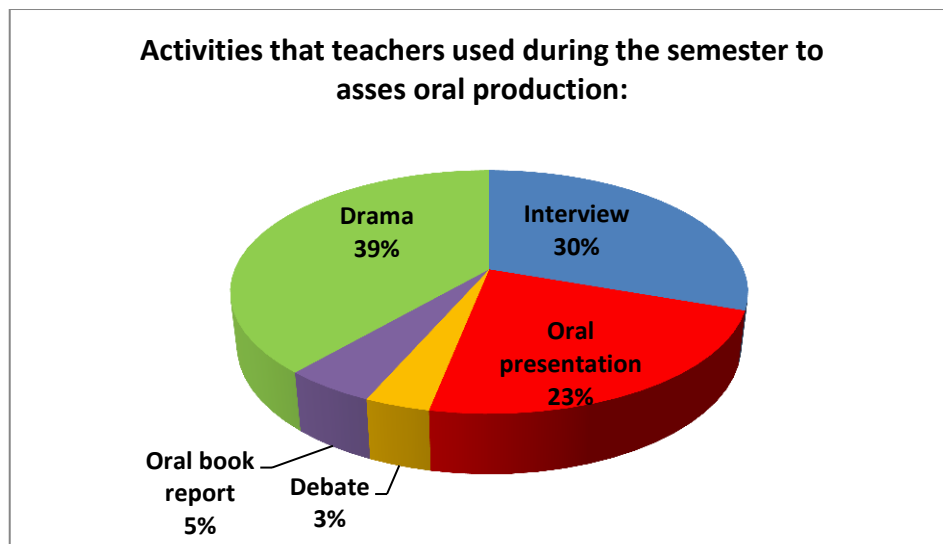
ANALYSIS:

When students were asked if they saw teachers taking notes, assigning a grade or giving any type of feedback while speaking activities were developed, most students agreed that they did not see the teachers performing the tasks previously mentioned. This takes more than half of the respondents with 56%. Then, we have 26% of the population that answered that they saw teachers taking notes or assigning some kind of grade based on their performances. And last but not least, the minority stated that teachers took notes, provided feedback or assigned a grade at a specific moment of the participation with the remaining 18%. This is perhaps one

of the key results of the survey since the purpose is to know if students' speaking was assessed or not, and how exactly it was done. It is relevant to mention that the groups where the surveys were administered had between 35 and 40 students. Based on this, "taking notes" or "assigning a grade" to every single student seemed not too practical, and it is important to remember that practicality is a key element in language assessment, so this is likely the cause.

SECTION II: LARGE SCALE ORAL EXAMINATIONS

GRAPHIC 4:



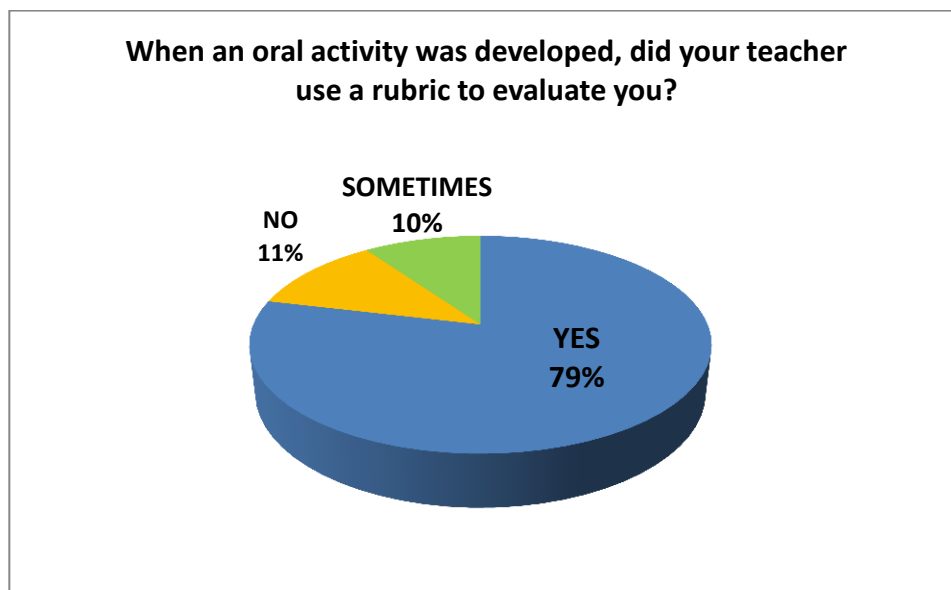
Graphic for survey question 10 from section 2

ANALYSIS:

When asking students if there were any activities used to assess their oral production on a large scale and outside the in-class format (*survey question 8 from section 2*), 93% of students answered that there were some activities developed during the semester to assess their oral production, while 7% of students said there were no activities used for this purpose.

From a list of activities provided to students in the survey, including: interviews, oral presentations, debates, oral book reports, and dramas, 39% of students answered that dramas were used to assess their oral production, followed by interviews with 30%, and oral presentations with 23%, being these three activities the ones with the highest percentages. When asking students how many activities were specifically developed during the semester (*survey question 9 from section 2*), 38% of students answered that only one activity was used during the whole semester, while 34% percent answered that there were three activities developed, and 28% answered two. When asking students how much time was approximately given to develop these activities (*survey question 11, section 2*) most students answered that it depended on the activity. For example, interviews lasted between 5 and 10 minutes and dramas lasted around 20 minutes. Most students were satisfied with the time they were given to develop these activities (*survey question 12 from section 2*): 56% of students said it was enough time while 20% said it was not enough time.

GRAPHIC 5:

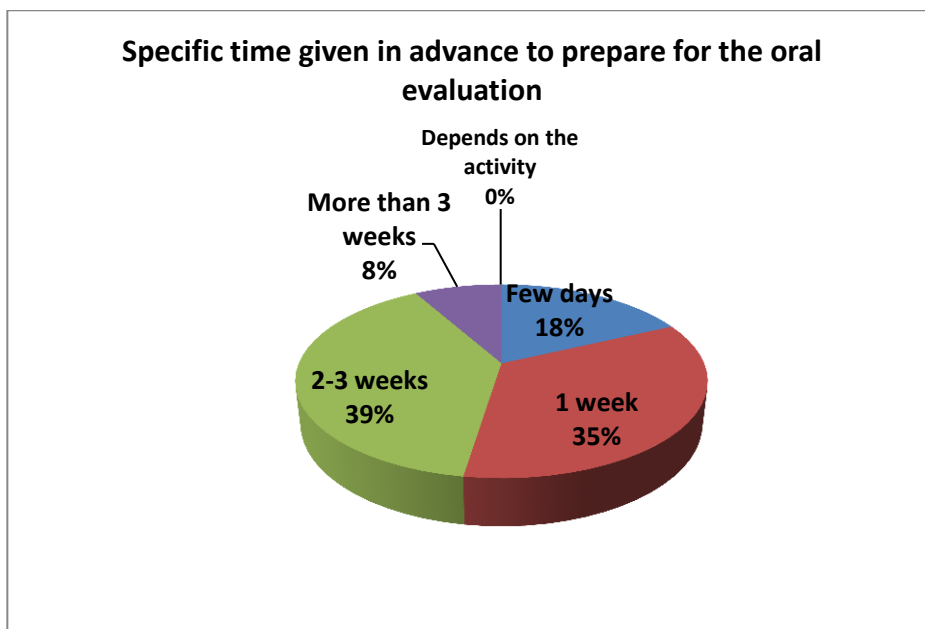


Graphic for survey question 14 from section 2

ANALYSIS:

