



**Universidad Católica de la Santísima Concepción
Facultad de Educación**

Magíster en Ciencias de la Educación Mención Evaluación de los Aprendizajes.

“Analysis of the procedures used by the teachers of English from a public high school from Curanilahue in order to assess speaking and its impact on students’ performance in this communicative skill”

By

Mariela Andrea Martínez Fuentealba

Research seminar to obtain the degree in Master in Education

Supervising teacher: Mg. Roxana Correa Pérez

Concepción, May 2018.

Content Index

	PAGES
Acknowledgements -----	i
Abstract -----	ii
Resumen -----	ii
Introduction -----	iv
CHAPTER I: RESEARCH PROBLEM AND OBJECTIVES-----	1
1. Research problem -----	2
2. Context -----	3
3. Research Question -----	4
4. Research Assumptions -----	4
5. Objectives -----	5
5.1 General Objective -----	5
5.2 Specific Objectives -----	5
6. Significance of the Study -----	5
7. Validity -----	6
8. Consequences -----	6
CHAPTER II: THEORETICAL FRAMEWORK -----	7
1. Assessment -----	8
2. Second Language Assessment -----	10
3. The Assessment of Speaking -----	14
3.1 Test Specifications -----	17
3.2 Scoring Instruments -----	19
3.3 Tasks -----	21
CHAPTER III: METHODOLOGY -----	24
1. Research Paradigm -----	25

2. Type of Research -----	26
3. Research Design -----	27
4. Research Context -----	28
5. Participants -----	28
6. Instruments and Procedures -----	29
6.1 Categories -----	31
7. Data Analysis procedure -----	34
CHAPTER IV: DATA ANALYSIS AND INTERPRETATIONS -----	35
1. Analysis of Specific Objective 1 -----	36
2. Analysis of Specific Objective 2 -----	42
3. Analysis of Specific Objective 3 -----	58
4. Analysis of Specific Objective 4 -----	63
5. Discussion -----	67
CHAPTER V: CONCLUSIONS -----	70
1. Conclusions -----	71
2. Limitations -----	76
3. Further Research -----	77
REFERENCES -----	78
APPENDIXES -----	83

Tables Index

	PAGE
Table III.1: Criteria to select the participant students -----	29
Table III. 2: Criteria to select the participant teachers -----	29
Table IV.1: Assessment activities -----	36
Table IV. 2: Assessment instrument -----	38
Table IV.3: Assessment criteria -----	39
Table IV.4: Assessment techniques -----	40
Table IV.5: Marks ranges -----	41
Table IV.6: Purpose to assess oral production -----	43
Table IV.7: Frequency of assessment -----	45
Table IV.8: Dimensions assessed -----	47
Table IV.9: Scoring instruments used -----	49
Table IV.10: Factors that affect oral assessment -----	51
Table IV.11: Frequency -----	52
Table IV.12: Types of evaluative tasks -----	53
Table IV.13: Scoring instruments -----	54
Table IV.14: Dimensions of assessment -----	55
Table IV.15: Feedback -----	57
Table IV.16: Teachers' accomplishment of test specifications -----	59
Table IV.17: Luoma's adequate aspects: Teacher 1 and Teacher 2 -----	59
Table IV.18: Luoma's minimally adequate aspects: Techer 1 and Teacher 2 -----	61

Table IV. 19: Luoma's *not observed* aspects: Techer 1 and Teacher 2 -----62

**Table IV.20: T1's assessment task, assessment instrument
and test specifications and students' performance in speaking -----64**

**Table IV.21: T2's assessment task, assessment instrument and
test specifications and students' performance in speaking -----65**

Acknowledgements

I would like to express gratitude to all the people who supported me during this process:

First of all, to my parents and family who believe in me and give me the support to start and continue this master degree. Without them, I would not have done this. I want to thank sincerely to my parents since they encouraged and helped me as much as they could and they would feel very proud of having a daughter with this degree.

Second, I want to express my sincere gratefulness to my ex-husband. He was with me during almost the whole process. I want to thank him for every time he went to take me to Laraquete to arrive in Curanilahue at 11:30 pm, for all his patience when I was stressed, for all the responsibility he had to take at home because I did not have time or I was tired.

Third, I would like to give special thanks to the high school where I work, especially to my principal who gave me the facilities to attend classes at the university and allow me investigate the reality inside the high school.

Forth, I would like to express gratefulness to my guide teacher, Roxana Correa, for her patience and encouragement in the process of developing the thesis project. She gave me the guidelines to start and finish this research. Without her, I would not have finished it.

Finally, to my boyfriend since he was the person who changed my life and encouraged me to finish this process.

Abstract

The present research intends to study how the way that teachers assess the speaking skill influences students' performance in the English subject at a public school. For this reason, it was made a qualitative research study where it was used the case study methodology since it allows collecting and analyzing data from various cases.

To study the influence on students' marks a teacher has when he or she assesses speaking, two teachers and twelve students participated in this research. Both teachers were observed using a checklist to identify which instruments and procedures teachers use at the moment of assessing speaking and a rating scale to identify the test specifications that teachers utilize when assessing speaking. Furthermore, the teachers and the students were interviewed using a semi-structured interview to identify what learners declared teachers use to assess speaking. After having all the data collected, it was compared and analyzed to distinguish the real influence that teachers have on students' performance when they assess speaking in the English class.