When students were asked if teachers used rubrics to assess them during any of the oral activities mentioned in the previous graphic, 79% answered yes, while only 11% answered no, and 10% of students answered sometimes. This is very positive, taking into account that the majority of students confirmed the use of these important instruments. Rubrics provide reliability to the assessment process, avoiding the possibility of bias. This means that teachers assessed their students' speaking in a reliable manner, as the theory explains it should be (Brown, 2001). Most students assured that teachers explained the contents of the rubric before carrying out the evaluation (*survey question 15 from section 2*); this represents 56% of the answers. While 20% of students answered that contents of the rubric were not explained before carrying out the activities and 24% answered that contents were explained sometimes. Talking about the contents of the rubrics, the majority of students with 80% answered that each criteria in the rubric included an explanation or description of what was expected of them to achieve for each criteria (*survey question 16 from section 2*). Only 11% answered that the criteria did not include a description or explanation. Explaining students beforehand what it is expected of their performance is very important because this makes the assessment process valid. Validity, which is one important element in language assessment, explains that if a test is exactly measuring what it should be measuring, it is valid.

GRAPHIC 6:

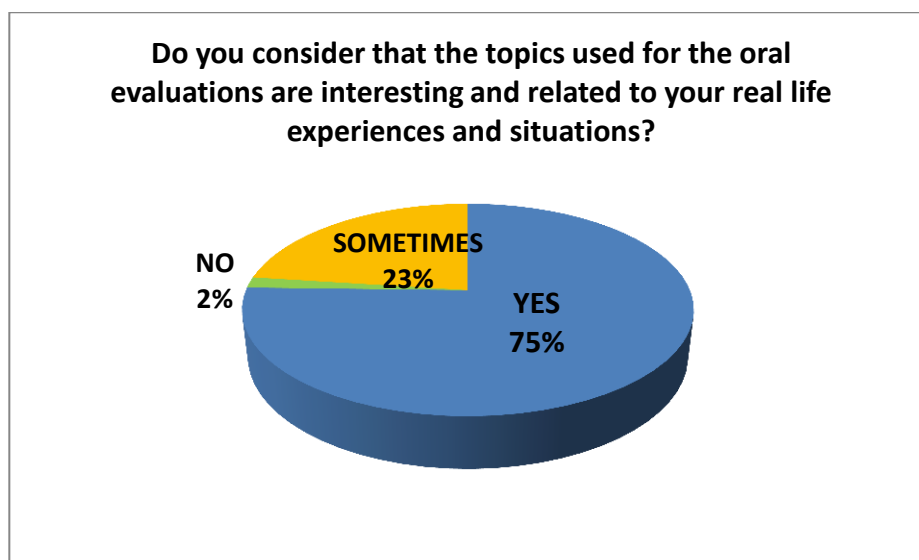


Graphic for survey question 20 from section 2

ANALYSIS:

Students were asked how much time was given in advance to prepare for any large scale oral evaluation. Their responses varied, for example, 39% of the population stated that they were given from two to three weeks to get ready for the evaluations, while 35% stated that they had only one week to prepare for any of the large scale evaluations. Finally, the remaining 26% of the responses are divided like this: 18% mentioned that they were given only few days to prepare for the evaluations and the last 8% responded that they had more than 3 weeks in advance to prepare. Having enough time to prepare for an oral evaluation is really important since students can have more possibilities of obtaining good results. A student might complain that his or her results are bad because he or she did not have enough time to prepare. But with enough time, students can feel that the assessment process was fair.

GRAPHIC 7:

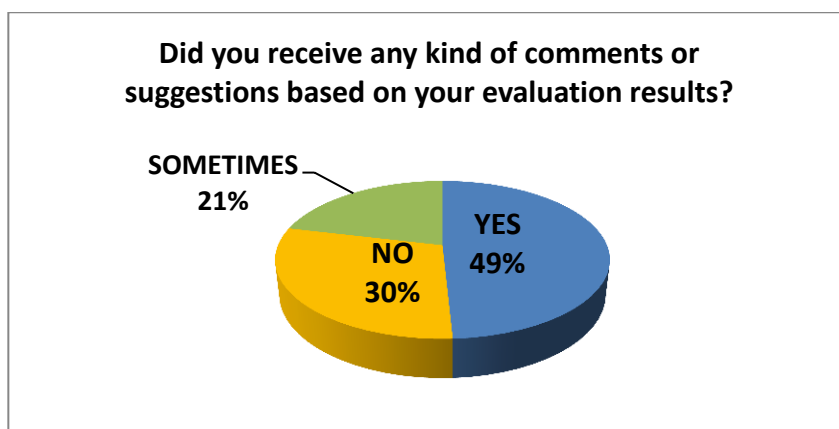


Graphic for survey question 22 from section 2

ANALYSIS:

When asking students if they considered that the topics used in the oral evaluations were interesting and related to their real life experiences and situations, 75% of students answered yes, and only a small percentage answered no. This is an incredibly positive result because when topics used in assessment are related to students' real life experiences and situations, the principle of authenticity is being applied. This gives more opportunities to students to use natural language and in real contexts (Brown & Abeywickrama, 2010), and also if students find the topics interesting, assessment can feel more meaningful for them and not forced or unreal. Students were also asked if these topics were the same or related to the topics they studied in their classes (*survey question 18 from section 2*) and 72% answered positively to this question, while 12% answered negatively, and 16% answered that the topics were sometimes the same or related to the ones studied in their classes.

GRAPHIC 8:



Graphic for survey question 23 from section 2

ANALYSIS:

When asking students if they received any comments or suggestions after their oral evaluations based on the results they obtained, 49% of them answered that they received some kind of feedback after their participation. When asked if these comments were mostly positive or mostly negative (*survey question 24 from section 2*), the majority of students answered that the kind of comments they got were mixed between positive and negative (61% of students), while 28% answered that the kind of comments they got were mostly positive, and only 11% answered that they got mostly negative comments after their participation in oral evaluation activities. In terms of the importance students give to these comments (*survey question 25 from section 2*), the majority of students answered that they have an influence in their learning of the language, an 82% of students to be exact. While only 7% of students answered that the comments or suggestions the teacher makes of their performance has no influence in their learning. Providing comments to students and the effect these comments might have in them is part of the principle of washback. The fact that students are receiving comments after their participation in speaking assessment activities indicates that washback is occurring, and it can be extremely beneficial to both students and

teachers (Brown & Abeywickrama, 2010). But teachers should be careful with the kind of comments they provide. Even if a students' performance is not so good, teachers should find a way to make encouraging comments and not discouraging ones, this because students mentioned that they also got negative comments.