In this way, it is intended to explore how a teacher assesses the speaking skill at a public high school and how it affects students' performance.

Keywords: Speaking, speaking assessment, tasks, scoring instruments and test specifications

Resumen

La presente investigación tiene la intención de estudiar cómo la forma de evaluar la habilidad de expresión oral influye en el rendimiento de los estudiantes en la asignatura de inglés en un liceo. Por esta razón, se realizó una investigación cualitativa donde se utilizó el estudio de caso lo que permitió recolectar y analizar datos de varios casos.

Para estudiar la influencia de la manera de evaluar de un profesor en la evaluación de la habilidad oral, dos profesores y doce alumnos participaron en esta investigación. Los docentes fueron observados usando una lista de cotejo para conocer qué instrumentos y procedimientos los profesores utilizan al momento de evaluar la expresión oral y una escala de apreciación para identificar las *test specifications*¹ que los profesores usan cuando evalúan la expresión oral. Además, los profesores y los estudiantes fueron entrevistados usando una entrevista semi-estructurada para saber qué declararon acerca de cómo los educadores evalúan la expresión oral. Después de tener toda la información recolectada, se comparó y analizó para saber si existía una influencia real en la forma de evaluar la expresión oral por parte de un profesor en las notas de los alumnos.

De este modo, se pretende explorar cómo la manera que el profesor utiliza para evaluar la expresión oral en un liceo afecta el rendimiento que los estudiantes obtienen en esta habilidad.

Palabras clave: evaluación, evaluación de la expresión oral, instrumentos de evaluación y test specifications

¹ Test specifications propuestas por Luoma (2011)

Introduction

When a person wants to learn English, his or her objective is principally to develop the speaking ability. However, it is the most difficult ability to teach and to assess for teachers. It may happen because students have few hours of English a week or maybe because there are many students in a classroom. Another issue could be that English is not an official language in Chile so learners do not use it outside the classroom.

During the last years, the Ministry of Education has changed its focus in relation to how to teach English. Nowadays, the Ministry wants students to develop the four communicative skills (reading, listening, speaking and writing) because all of them are necessary to understand this language and to be able to communicate in English.

In my experience as a teacher of English, I have observed that the productive skills are not taught at all in public high schools, especially speaking. Besides, teachers do not often assess speaking in classes. The main aspect that influences it is the lack of time of class they have and the quantity of students inside the classroom. In addition, when they assess, they do not have a suitable scoring instrument and they do not use the proper criteria since they usually take them from the English course book or the English program. As a consequence, the way how teachers assess could interfere in the students' marks.

For this reason, in this research, it will be analyzed how teachers assess speaking in the subject of English, which procedures, instruments, tasks and test specifications they use and how it impacts the students' performance in the context of public education.

CHAPTER I
RESEARCH PROBLEM AND OBJECTIVES

1. Research Problem

The Chilean Ministry of Education (2011) affirms that the skills that should be developed in the English subject are listening, speaking, reading and writing. Nevertheless, according to teachers who work at public schools, the most difficult skill to develop is speaking. It occurs because English is taught as a foreign language (Ministry of Education, 2011) instead of as a second language in Chile. As a consequence, English is not used outside the classroom. Hence, students do not consider speaking activities relevant since they are not going to use this skill in their daily routine.

The Ministry of Education (2011), in the English 10th grade school program, states that speaking is acquired through interactions which are produced in communicative situations which should not be long and complex. Furthermore, The Ministry of Education (2011) adds that speaking consists in communicating in English with a clear pronunciation in order to participate in monologues, oral exchanges and conversations.

In the same way, speaking assessment presents difficulties. According to Coombe, Folse and Hubley (2010), assessing speaking has big challenges as to subjectivity at the moment of assessing, the little time, the number of students and administrative difficulties. The authors manifest that, because of these reasons, some teachers may not assess speaking. These problems are also present at public high schools in Chile since there are many students per class so more time is needed to assess this communicative skill. Besides, there are vulnerable students with low level of English. According to English SIMCE Report (2014), 53.2% of 3rd grade students are below A1 level in the CRFR (Common European Framework of Reference for Languages) and 22.3% are in A1 level. Another problem that exists in assessing speaking is the subjectivity at the moment of grading (Coombe, Folse and Hubley, 2010). However, these authors state that

assessing speaking is crucial since it is an important communicative skill because, every person who wants to learn English, studies it with the aim of being able to speak in English.

In this context, it is necessary to research the following issues: how is speaking assessed at public high schools? and how does it influence on students' performance in speaking? There is not enough information about this problem. Also, it could be explored which procedures, activities, assessment instruments and test specifications are more useful to assess speaking in this context. In addition, it is a viable study as it will be based principally on observation and on interviews analysis. The investigation will be done at a public high school in Curanilahue since the researcher has observed the speaking assessment is not done frequently and teachers apparently do not use assessment instruments. Moreover, there are many vulnerable students and a study has not been done in this context. For this reason, the research will be done at a specific high school in Curanilahue because it is the only public high school in the city and it is in this context where the problem emerged.

2. Context

Assessment is very important in the process of teaching and learning because assessment is part of the learning process in the class (Fulcher and Davidson, 2007). It is a powerful tool for learning so if teachers know how to orient it, it could foster learning in students since it helps teachers to make decisions to improve their way of teaching. That is why teachers assess periodically during the year. As written before, the Ministry of Education (2011) asks teachers to develop the four skills: listening, speaking, reading and writing. Therefore, all these abilities should be assessed.

According to what the researcher has observed during her working experience, teachers do not assess the four skills in an equal way and the skill

which is less assessed is speaking. Luoma (2011) manifests that it is hard to assess speaking because reliability could be affected since the assessment is done face to face between the student and the teacher. Therefore, the teacher has to listen to the student and assess at the same time so the teacher could avoid an issue that may affect the student's mark. Another problem that is present at public high school is the number of learners per class so more time and resources are needed (Coombe, Folse and Hubley, 2010).

3. Research question

The main question of this research is how is speaking evaluated at public high schools and how it influences on students' performance in this communicative ability?

To answer that, the following specific questions should be answered: which procedures and assessment instruments do the teachers of English use to assess speaking in English? Which activities do teachers utilize to assess speaking? Which test specifications do teachers use to assess speaking? How do activities, assessment instruments and test specifications influence on students' performance in speaking?

4. Research assumptions

- 4.1 Teachers do not use rubrics to assess the abilities of the language.
- 4.2 Teachers do not assess communicative aspects of speaking, they assess part of them.
- 4.3 The way teachers assess influences on students' performance.

5. Objectives

5.1 General objective

To analyze the test specifications, instruments and tasks used by teachers of English of a public high school to assess speaking and its effect on secondary students' performance.

5.2 Specific objectives

5.2.1 To identify the instruments and procedures that teachers of English apply to assess speaking.

5.2.2 To characterize the tasks that teachers of English and students declare to use to assess speaking.

5.2.3 To identify the test specifications that teachers of English utilize to assess speaking.

5.2.4 To analyze the relation among tasks, assessment instruments and test specifications with the students' performance in speaking.

6. Significance of the study

In Chile, there are not enough studies about what is happening in the teaching and learning process of English. It could be since it is taught as a foreign language so it is not as important to investigate as in countries where this language is taught as a second language. For this reason, it is relevant to research about how English is being taught and assessed as a foreign language in order to examine what teachers do in classes.

The focus of this research is the speaking assessment at a public high school since it is the most difficult skill to assess due to the students' low level of English (SIMCE Report, 2014), the time that it requires and the number of students per class. In relation to what the researcher has observed, it is a difficult task

because teachers work alone in the classroom and the hours of English are not enough, between three and four per week. Hence, teachers have to deal with the lack of time at the moment of assessing.

The main objective of this investigation is to discover how teachers are assessing speaking and how it impacts on students' performance. Furthermore, it could be concluded which procedures, activities, assessment instruments and test specifications are recommended to assess speaking at a public high school.

7. Viability

This research is viable because it will be based on observation and data analysis. In addition, the study will be done in the high school where the researcher works since it is the only one in the city and the problem was presented in this place. Therefore, problems to implement this research should not emerge.

Otherwise, economical funding would not be needed since the materials required are basic. For example, sheets, pens, printer, among others. Likewise, the participants that will be part of this investigation are accessible.

8. Consequences

The investigation expects to have certain consequences related to speaking assessment. One of the consequences may contribute to discover the frequency of the speaking assessment in the academic year. Also, the research can reveal the principal procedure used is the observation. Another consequence could be to realize that teachers do not use the assessment instruments correctly or that they do not utilize any instrument. Furthermore, to identify that teachers use a variety of activities to assess but that they do not have test specifications defined which impacts directly on students' performance. The last consequence would be to affirm that speaking could be assessed at public high schools.

CHAPTER II

THEORETICAL FRAMEWORK

The following study has as objective to know how speaking is assessed in a public high school and how it impacts on students' performance. Therefore, the theoretical framework will be based on three principal concepts: assessment, second language assessment and speaking assessment. First of all, the concept of assessment is going to be stated.

1. Assessment

In second language teaching, there is a difference among the terms evaluation, assessment and testing. Evaluation is a general concept which collects the information from education like the syllabus objectives, course design, materials and the students (Coombe, Keith and Hubley, 2010). For Gottlieb (2006), evaluation is a wider concept that involves a methodical process which means that it goes along with assessment so it is not isolated.

It can be affirmed that assessment takes part of evaluation since it is related to what learner does. It is the different ways that a teacher uses to gather information on a student's achievement. Assessment also involves the concept of testing (Coombe, Keith and Hubley, 2010). Brown (2003) states that assessment is an ongoing- process that incorporates a domain which is wider. The same author manifests that the assessment can be done at whatever time of the teaching process. Fulcher and Davidson (2007) add that "assessment and learning are integrated in the classroom" (p. 27). Also, they declare that educators are always assessing and that the main objective of assessment is to give information in order to have a better teaching and to achieve a more efficient. Besides, Gottlieb (2006) considers assessment as the process of planning, collecting, reporting and interpreting this information in the teaching and learning process. Moreover, through assessment, teachers collect information from different data sources.