7.2.2 ANALYSIS OF CLASSROOM OBSERVATION RESULTS

Over the course of two weeks, four different intermediate English I groups were observed in order to identify the methodology and instruments used by teachers to assess students' speaking, in this particular case inside the classrooms. Each group was observed four times, making it a total of sixteen classes or thirty-two hours of observation. These are some of the most important aspects observed related to the use of speaking activities in the classroom and their assessment.

➤ Assessment of speaking activities used in the classroom

The type of activities observed in most of these classes involved pair work and small group work. These activities in their majority were: the practice and creation of dialogues, asking and answering questions made by the teacher, the discussion of different questions taken from their text books, and sharing opinions about different situations also presented in their text books. It was unusual to observe teachers including any other activities in their methodology. It is important to mention though that these activities were developed in almost every class, and at least three speaking activities were developed in a single class. In most cases, the time given to develop these activities was between 10 and 15 minutes. Based on the theory that suggests a number of different activities should be used when assessing

speaking (Lazaraton, 2001; Hughes, 2002), we consider that there seems to be a lack of variety in the activities developed during class to assess speaking, but it is positive to mention that these activities, although repetitive, were developed almost in every class.

➤ **Students' participation in the activities**

It could be observed that while most students were given the opportunity of participating, in some cases they were the same students participating over and over again. Another situation was that, since classes are around 35 or 40 students, it was difficult for a teacher to let every single student participate in a *question and answer* activity for example, so only some students could participate per class. It is not that teachers did not want to include everyone in the activities, but it was rather time consuming to have every single student participate or give their opinions. Letting them do that would not have been practical, which takes us back to the principles of language testing: the principle of practicality. It was also observed that while speaking activities in pairs or groups were taking place, if students were not being monitored they easily went off-topic and started talking about other things and even some of them stopped talking and remained silent until the teacher got close to the pair or group. It is positive to mention though, the majority of teachers were going around groups or pairs when speaking activities were being developed. Monitoring and observing students' participation is of great importance because it is part of informal assessment, which as the theory suggests is as important as formal assessment (Spratt, Pulverness, & Williams, 2006). Based on this, it is important that teachers monitor the speaking activities carefully. If activities are not being paid the necessary attention they might feel useless and without purpose, and as if they are used only to consume time.

➤ **Teachers assessment of students' participation**

It was observed that assessment of students' participation in classroom activities was mostly done by observing and monitoring the speaking activities. In some occasions after a student participated teachers would praise them by using comments like: "good", "good job", or "very well". But we could observe that there is not a score or grade given to a student's participation.

Based on the course syllabus, Intermediate English I classes are promoted as "communicative classrooms" and the methodology used as "Communicative Language Teaching", and even in the classroom policies active participation is required. Based on our observation, classes were definitely communicative and communicative activities were developed but as it was mentioned before, there was not any type of score or grade given to this. It is interesting to observe that while according to the syllabus a lot of importance is given to communication in the classes, participation in communicative activities in the classroom is not included in the evaluation system of the course. On the other hand, it is relevant to mention that based on the number of students in the groups, taking notes on every student's participation or giving them a specific grade or score is not so practical, which makes the situation understandable, and as the theory suggests assessment should be practical.

➤ **Topics and situations used in speaking activities**

In most cases, the topics used for the dialogues, conversations, and question/answer activities were taken from the textbooks. Some of those topics included: the environment, relationships, traveling, etc. In most cases, students were asked to place themselves in the

situations presented, kind of asking them what they would do. In dialogues for example, while the activity started by using the dialogue as presented in the book, they were later asked to change parts of it by including their own ideas. In some cases, they were asked about their preferences too. It was rarely observed teachers using other ideas that were not in the textbooks for discussions or conversations which again shows us a lack of variety in the activities used in the class. Still, teachers often tried to relate the topics to students' real life situations, even if the topic did not seem so interesting. This matches the objectives presented in the course syllabus which talked about communicating "in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters". The fact that familiar and routine matters are used in communicative activities is positive since the principle of Authenticity suggests that it leads to the use of natural language as opposed to mechanic or forced language.

Overall, there are both positive and negative aspects in the assessment of speaking inside the classrooms. Among the positive aspects we found that first, teachers used every-day communicative activities in the classroom. Second, the activities were assessed informally by teachers, and even though is not formal assessment, it is still important. And third, real-life situations and contexts were used in these activities. Among the negative aspects we found, first, that there is a lack of variety in the activities used and in the methodology. Second, that feedback, which is part of the washback principle, is used very little. And third, that participation is not included in the evaluation system of the course.

7.2.3 ANALYSIS OF TEACHER'S INTERVIEWS

As part of this research, different teachers of the Foreign Languages Department working in the Intermediate English I courses were interviewed. The purpose of doing this, just as with the other two techniques used in this research, was to ask teachers about the methodology and instruments used to assess students' speaking. It is important to mention that the teachers interviewed have a vast experience in the English teaching field, as well as many years working in the Foreign Languages Department. This is the information we collected from the interviews:

➤ **Assessment of speaking activities inside the classroom**

When teachers were asked about the activities that are part of their methodology to assess speaking, they all had something in common: all of them used activities such as dialogues, discussion questions, role plays, and conversations in their classes. One of them mentioned that other activities such as trivia contests and debates were used in his classroom. This teacher in particular mentioned that these two types of activities can be sometimes time consuming so they are not developed that often. This analysis is important because as the principles of language assessment suggest, the process of assessing should be practical (Brown & Abeywickrama, 2010). The principle of practicality explains that there should be an appropriate amount of time for students to complete a test, in this particular case not a test but an activity, and also, there should be enough time to assess it. We have to take into consideration again the number of students inside the classrooms, which as we have mentioned before is between 35 and 40 students per class. The activities teachers mentioned

are good speaking activities and they are in fact the most common activities inside English classrooms to develop speaking (Harmer, 2001). According to teachers an average time of 10 minutes was used to develop an individual speaking activity, sometimes 15 minutes. Two teachers in particular mentioned that they try not to abuse time because if an activity is way too long students get bored and stop participating. This is also very good because the objective is to have students participating and not to discourage them from doing it.

All teachers mentioned that speaking activities are developed in every class and of course all of them highlighted their importance. When teachers were asked if students' participation in the speaking activities in the class had any percentage in the evaluation system of the course or if participation was included, all of them answered no. One of them mentioned that sometimes participation is included in the evaluation system together with class attendance, but not in this particular semester. This is important information because as it has been previously mentioned, communication is promoted in the course syllabus as one of the most important factors in the Intermediate English I courses, even the objectives are aimed at communication, but if this is so important, should participation in said communicative activities be included in the evaluation system of the course as participation? We know based on the survey results and from the interviews themselves that communicative activities are indeed very important in the courses because different oral tasks are developed during the semester, but only as part of what we have been previously defined as large scale oral examinations, which happen around two or three times in the semester and not as often as the classroom speaking activities happen. What happened with speaking activities in the classroom is that they were informally assessed as opposed to formally assessed. It is

important to note that informal assessment is still assessment and it is as relevant as formal assessment is.

When teachers were asked if they made any comments on students' participation or if students were given any sort of feedback, all teachers said they gave some feedback to their students. One teacher in particular mentioned that sometimes it is difficult to provide feedback mostly because of time limitations. It was very difficult for this person to give feedback to every single student because that would take a lot of time. Another teacher mentioned that he tries to correct students when their participation is not grammatically accurate for example. He does this by interacting with them and talking to them rather than by telling them "this is wrong" or "you made a mistake". This is good because if somehow students are given feedback or comments the principle of washback is being applied, and as the theory suggests (Brown & Abeywickrama, 2010), washback is of great importance because it can have a good influence on students' learning and it gives the students the possibility of improving, when done properly of course.

When teachers were asked about the topics used in the speaking activities, all of them mentioned that the topics were mostly taken from the textbooks, but they were not limited to just discussing or commenting on the situations presented there. Often, they included additional situations but always related to the central topic from the book. When teachers were asked if they considered the topics to be related to students' real-life situations and contexts, they all agreed that most of them are but not one hundred percent of the time. For example, one of the teachers mentioned that textbooks usually included some situations related to the United States or England and students sometimes could not relate to those

situations. Using real-life situations and contexts in speaking activities is of great importance since this is part of the principle of authenticity. If activities seem more authentic, students can feel more encouraged to participate and as mentioned before, it gives them the opportunity of using natural language.

➤ **Assessment in large scale oral examinations**

When teachers were asked about the activities used as part of large scale oral examinations, most of them answered that three activities were developed. Two of the teachers mentioned that two oral interviews were carried out and then a drama. One of them specifically said that the interviews were part of the midterm and final exam, and the drama was a separate task. The rest of the teachers only mentioned the two interviews and the drama but did not specify as part of what they were done. One of the teachers even gave vague answers and did not go too deep on answering questions. Based on this information it can be said that just as with the speaking activities in the class, there is a lack of variety in activities used as part of large scale oral examinations.

About the interviews, when asked if there was only one teacher interviewing or if there were any other teachers helping out with interviews all four answered that they were the only teachers carrying the interviews with their particular group. One of the teachers explained that the idea of having more than one teacher carrying out interviews was discussed but in the end it was not done like that because they were worried about students complaining that it was not their teacher evaluating them. It is important to mention that as the theory says, having more than one tester in an oral activity makes the process more

reliable (Hughes, 2002), and reliability is one of the five principles of language assessment. In the case of the interview, the problem of having only one person carrying out an interview is that in that moment the interviewer has two tasks that he or she has to do at the same time: conducting the interview and keeping track of the student's performance (paying attention to accuracy, pronunciation, writing observations, etc.) and it might happen that the teacher will focus too much on one or the other. A student might question the fact that he or she was not being paid enough attention because the teacher was taking notes for example. This student might feel that his assessment was not reliable.

About the time used to carry these assessment activities, all teachers said that interviews did not have a specific time for each student but most of them took approximately 5 to 10 minutes. Two of the teachers mentioned that since there were students who liked to speak a lot and there were others who spoke very little, time varied. The performance of the drama took around 20 minutes for each group. When asked about how much time of preparation was given to students to prepare for the oral activities, the drama in this particular case, teachers said that they had between three weeks and a month. One teacher mentioned that he gave the rubric that was going to be used to score the activity also a month before.

Regarding rubrics, when asked what kind of instruments they used to assess their students they all mentioned rubrics. Using rubrics is extremely positive because it makes the assessment process more reliable (Harmer, 2001). It is not the same to tell a student that he or she got a seven as a score for an interview without any sort of explanation or record than to tell him or her that they got a grade based on specific criteria previously established and presented in scales (either numerical or non-numerical). When teachers were asked if they

explained the contents of the rubrics to their students before carrying out the activities, three of them had in common that they only mentioned the criteria to be evaluated. But they said that students were told that they could ask any questions if something was not clear. One teacher in particular said that it was a must for him to explain what was going to be evaluated in detail. This teacher said that if he did not do that his students could lose the idea or purpose behind the task. This is an important conclusion that the teacher makes because as mentioned in the theory, results can be better if a student knows what is expected of him or her by using and explaining beforehand the appropriate assessment tools (Harmer, 2001).

When teachers were asked how much time they took to score and give results back to students, all of them agreed that when a class is too big it takes more time to give scores back, around two weeks sometimes, for some of the teachers is even more difficult if they are teaching other classes in the Department. But if a group does not have too many students the teacher could take only a few days to get the scores and give them to students. Finally, when teachers were asked if they made any comments or gave any feedback after speaking tasks are developed, two of them said that what they did was to write some observations in the rubrics they used to score students. One teacher in particular said that it is almost impossible to give comments to each student individually because of time reasons, and making comments in general (for the whole class) can make some students feel uncomfortable. Another teacher said that what he did was to take around 30 minutes of a class after the activities were done and he wrote on the board the most common mistakes made by students and he provided some correction. This last technique is particularly good because it is understandable that teachers did not give comments to students individually because it was

time consuming but giving general feedback is more practical in this situation, and again, it can be noticed that the principle of practicality is applied in assessment of speaking.

Overall, based on the teachers' interviews we found both positive and negative aspects in the assessment of speaking in the classroom and in the large scale oral examinations. However, we want to highlight some of the positive aspects. First, we found that even though classes are large in the Department, teachers work to be as practical as it is possible. Second, speaking activities are developed almost every class, meaning communication and speaking is in fact happening there and it is important. Third, the washback principle is applied in some form, mostly by providing comments to students after their participation in a speaking activity. Fourth, even though teachers are led by their textbooks in the classrooms, they try to relate the contents in them to students' real life situations and contexts. And finally, there are appropriate tools used to assess students' participation in the oral activities part of large scale examinations.