Along these lines, test is the next concept which is going to be explained. Test is a way of assessment which is formal and systematic. The most known test is the paper-and-pencil. This kind of test is also used to get information about students' performance with the objective of measuring (Coombe, Keith and Hubley, 2010). For Brown (2003), a test is "a method of measuring a person's ability knowledge or performance in a given domain" (p. 3). Brown adds that a test is an instrument which has a group of techniques, procedures, or items that a test-taker must perform. Also, this author clarifies that the aim of a test is to measure a general ability, an objective or competence in specific, the student's performance or a domain among other aspects. Gottlieb (2006) manifests that testing gets a student's behavior sample in a specific time. It is done systematically.

Therefore, there is a difference among the concepts of evaluation, assessment and testing in English. Evaluation involves all the concepts since it gets information from the national curriculum hence it is more general. Assessment is in charge of collecting information from the learning process with the aim of improving students' learning. Testing is more formal and the purpose of it is to measure how well learners achieve the objectives of the course or unit.

The concept of assessment considers feedback as a crucial element. In a classroom learning environment, the feedback that students receive is crucial since it helps them to know what they need to learn and how to improve to become an independent user of language in different contexts. For this reason, feedback must have diagnostic information which is not normally obtained in formal tests (Fulcher and Davidson, 2007). Ur (2006) adds that "...feedback, in the context of teaching in general, is the information that is given to the learner about his or her performance of a learning task, usually with the objective of improving this performance" (p.242). Hence, teachers provide feedback to allow students know how well or bad their performance was in order to improve it. The idea is to encourage students' autonomy. Harmer (2003) emphasizes that feedback is given according to the activity required by the teacher and it involves correcting and assessment. The

author explains this stating that “Feedback encompasses not only correcting students, but also offering them an assessment of how well they have done ... The way we assess and correct students will depend ... on the type of activity the students are taking part in” (p. 99). Therefore, the idea of feedback is to let students know how they performed a task and, in occasions, to correct them if there is any mistake.

As this research is about how English is assessed at public high schools in Chile where English is not the first language of the country, it is important to know how to assess a second language. But before explaining that concept, it has to be distinguished the difference between ESL and EFL. ESL is English as a Second Language. Phillipson (2003) explains that ESL is taught in countries where “English is not a native language but where it is used widely as a medium of communication in domains such as education and governments (p.24). The author adds that it is taught at schools. Therefore, people who live in countries which have ESL can communicate in their native language and in English. Otherwise, EFL means English as a Foreign Language and it is also taught at formal education after students learn their native language. However, it is not talked or used outside the classroom (García and García, 2003). The concept of second language assessment will be developed in the following section of the study.

2. Second Language Assessment

The information obtained in assessment made to English Language Learning (ELL) students helps administrators, teachers, students and parents to get to know which the appropriate program is or what the correct instructional activities are for them. Also, it is useful to monitor learners’ progress. Besides, it is harder to assess ELL students than native speakers (O’Malley and Valdez, 1996). In consequence, with the information gathered from the learners through English language assessment, teachers can choose suitable activities for them and it can help teachers to know if the students are learning.

Besides, O'Malley and Valdez (1996) mention that assessment of ELL students has the following purposes: screening and identification, placement, reclassification or exit, monitoring student progress and program evaluation and accountability.

According to Gottlieb (2006), the assessment in English as a second language has four phases: planning, delivery, interpretation of the results and use of the results for feedback (Gottlieb, 2006). In planning, teachers have to decide the academic content, the language proficiency standards and the language aims to assess. Then, they match this with the tasks and instructions. In delivering, teachers define which tasks students will perform. Also, they provide materials with the objective of facilitating students' learning and how learners are going to be grouped (individual, pairs or groups). In interpreting results, teachers can make learners to do peer or self-assessment in order to realize their progress. Also, the author suggests that teachers should separate content and language in order to know their achievements in each area. In the feedback stage, teachers need to give constructive advice. Moreover, it is an interactive process which helps students to be nearer to achieve the aim.

When authors discuss about assessment, they define some qualities or principles of assessment. According to Bachman and Palmer (1997), test usefulness is relevant for the quality of a test because it gives a metric to evaluate the tests that teachers develop and use apart from every aspect of the test development and use. Coombe, Keith and Hubley (2010), also agree that usefulness is the most relevant quality of a test. The test usefulness has six test qualities: reliability, construct validity, authenticity, interactiveness, impact and practicability which have to be balanced when developing a test (Bachman and Palmer, 1997). On the other hand, Coombe, Keith and Hubley (2010) recognize nine principles that a good test must have and they are: usefulness, validity,

reliability, practicality, washback, authenticity, transparency and security. The first term to be developed is going to be reliability.

Reliability is the consistency of a test score even though the testing situation changes. Therefore, if a teacher takes a test to the same group of students on two different circumstances, there should not be any difference (Bachman and Palmer, 1997). According to Coombe, Keith and Hubley (2010), reliability is present in a test when its score is consistent even if the test is taken to the same group in other day and in another environment. Reliability in oral assessment is weak as it is subjective. Hence, to make it more reliable, teachers have to use rubrics and have another person as a rater.

The second quality is validity which is reflected when the test assesses what it wants to assess so the test must reflect what was taught in classes (Coombe, Keith and Hubley. 2010). The authors classify validity in three types: content validity, face validity and construct validity. Content validity refers to measure the contents and goals taught in classes using a format known by the learners. Face validity purports that the test assesses what it should assess. This characteristic gives more credibility to the test. Construct validity is the interpretation that the test takers make on the test score which has to be meaningful and appropriated. However, this interpretation has to be validated through evidence. Therefore, in a language test, the score must evidence the language ability performance which is measured (Bachman and Palmer, 1997). According to Coombe, Keith and Hubley (2010), construct validity is the relation or coherency that must exist between the methodology used by the teacher to teach the language and the kind of assessment.

The third quality is authenticity. According to Bachman and Palmer (1997), authenticity is the degree of correspondence between the characteristics of the target language use and the characteristics of the test task. It is relevant since it connects the test task and the domain to the generalization of the score

interpretations we expect. Another important aspect is that it has a great effect on the takers' point of view of the test. In other words, how the test taker's performance was. For Coombe, Keith and Hubley (2010), authenticity is reflected in tests when they have tasks that can be experiences in real-world situations.

The fourth quality is interactiveness which is related to how the test taker's individual characteristics are involved to achieve a test task. In terms of language testing, the most important characteristics are the test taker's language ability (which can be language knowledge and strategic competence or metacognitive strategies), knowledge and affective schemata (Bachman and Palmer, 1997).

The impact is the fifth quality. For Bachman and Palmer (1997), the impact of the test refers to the effect that it has on the society and educational systems and on the people who are part of both of them. The impact can happen in two levels: micro level (referred to individuals affected by a specific test use) and macro level (referred to educational system or society).

The next quality is practicability. It is related to how the test is implemented. There should be a balance among the test qualities. Moreover, it has been determined for the test the resources to make an operational test and the management of the resources available. Therefore, to have a practical test, it has to be spent few resources which are available. They are the design and development (Bachman and Palmer, 1997). According to Coombe, Keith and Hubley (2010), practicability is an important test characteristic; hence, a test should be "teacher friendly". Some of the features that a test should have are: the reasonable time, resources, how much the test development and maintenance cost, easy to mark, among others.

The following quality is washback which refers to the impact that a test has on teaching and learning. This can be positive or negative. It can be positive if teachers and students notice that testing and curriculum design are created by

course goals which are clear and known by learners and educators. It can be negative since teachers and students can believe that tests are guided by the curriculum and students have to learn and study what is necessary just to pass the test. (Coombe, Keith and Hubley, 2010).

Transparency is the next quality. It consists on saying to students what is going to be assessed and how they will be assessed. Therefore, learners should know the outcomes that are going to be measured, the format that will be utilized, the time that they will need to complete the test, the grading criteria, how much each section or item is going to weigh. Transparency lets students be part of the testing process. (Coombe, Keith and Hubley, 2010).

The last quality is security which is part of reliability and validity. It occurs when a teacher creates a good test which assesses the course goals so this test can be used again (Coombe, Keith and Hubley, 2010).

Another important aspect for assessment is that it should be authentic. According to O'Malley and Valdez (1996), authentic assessment should be coherent with the classroom objectives, curricula and instruction. O'Malley and Valdez states that authentic assessment is used "to describe the multiple forms of assessment that reflect student learning, achievement, motivation, and attitudes on instructionally-relevant classroom activities (1996, p. 4)." The authors mention that some examples of this kind of assessment are performance assessment, portfolios, and student self- assessment.

3. The assessment of speaking

The first concept that should be clarified is speaking since it is the communicative skill that the teacher will assess. When defining the speaking ability, most of the authors revised in this research agree on the communicative component of this skill (Gottlieb, 2006; Harmer, 2003; Luoma, 2011; O' Malley and

Valdez, 1996). O'Malley and Valdez (1996) manifest that the most important responsibility that teachers of ELL students have is to make them to communicate effectively using oral language. Gottlieb (2006) also states that speaking is a communicative action which is done with a purpose. It emphasizes the specific use (language function) and performance. It includes several elements like grammar, vocabulary, intonation and stress, pronunciation, fluency, and accuracy. The author adds that speaking and listening go together because when a person speaks, the other one has to listen in order to get comprehension. Therefore, listening is an input for speaking.

Authors like O'Malley and Valdez (1996) explicit that a proficient speaker has to understand what is said orally. They consider that "the oral communication involves the negotiation of meaning between two or more persons, it is always related to the context in which it occurs" (1996, p. 59). They add that in a classroom, speaking implicate interaction with the teacher and classmates, it depends on how activities are planned.

Harmer (2003) introduces some other elements to oral production and they are: connected speech, expressive devices, lexis and grammar and negotiation of language. Connected speech refers to the speaking ability to produce individual phonemes as well as so students should modify, omit, add or weaken sounds in English. As a consequence, the teacher should give activities where learners can develop and improve their connected speech. On the other hand, expressive devices are the second element: English native speakers use a variety of ways to communicate orally. For instance, they change their pitch, volume, speed and stress certain parts of the speech. Also, to express the mood, they use physical and non-verbal language. To seem natural, students should learn, practice and apply some of these devices. Another element, lexis and grammar which refers to the functions that the person wants to express: natural speaking is formed by various common lexical phrases according to the function that the person wants to express. It is important that teachers teach learners lexical phrases for different

contexts like job interviews, shock, surprise, agreeing, disagreeing, among others. The last element is negotiation of language. To speak effectively, people negotiate language through asking for clarification and what they are saying. Therefore, teachers should teach expressions for asking clarifications in order to make the conversation more understandable and fluent.