7.2.4 TRIANGULATION OF INSTRUMENTS RESULTS

What the authors propose:	What students said:	What teachers said:	Our analysis:
<p>About the activities that can be used to assess students' speaking in the classrooms, different authors suggest the following activities: discussions, role plays, and conversations (Harmer, 2001); intensive speaking activities, responsive activities, dialogues, conversations, and monologues (Brown, 2001); elicitation activities such as questions and requests for information, use of pictures to elicit descriptions, role plays, and discussions (Hughes, 2002).</p>	<p>According to the survey results, students said the classroom activities used to assess their speaking were: role plays, questions and requests for information, dialogues, conversations, and discussions.</p>	<p>According to teachers of the Intermediate English I courses in the Foreign Languages Department, the activities they used to assess their students' speaking during class were: dialogues, discussions, role plays and conversations, and in some occasions games, and debates.</p>	<p>Based on our observation, the classroom activities used by teachers to assess students' speaking were: dialogues, questions and requests for information, and discussions. Although what both students and teachers said seems to differ in some activities from what was observed, it is important to mention that not every single class of the semester could be observed; therefore we can assume that the other activities both students and teachers mentioned were developed during classes in which we were not present. But it is also important to highlight that some of the activities suggested by the different authors were used in the classes. Maybe more of those activities could have been used to add variety to the classes.</p>
<p>About assessment, authors Spratt, Pulverness, & Williams (2006) define it as the process of collecting information about learners' performance in order to make judgments about their learning. They also mention that learners can be assessed: informally and formally. Among informal assessment procedures we find in-class observation and monitoring of activities. Authors Brown & Abeywickrama (2010) mention that as part of the washback principle giving feedback, making comments, praising students when necessary and even making suggestions is part of assessment.</p>	<p>According to the survey results, most students responded that they did not see their teachers taking notes on their participation, or assigning a grade, which is part of formal assessment. But a small percentage of students responded that their teachers did give them a grade for their participation. Also, the majority of students responded that their teachers gave them comments on their participation; this is part of informal assessment.</p>	<p>When teachers were asked if students' participation in classroom activities was included in the evaluation system of the course, all of them answered no. This means, that classroom activities were not formally assessed. On the other hand, all teachers said that they gave feedback to their students whenever possible and corrected them whenever they could too, which is part of informal assessment</p>	<p>During our observation sessions we could see that teachers did not take notes on students' participation nor gave them a grade for it, which is part of informal assessment. But on the other hand, we could observe that teachers monitored speaking activities, gave comments on students' participation, gave them praise when necessary, and also made corrections whenever possible, which is part of informal assessment. While classroom activities were not formally assessed, there was informal assessment, which is valid and just as important as suggested by the authors. Giving comments making suggestions, and even praising like the teachers did means that the washback principle is being applied.</p>
<p>When assessment occurs, five very important principles should be applied according to authors Brown & Abeywickrama (2010). One of those principles is practicality, and as the authors</p>	<p>According to the survey results, the majority of students responded that their teachers gave them around 5 or 10 minutes to develop a speaking</p>	<p>According to teachers' interviews, the time usually given to develop a speaking activity in the classroom was 10 minutes, and in some</p>	<p>Based in our observation, the time given to develop speaking activities in the classroom was between 10 and 15 minutes in most cases. And as we were told by teachers in the interviews, they</p>

<p>mention it involves elements such as giving an appropriate amount of time for students to complete an activity or test.</p> <p>About the use of time in large scale oral examinations, Hughes (2002) suggests that a test should be as long as possible because not too much information will be obtained in less than 15 minutes for example. To get the necessary information from a students' performance around 30 minutes would be needed approximately. And as mentioned before the use of appropriate time involves the practicality principle.</p>	<p>activity in the classroom.</p> <p>According to survey results, the majority of students responded that the time given to the large scale oral examinations depended on the kind of activity developed. Interviews lasted between 5 and 10 minutes, and dramas lasted around 20 minutes.</p>	<p>cases 15 minutes.</p> <p>According to teachers' interviews, time given to develop a large scale oral examination also depended on the activity. Interviews lasted between 5 and 10 minutes, but not always because some students spoke more than others, so in those cases time varied. Dramas lasted around 20 minutes.</p>	<p>provided around 10 or 15 minutes to develop in-class activities. Since these activities are part of informal assessment and not formal assessment, we consider that the time given for the classroom activities was appropriate. Regarding the large scale oral examinations, the theory suggested that an oral test should last around 30 minutes in order to gather the necessary information from a student's abilities, but if we take into account what Hughes (2002) suggests, in a class with around 35 students giving them around 30 minutes for each (for example in an interview) would not be practical because it would take weeks to assess all students in a single group. Teachers adjust the time they have to assess all students so each one of them can have the evaluation done. This is what's practical in our department and therefore, practicality is being applied.</p>
<p>About the Authenticity principle, authors Brown & Abeywickrama (2010) mention that it involves the use of "real world" elements in testing. These real world elements can include real-life contexts or situations that a student can recognize and relate to, natural language and meaningful and interesting topics.</p>	<p>According to survey results, the majority of students responded that the topics used in both the classroom activities and the large scale oral examinations were interesting and related to their real life situations and contexts.</p>	<p>According to teachers' interviews, they all agreed that most of the topics used in speaking activities were based on real-life situations and contexts but not one hundred percent of the time. Since most topics were taken from the textbooks, sometimes situations related to other countries or cultures were presented, and in some occasions students felt they were not significant for them, but teachers tried to make those situations relatable.</p>	<p>Topics for classroom activities, which are the ones we observed, were indeed taken from textbooks, but in spite of this, topics were relatable for students most of the time and teachers tried to associate students' experiences, opinions, and situations to them. It was very common to see a teacher saying things like "what would you do in that situation?" or "what do you think?" Some of the topics used in the classroom speaking activities were: the environment, relationships, traveling, etc. which can be considered "global" topics. We can conclude that the authenticity principle was applied in most cases. Another important observation regarding how teachers dealt with the authenticity issue, we could see that teachers prepare the activities they will use in their classes with a lot of thought and taking into consideration aspects like authenticity.</p>
<p>According to Hughes (2002), three possible formats that can be used in large scale oral testing are interviews, activities that</p>	<p>According to survey results, students responded that the activities used in large</p>	<p>According to teachers' interviews, the activities they used as part of large scale oral testing</p>	<p>Even though it was not possible to observe the development of large scale oral examinations, we can analyze the situation by looking at</p>