In the situation of assessing speaking, teachers have to deal with some issues like the following: the scarce time that they have, activities that are appropriate to use and the suitable criteria for the evaluation (O'Malley and Valdez, 1996). These authors add that the assessment of speaking should be focused on "student's ability to interpret and convey meaning for authentic purposes in interactive contexts. It should include both fluency and accuracy" (p. 61). In other words, teachers have to face different situations when assessing speaking which are to lack of time, the activities which could not reflect what teachers want to assess and the criteria to assess. Moreover, when teachers assess speaking, they have to pay attention if the learner is able to understand what is said and to express what he or she wants to manifest in a context which is authentic and interactive. This action shall be fluent and accurate.

Referring to the requirements for assessing speaking, Fulcher (n.d) points out that a speaking test should have the following components: a task, rating scale and a rater. The task involves something we ask the learners to do so it makes learners talk. The rating scale is used to grade students' performance which is compared with the descriptors that the rating scale has. The rater is frequently the person who speaks with the learner.

At the moment of assessing speaking, authors like Brown (2003), Fulcher and Davidson (2007) and Luoma (2011) emphasize that teachers should have an outline to prepare it. For this reason, the concept of test specifications will be introduced.

3.1 Test Specifications

According to Fulcher and Davidson (2007), test specifications are documents in which teachers find the bases to create the test items, to decide how the test structure is going to be and to plan the materials that are going to be needed. The authors manifest that tests specifications are known as blueprints too which implicates the beginning to build something, like building, for instance. In addition, they state that test specifications are useful to have an equivalent test which means that it should have a similar difficulty and the same aims even though the contents are different so the results may not vary. Hence, the specifications can allow the presence of reliability and validity in a test. Consequently, test specifications are very important when designing a test.

Brown (2003) manifests that test specifications are a test outline which is simple and practical. When a test is built, teachers have to consider that the test specifications are a general draft, where they have to think which skills they are going to assess and how the items would be and which tasks would be used to make students do. Test specifications help educators to organize the topics or objectives they want to assess, the kind of items and the number of them, the time needed for each one, the score that everyone is going to have and the weight in the test.

Luoma (2011) specifies that test specifications should be written at the beginning of a test design. For this author, specifications for speaking assessment indicates what type of speaking will be assessed , how it is going to be assessed and how it is going to be scored. She clarifies that “The specifications records the rationale for why the assessment focuses on certain constructs, and how the tasks and criteria operationalise them” (p. 113). Hence, the test specifications give the reason why the teacher is going to assess an objective and how the activities and the criteria are going to be linked to assess that aim.

Luoma lists the following test specifications (2011, p. 114):

- The test's purpose.
- Description of the examinees.
- Test level
- Definition of construct (theoretical framework for the test)
- Description of suitable language course or textbook
- Number of sections/papers
- Time for each section/paper
- Weighting for each section/paper
- Target language situation
- Text-types
- Text length
- Language skills to be tested
- Language elements to be tested
- Test tasks
- Test methods
- Rubrics
- Criteria for marking
- Descriptions of typical performance at each level
- Description of what candidates at each level can do in the real world
- Sample papers
- Samples of students' performance on task

Luoma (2011) states that all the specifications should be present in one document which can be divided in three modules: "construct specifications, assessment specifications and task specifications" (p. 115) with the aim of focalizing on one conceptual part when designing a test. Construct specifications define the connection that exists between skill and how this one is going be performed and assessed. In construct specifications, it can be found the following specifications: test purpose, description of examinees, test level, definition of

construct, description of suitable language course or textbook, target language situation, text types, text length and test methods. In assessment specifications, the criteria that will be used is defined and how they will be implemented at the moment of assessing. The test specifications present are: rubric, criteria, description of typical performance and description of what candidates can do in real world. Finally, the task specifications describe the task that will be implemented when assessing and the instructions that it will have. In task specifications, the specifications grouped are: number of sections, time for each section, weighting for each section/paper, language skills, language elements, test tasks, sample papers, and samples of students' performance on task. Furthermore, the author manifests that specifications conduct the test development and give information about how score should be used and validate. Moreover, they allow to link the teaching and learning in the assessment. She emphasizes that specifications should be written since they can be "a coherent system whose parts fit together" (p. 115). In other words, they can help the teacher to organize better the specifications when developing a test in the future. Also, they can help to connect the theory of assessment and the practice of speaking assessment in a specific context.

As Fulcher (2017) affirmed, rating scales are one of the important elements when assessing speaking. Thereupon, scoring instruments are going to be defined in the next part of the study.

3.2 Scoring Instruments

The assessment instrument is a tool used to get information to know if the learner achieves every outcome in a class, unit or year (Gallaudet University, n.d). According to Luoma (2011), assessment instruments are used to record how well students performed an activity. Specifically, for speaking, Gottlieb (2006) states that observation or recording is used to document students' oral production and interpret it. To do this, teachers need to use rubrics, rating scales and checklists.

The author defines rubric as the instrument which contains the goals that teachers want to assess and the description of how well learners achieve them. For Fulcher and Davidson (2007), rubric is the instrument which locates students' performance in a specific point on a rating scale with the help of the descriptors.

Gottlieb (2006) suggests that four types of rubrics exist: checklist, rating scales, analytic scales and holistic scales. Checklists are scales of two options in which teachers can mark if skills, competencies, strategies or language functions are present or not. They do not have descriptors.

Rating scales indicate the degree of how target skills, competences, strategies, or language functions are performed. They can have a range of frequency or quality which represents the competence or skill that the learner has at the moment of assessment (Gottlieb, 2006). Analytic scales have the format of a matrix which contains the description of specific dimensions or characteristics of the contrast that is going to be measured. These are known as descriptors. They describe the student's performance at every level of language proficiency. Therefore, they should be clear, observable and measurable. Usually, analytic scales have between four to six performance levels (Gottlieb, 2006). Holistic scales give a general description of a student's competence by performance level. There are various descriptors since the least to most proficient. This scale provides a summary score which is easy to communicate (Gottlieb, 2006).

All the information obtained using these instruments could be transformed to grades. Marks or grades are going to be understood as a way to assess learners' performance and learning individually. It is usually done by teachers. They could be presented as letters, percentages, and numbers, among others (Concordia University, 2013). O'Malley and Valdez (1996) suggest that the criterion levels of oral language proficiency should be based on the classroom objectives.

As the speaking skill is assessed through the development of tasks by students, it is relevant for this research to go in depth in this concept.

3.3 Tasks

Bachman and Palmer define this concept as "...a language use task as an activity that involves individuals in using language for the purpose of achieving a particular goal or objective in a particular situation" (1997, p. 44). For Ellis (2009), tasks are samples of language that give the opportunity to use the language in order to have evidence of students' interlanguages development and also to communicate in a fluent and effective way. Therefore, a task is an activity where students have to use the language in order to achieve an objective in a specific situation. Moreover, tasks have some characteristics which are "... the setting, the test rubric, the input, the expected response, and the relationship between input and response." (Bachman and Palmer, 1997, p. 59). In the same way, Gottlieb (2006) makes a difference between activity and task. For the author, an activity is a single or group of questions that are related. They have a learning objective. Otherwise, a task contains a combination of two or more activities. In this, the cognitive engagement levels are higher.

For O'Malley and Valdez (1996), speaking activities should give the students the opportunities to use the language for authentic purposes, in other words, to give the chance to exchange information with a purpose instead of providing known information to the listener. The authors add that teachers have to consider the following element when assessing: the assessment purpose, if it is going to be individual, in pairs or groups; students' level of English, language functions used commonly in the class activities; and how student are prepared for that activity.

Apart from the characteristics above, Ellis (2009) propones the following dimensions:

- Scope: the task is focalized on eliciting the language use and on completing tasks that involve language. Moreover, it presents a difference in the meaning of exercise and task. On one hand, the “exercise” involves activities which require form-focused language used. On the other hand, the “task” has activities which involve meaning focused language used. Nevertheless, both have the same purpose which is to learn a language.
- Authenticity: it refers to an activity that reflects the real world. However, there are some activities that do not represent completely the real world. For example: identifying the differences in two pictures or describing a picture.
- Language skill: even though skills like reading and writing are involved in the activities, the most used skill is speaking since they are performed orally.
- Cognitive processes: doing a task implicates developing cognitive processes, for example “selecting, reasoning, classifying, sequencing information, and transforming information from one form of representation to another” (Ellis, 2009, p. 7). Therefore, there should be a relationship between the cognitive processes that students need and the language that wants to be developed in the tasks.
- Outcomes: all tasks have a clear objective that could be related to language or content. It is important to realize that there is a difference between an outcome and an aim. The “outcome” is related to what the learner achieves when he or she finishes the task. The “aim” is related to the teaching purpose of the task, in other words, how the language is used, if it is receptive or productive.

For assessing speaking in class, there are several activities in which students can demonstrate their speaking performance. Harmer (2003) names the

following: acting from script, communication games, discussion, prepared talks, questionnaires and simulation and role play.

To sum up, the assessment of speaking should have three components: a task, a scoring instrument and a person to assess. The task gives students the opportunity to demonstrate their speaking ability exchanging information with a listener. The scoring instruments are used to obtain information to know how well the students achieve the speaking objective and they are: checklist, rating scales and rubrics. The person who gets that information is the teacher who observes students' performance. Otherwise, to assess speaking, teachers should have an outline. For that reason, they should use the test specifications which indicate what is going to be assessed and how is going to be assessed.

CHAPTER III
METHODOLOGY

1. Research Paradigm

The paradigm that this research adopted is the interpretative qualitative since the aim of it is to interpret all the information collected from a group of participants which is limited. Furthermore, to describe and analyze profoundly of single units from natural settings.

Phothongsunan (2010) claims that the interpretative paradigm researches how people understand the world. The author adds that the person who investigates interprets the data collected from a small group of participants since his or her objective is “to explore the meanings which participants place on the social situation under investigation” (p.2). In the same way, Nguyen and Tran Thi Le (2015) manifest that the interpretative paradigm allows the investigator to see a reality using the participants’ perception which he or she uses to build his or her own understanding. Moreover, Cohen, Manion and Morrison (2007) state that the aim of the interpretative paradigm is “to understand the subjective world of human experience” (p. 21). In addition, these authors establish that the researchers who utilize this paradigm study individual with the purpose of knowing in order to later comprehend their points of view of what is happening around them.