<p>involve interaction with peers like discussions or debates, and response to tape recordings. And Brown (2001) also suggests activities that involve one-on-one tester/ test-taker relationship, like an interview, and also “live” performances.</p>	<p>scale oral testing were dramas, interviews, oral presentations, oral book reports and debates.</p>	<p>were interviews and a drama.</p>	<p>the results from both the survey and teachers interviews. We can see that the activities students mentioned were used for the large scale oral examinations differ in some way with the activities mentioned by teachers. The only coincidences are interviews and dramas. Student’s responses suggest variety, but teacher’s responses do not show this. We consider that variety is important and that it should be taken into account when deciding what activities will be used as part of the large scale oral examinations.</p>
<p>About having reliability in testing, Hughes (2002) suggests that having more than one tester at the time of assessing students could make the process more reliable. He explains that in an interview for example, it is difficult conducting it and keeping track of a learner’s performance at the same time, and also, having only one person testing could be not as reliable as it is desired.</p>	<p>When students were asked how many testers were present at the time interviews were carried out, they responded that only one teacher was present during the interview.</p>	<p>According to teacher’s interviews, at the time of carrying out an interview with students, they were the only testers in their respective groups.</p>	<p>As theory suggests, if two testers are assessing students the results can be considered more reliable. When a teacher is carrying out an interview there are two main things he or she has to do: formulate the questions and listen to students’ answers while also paying attention to aspects like grammar, pronunciation, etc. and the other one is getting a record of students performance, sometimes even taking notes of the mistakes he or she made. One of these two tasks might distract the teacher from doing the other one or maybe the teacher could focus on one more than the other and this can be perceived by the students as unreliable to the point where he or she might question the results. If you have two testers, each one of them can be in charge of one task and therefore provide more reliability to the assessment process.</p>
<p>According to Brown & Abeywickrama (2010), testing reliability involves not only the test itself but also other factors like the scoring process. Since testing speaking is a subjective process, subjectivity can be reduced by using the appropriate assessment tools (Harmer, 2001). The use of rubrics or assessment scales can make the scoring of oral performance more valid and reliable.</p>	<p>According to survey results, almost 100% of students responded that rubrics were used in order to assess them in large scale oral examinations.</p>	<p>According to teachers’ interviews, rubrics were used every time a student was assessed orally in a large scale oral examination.</p>	<p>We can see that both teachers and students agree on the fact that rubrics were used every time a large scale oral examination was carried out. Using rubrics is extremely positive because as the theory suggests, it makes the assessment process more reliable. It is not the same to tell a student that he or she got a seven as a score for an interview without any sort of explanation or record than to tell him or her that they got a grade based on specific criteria</p>

			previously established and presented in scales (either numerical or non-numerical). We can conclude based on this that the reliability principle is applied.
About the washback principle, when washback occurs in assessment, it will positively influence how teachers teach and how students learn (Brown & Abeywickrama, 2010). Giving feedback, making comments and not just giving a grade, praising students when necessary and even making suggestions are part of washback.	According to survey results, most students responded that they received comments or suggestions after the evaluation activities were done.	According to teachers' interviews, in some occasions they gave their students written feedback or comments, and when it was possible they did it verbally by taking some time in the class to address all students. Time sometimes made it difficult for teachers to give comments individually.	Based on both students and teachers responses, we can see that some sort of feedback was given and also comments were made on students' participation. Also, that these comments or feedback were done in both ways: spoken or written. And in most cases, when it was spoken, it was done in general rather than individually because of time reasons. This technique is particularly good because it is understandable that teachers did not give comments to students individually because it was time consuming but giving general feedback is more practical in this situation, and again, it can be noticed that the principle of practicality was applied. Also, it shows that the washback principle is applied.

CHAPTER VIII

8. CONCLUSIONS, RECOMMENDATIONS, AND LIMITATIONS

8.1 CONCLUSIONS

After gathering and analyzing the data collected from students surveys, teacher's interviews and class observation, and based on our research objectives the following list of conclusions was reached.

It is important to mention that several of the conclusions are about the application of the five principles of language assessment included in authors Brown & Abeywickrama's book "Language Assessment: Principles and Classroom Practices". The authors suggest that when it comes to language assessment these five principles should be applied not only to formal assessment but also to any other kind of assessment. The five principles are: validity, reliability, practicality, authenticity, and washback. These terms were defined in the theoretical framework of this report.

- While teachers of the Foreign Languages Department used some of the activities recommended in the TEFL literature, there seemed to be a lack of variety in both the classroom speaking activities and the activities used in large scale oral examinations developed during semester II-2014.
- Even though the same classroom speaking activities were used frequently, they were developed in every class and an appropriate amount of time was provided to develop

these activities proving that as the course syllabus suggested, Intermediate English I courses are communicative classes.

- Classroom speaking activities were not formally assessed but rather informally assessed. There were not any specific grades assigned to a student's participation and participation was not included in the evaluation system of the course, but comments on students' participation were made by teachers and activities were monitored most of the times they were developed.
- Rubrics were used as instruments by teachers in order to assess students' speaking abilities in the large scale oral examinations. This instrument included different criteria that described what students had to do and they were also explained in advanced to students in most cases.
- The validity principle was applied when students' speaking abilities were assessed by teachers using appropriate assessment tools, particularly in the large scale oral examinations activities in which rubrics were used. Also, the contents that were presented in the classes, through their books for example, were used for the evaluation activities meaning that students were assessed at their level with contents that belonged to their level assuring that the evaluation process was valid.
- The reliability principle was applied when students' speaking abilities were assessed by using appropriate assessment tools. These instruments included a variety of criteria with their appropriate explanation and were explained to students before the

assessment activities were carried out. The same criteria were established for all students and the same outcome was expected of each one of them.

- The practicality principle was applied when students' speaking abilities were assessed in both the classroom speaking activities and the large scale oral examinations. Regarding time and considering the number of students in each class which was around 35 students, teachers used an amount of time that allowed them to have almost all students participate. In large scale oral examinations, even though TEFL literature suggests using around 30 minutes for a speaking evaluation activity, for teachers it was not practical to use this amount of time with each student, so they used an amount of time that was practical for them considering the number of students they had.
- The authenticity principle was applied at the time of assessing students' speaking abilities by using real-life contexts and situations in the different activities both in the classroom speaking activities and in the large scale oral examinations, and even when the contents were taken straight from the textbooks, teachers tried to relate these contents to students' experiences and lives.
- The washback principle was applied at the time of assessing students' speaking abilities by giving them feedback, making comments and providing suggestions after an activity was carried out. Students expressed that these comments had an influence in their learning and that they were mostly positive comments.

8.2 RECOMMENDATIONS

Based on our conclusions, the following recommendations are proposed:

- In order to add more variety to the activities used in both the classroom speaking activities and the large scale oral examinations, teachers of the Intermediate English I courses could try consulting theory from different authors to gather ideas on speaking activities they could use in their classes and in their speaking tests so monotony can be avoided.
- Students' participation in speaking activities developed inside the classroom should be included in some way in the evaluation system of the course. This could avoid having students' speaking abilities being assessed only two or three times during a semester and making the assessment of speaking more constant while using unprepared and more natural situations as opposed to a presentation or drama.
- Teachers should go beyond the books and try to include topics of their own in the speaking activities they develop in both the classrooms and as part of the large scale oral examinations. We understand that contents in books must be covered but adding something extra could be good for classes.
- Teachers should take into consideration using more than one tester whenever possible at the time of assessing their students speaking abilities, particularly during interviews. Conducting interviews as an assessment activity is not easy and the tester

is doing at least two different tasks at the same time: listening to the student being assessed and taking notes on the student's performance paying close attention to aspects like grammar and pronunciation, among others. By having a second tester these tasks can be split and each tester can focus on one particularly, making the results of this activity more reliable and valid and also making the student feel more confident that his or her results are reliable.

- Teachers should continue with all the positive aspects found in their assessment of students' speaking abilities, like continue applying the principles of language teaching and trying to improve their application in each of the activities they develop in their classes and as part of assessment.

8.3 LIMITATIONS

During the research process the group encountered some difficulties or limitations, among the most important ones we have:

- It was difficult to get the appropriate permissions to work with some of the Intermediate English I groups. Out of the eight groups of Intermediate English I we needed four groups to work with, but it was necessary to ask more than four groups because in some cases our request to work with a specific group was rejected by the teacher. Teachers usually claimed that they did not have time for doing this in their classes.
- At the time of carrying interviews with teachers we encountered some hostility, in one case particularly. Our questions were answered superficially and very limited answers were provided by the person being interviewed. Also, in this same case, it was necessary to re-schedule the day of the interview at least three times since we were not attended the day that it had been originally scheduled.
- Not all students in the group attended classes the day the surveys were administered.

APPENDIXES

APPENDIX A: students' survey



Universidad de El Salvador
Facultad de Ciencias y Humanidades
Departamento de Idiomas Extranjeros

Nombre de la encuesta: Metodología e instrumentos para evaluar la producción oral de los estudiantes de Inglés Intermedio I.

Objetivo de la encuesta: Recolectar información relacionada con la evaluación de la producción oral de los alumnos de los cursos de Inglés Intermedio Intensivo I del Departamento de Idiomas Extranjeros de la Universidad de El Salvador.

Código: _____

No: _____

Facultad de Ciencias y Humanidades

Departamento de Idiomas Extranjeros

1. Durante el semestre en curso, ha desarrollado su profesor o profesora actividades para promover la producción oral y la comunicación?

- a. SI
- b. NO
- c. Algunas veces

3. ¿Qué tan seguido se desarrollan este tipo de actividades (pregunta 2) en tus clases?

- a. Siempre
- b. Usualmente
- c. Seguido
- d. Algunas veces
- e. Raramente
- f. Casi nunca

5. ¿Consideras que el tiempo que tu profesor o profesora proporciona para desarrollar las actividades antes mencionadas es satisfactorio y suficiente?

- a. SI b. NO c. Algunas veces

6. ¿Tu participación en clases tiene alguna ponderación o porcentaje en tu evaluación del semestre?

- a. SI b. NO c. No sabe

2. De la siguiente lista de actividades, selecciona las que tu profesor o profesora ha utilizado durante la clase para promover la producción oral y comunicación:

- a. Discusiones
- b. Conversaciones
- c. Juegos de roles (role play)
- d. Diálogos
- e. Preguntas y respuestas
- f. OTROS: _____

4. Cuando alguna de las actividades antes mencionadas se desarrolló ¿Cuánto tiempo fue proporcionado para hacerlo?

- a. 1-5 minutos b. 5-10 minutos c. 10-15 minutos
- d. 15-20 minutos e. más de 20 minutos
- f. Dependía de la actividad (aclarar):

7. ¿Has podido observar a tu profesor o profesora tomando nota, asignando una calificación, o proporcionándote comentarios por tu participación en clases?

- a. SI
- b. NO
- c. Algunas veces

Facultad de Ciencias y Humanidades

Departamento de Idiomas Extranjeros

2

2

SECCION 2: EVALUACIONES ORALES

SECCION 2: EVALUACIONES ORALES

8. ¿Se han desarrollado durante este semestre algún tipo de evaluaciones formales que promuevan la producción oral?

- a. SI b. NO

9. ¿Cuántas actividades o evaluaciones orales has realizado durante el semestre?

- a. Una
b. Dos
c. Tres o más

11. Cuando alguna de las actividades antes mencionadas se realizó ¿Cuánto tiempo fue proporcionado para hacerlo?

- a. 1-5 minutos b. 5-10 minutos c. 10-15 minutos
d. 15-20 minutos e. más de 20 minutos
f. Dependía de la actividad (explicar):

14. ¿Cuándo se va a realizar una evaluación oral, ¿tu profesor o profesora utilizó una rúbrica* para evaluarte?

- a. SI b. NO c. Algunas veces

*RUBRICA: Una tabla donde se describen criterios y niveles de calidad de cierta tarea, objetivo, o competencia que clarifica lo que se espera del trabajo del alumno.

****NOTA: si la respuesta a esta pregunta es "NO", no responder preguntas 15, 16, y 17.**

10. Marca de la siguiente lista las actividades que tu profesor o profesora ha utilizado durante el semestre para evaluar tu producción oral:

- a. Entrevista f. OTRAS:
b. Presentación oral _____
c. Debate _____
d. Reportes orales de libros _____
e. Dramas o actuaciones

12. ¿Consideras que el tiempo que se proporciona para realizar alguna de las actividades antes mencionadas es satisfactorio y suficiente?

- a. SI b. NO c. Algunas veces

13. ¿Cuándo se va a realizar una evaluación oral, ¿tu profesor o profesora explica detalladamente lo que tendrás que hacer?

- a. SI b. NO c. Algunas veces

15. Si se utilizó una rúbrica, ¿el profesor o profesora explico sus contenidos con anticipación a la actividad de evaluación?

- a. SI b. NO c. Algunas veces

16. ¿Los criterios o elementos incluidos en la rúbrica incluían una explicación o descripción de lo que se te evaluaría?

- a. SI b. NO c. Algunas veces

17. De la siguiente lista, marca los elementos o aspectos que fueron incluidos en la rúbrica utilizada para evaluar las actividades orales:

- a. Gramática e. Tiempo
b. Vocabulario f. Uso del idioma
c. Pronunciación g. OTROS: _____
d. Fluidez _____

18. ¿Los temas de las evaluaciones orales eran los mismos que estudiaste durante clases?

- a. SI b. NO c. Algunas veces

19. ¿Fueron las instrucciones y objetivos de la actividad oral entendibles cuando tu profesor o profesora los explico?

- a. SI b. NO c. Algunas veces

20. ¿Cuánto tiempo de anticipación te proporciono tu profesor o profesora para prepararte para la evaluación oral?

- a. Pocos días b. Una semana
c. Dos o tres semanas d. Más de tres semanas
e. Dependía de la actividad (explicar: _____

21. ¿Cuánto tiempo después de la evaluación oral recibiste tus resultados o notas?

- a. Uno o dos días después
b. Una semana después
c. Dos o tres semanas después
d. Más de tres semanas después

22. ¿Consideras que los temas que se utilizan para las evaluaciones orales son interesantes y relacionados a situaciones y experiencias de la vida real?

- a. SI b. NO c. Algunas veces

23. Después de que se realizaron las evaluaciones orales, ¿recibiste algún tipo de comentario o sugerencia en base a tu participación?

- a. SI b. NO c. Algunas veces

24. Si recibiste algún tipo de comentarios o sugerencias por parte de tu profesor p profesora, estos fueron:

- a. Positivos b. Negativos c. ambos positivos y negativos

25. ¿Los comentarios que tu profesor o profesora hace de tu participación tienen algún tipo de influencia en tu aprendizaje del idioma?

- a. SI b. NO c. Algunas veces

Facultad de Ciencias y Humanidades

Departamento de Idiomas Extranjeros

Si la entrevista fue la actividad utilizada para evaluarte:

26. ¿Cuántas personas estuvieron presente para evaluarte al momento de llevar a cabo la entrevista?

- a. Una persona (tu profesor)
- b. Dos personas (incluyendo tu profesor)
- c. Tres personas (incluyendo tu profesor)
- d. Más de tres personas (incluyendo tu profesor)

27. Al momento de llevar a cabo la entrevista, tu profesor o profesora:

- a. Utilizó una guía con posibles preguntas
- b. Improvisó las preguntas

28. ¿Las preguntas estaban relacionadas a un mismo tema o situación?

- a. SI
- b. NO
- c. Algunas veces

Si un debate fue la actividad utilizada para evaluarte:

29. ¿Recibiste asesoría o algún tipo de ayuda por parte de tu profesor o profesora antes de llevarlo a cabo?

- a. SI
- b. NO

Si una presentación oral fue la actividad utilizada para evaluarte:

30. ¿Cuántos temas tuviste que preparar?

- a. Un tema
- b. Dos o tres temas
- c. Más de cuatro temas

¡Las investigadoras agradecen su colaboración al completar esta encuesta!

APPENDIX B: observation sheet



**UNIVERSITY OF EL SALVADOR
SCHOOL OF ARTS AND SCIENCES
FOREIGN LANGUAGES DEPARTMENT**



OBSERVATION SHEET

Date of session: _____
 Schedule: _____
 Subject: _____

Purpose: To observe the methodology and instruments used to assess speaking during classes of the Intermediate English I courses at the Foreign Languages Department of the University of El Salvador.

Instruction: Complete items according to what is observed during the class.

1. Teacher uses activities to develop speaking during class.
 Always Usually Often Sometimes Rarely Almost never

2. Activities used by teacher during the class to develop speaking and time used to develop them:

<u>ACTIVITY</u>	<u>TIME</u>
_____	_____
_____	_____
_____	_____
_____	_____

3. There is a variety of activities used to develop speaking during class.
 Always Usually Often Sometimes Rarely Almost never

4. Number of activities used in a class to develop speaking:

5. Students are given the opportunity of participating in some way during the class.
 Always Usually Often Sometimes Rarely Almost never

6. Number of students who are given the opportunity of participating during the class.
 ALL STUDENTS MOST STUDENTS SOME STUDENTS VERY FEW STUDENTS

7. Teacher takes notes or assigns any type of grade to student's participation.
 Always Usually Often Sometimes Rarely Almost never

8. Teacher provides comments during/after student's participation in class. Always Usually Often Sometimes Rarely Almost never
9. If there are comments made to student's participation, they are:
MOSTLY POSITIVE BOTH POSITIVE AND NEGATIVE MOSTLY NEGATIVE
10. Topics used in speaking activities are related to real-life situations and contexts.
Always Usually Often Sometimes Rarely Almost never

APPENDIX C: Teachers interview



UNIVERSITY OF EL SALVADOR
SCHOOL OF ARTS AND SCIENCES
FOREIGN LANGUAGES DEPARTMENT



INTERVIEW FOR TEACHERS

Date of session: _____

Schedule: _____

Teacher's name: _____

Purpose: To interview teachers of the Intermediate English I courses of the Foreign Languages Department regarding the methodology and instruments they use to assess students' speaking skills during class and as part of large-scale oral examinations.

1. How long have you worked here / what's your experience? What courses have you taught?
2. What kind of activities do you use to practice or develop speaking in your classes?
3. How much time do you provide to develop these activities?
4. How often do you develop speaking activities during your classes? Is this an every-day thing?
Do you always have time to do it?
5. Do you take into account students' participation during these activities as part of their evaluation and are they being scored/ assigned any kind of grade?
6. How do you decide or select the topics that are going to be used in your speaking activities?
7. Do you make comments or give any kind of feedback to students' when they participate in speaking activities during class?
8. What kind of activities do you use for the speaking tasks and midterms?
 - If the interview is used, are you the only tester or is there someone else present?
 - If the interview is used, do you prepare your questions beforehand?
 - If a debate is used, do you give your students any kind of advisory or tutoring before?

- If an oral presentation is used, how many topics do students have to prepare and how many do they actually develop? How are the topics for these presentations chosen?
9. Do you consider that the topics used for the task are related to students' real-life situations and contexts?
 10. How many speaking tasks are done during the semester?
 11. How much time approximately is it given to students for developing a speaking task?
 12. How much time before doing the tasks do you let your students know about them and do you explain what they have to do in detail?
 13. What kind of instruments do you use when you are evaluating your students' speaking and what are the main criteria to be evaluated?
 14. Do you explain students what you are going to evaluate exactly (what grammar structures, what vocabulary, etc.) and the criteria to be evaluated?
 15. How much time does it usually take you to get the scores and give them back to your students?
 16. Besides giving them a score, do you also provide your students with any comments or feedback on their performance after carrying out the tasks?

BIBLIOGRAPHY

- Bailey, K. (1998). *Learning about Language Assessment: Dilemmas, Decisions, and Directions*. Boston, MA: Heinle&Heinle.
- Brown, H.D. (2001). *Teaching by Principles: An Interactive Approach to Language Pedagogy* (2nd ed.). Longman.
- Brown, H.D. & Abeywickrama, P. (2010). *Language Assessment, Principles and Classroom Practices* (2nd ed.). Pearson Longman
- Harmer, J. (2001). *The Practice of English Language Teaching* (Rev. ed.). Essex, England: Longman ELT.
- Hughes, A. (2002). *Testing for Language Teachers* (Rev. ed). Cambridge University Press.
- Lazaraton, A. (2001). Teaching Oral Skills. In M. Celce-Murcia (Ed.), *Teaching English As a Second Or Foreign Language* (pp. 103-111). Boston, MA: Heinle&Heinle.
- Spratt, M., Pulverness, A., & Williams, M. (2005). *The TKT Course* (3rd ed.). Cambridge University Press.

WEB BIBLIOGRAPHY

- Definitions were retrieved from *dictionary.reference.com*.