In relation to the qualitative method, Thomas (2010) announces that this method allows investigators to study an issue in a social and cultural way. Also, the author declares the qualitative research tries to investigate the routine of certain groups of people in an environment which has to be as natural as it can be. Likewise, the author adds that the qualitative method uses observation, interviews and questionnaires, documents and texts, and the researcher's impressions and reactions as the researcher is in charge of gathering the information and interpreting it. Similarly, Burns (2003) claims that qualitative research provides “descriptions, interpretations and clarifications of naturalistic social contexts” (p. 23).

The aim of this research is to analyze the test specifications, instruments and tasks used by teachers of English of a public high school to assess speaking and its effect on secondary students' performance. For that reason, the qualitative data collection method was adopted, even though some quantitative data will be collected to enrich the analysis, and, as the researcher has to interpret the data gathered, the interpretative approach was used. In order to collect the data, two semi structure interviews, a check list and a rating scale were applied to teachers and students.

2. Type of Research

The following research is categorized as exploratory and descriptive because it studied a problem that was not investigated before in that context and it described and interpreted what it was observed.

According to Sahu and Singh (2016) an exploratory research is used when a problem has not been studied or an issue is not clear. This kind of research makes the researcher be more familiar with what is being studied and develops a hypothesis to prove. Jupp (2006) also states that the exploratory research is methodological whose aim is to discover and generate theory.

Kumar (2007) affirms that the descriptive study "describes and interprets what is" (p. 229). Moreover, the author states that this kind of research studies the conditions that are present, the opinions, what is occurring, the effects that exist in a certain moment. Knupfer and McLellan (1996) add that descriptive research is useful for quantitative and qualitative research since the researcher can collect information in order to describe an issue which is organized, tabulated and described.

Hence, the research was exploratory because it expected to study the test specifications, instruments and tasks used by teachers of English of a public high school to assess speaking and its effect on secondary students' performance and

this problem was not investigated before in a high school in Curanilahue. Besides, this study wanted to explain what teachers did at the moment of assessing speaking and how this influenced on their students' performance.

3. Research Design

The research is a case study since it was applied to teachers of English and to their students at a high school in Curanilahue so all the data was collected in a real context.

Cohen, Manion and Morrison (2007) affirm that a case study consists of giving information of how real people behave in natural circumstances which allow readers to comprehend the information in a clearer way. Moreover, the case study makes easier to understand the relation that exists between the ideas and abstract theory. Zainal (2007) proclaims that a case study permits the investigator to study the information in a precise environment. In addition, the author declares that the case study needs a small number of people or a limited geographical area. Besides, the issue to study should be contemporary and real. For Gerring (2004), a case study is to study profoundly one fact in order to get a general conclusion through analyzing various facts.

Since the research expects to know the test specifications, instruments and tasks that teachers use and their influence on students' performance, two teachers from a specific high school and their students were studied. The idea was to get a general conclusion from what was observed. Therefore, the information obtained was real and from an authentic environment. The group studied is a sample from the high school.

4. Research Context

The research was conducted at a high school in Curanilahue, Bio Bio Region, Chile. It is a multipurpose high school and it considers three development areas: Scientific-Humanistic, Technical-Professional and Artistic. This is a public high school which was founded in 1964. The enrollment is of 1200 students whose vulnerable rate is about 62%. It has grades from 1st and 4th in secondary school with eight grades per level. The grades are mixed and have between 35 and 42 learners.

The high school follows the programs given by the Chilean Ministry of Education. In relation to the English subject, it is taught as a Foreign Language (EFL). In 1st and 2nd level, there are four hours per week and in 3rd and 4th level, there are three hours per week. In classes, the students develop the four language skills (reading, speaking, listening and writing). The most developed skills are reading and listening. Also, grammar and vocabulary are taught. The class is organized in three stages: beginning, development and end. In terms of language used, the classes are 50%-50%, half in Spanish and half in English. The teachers of English follow the Ministry program and they prepare the material needed. Sometimes, they use the course books provided by the government.

5. Participants

All the participants in the research belong to the same public high school. They were two teachers of English and twelve students, six learners per teacher. One teacher (Teacher 1) has 34 years of experience and the other teacher (teacher 2) has 20 years of experience. In relation to students, one group belonged to 2nd grade, whose age was between 15 and 16 years old, and the other group belonged to 11th grade, whose age was between 16 and 17 years old. The 10th grade was Scientific-Humanistic and their level of English is elementary and the

11th grade was Technical-Professional with, electronic specifically, and their level of English is break-through.

McDonough, J. & McDonough S. (2005) manifest that a group interview should have between three and four participants because it is not necessary to have the whole group of class. In this study, the number of participants was limited to 14 (two teachers and twelve students). The teachers and the students were selected using the non-probability sample. This kind of sample allows the researcher to select the participants for the study from a group (Cohen, Manion and Morrison, 2007).

Table III.1 Criteria to select the participant students.

Criteria of selection
Students from the participant teachers' public high school
Student from a participant teacher

Table III. 2 Criteria to select the participant teachers.

Criteria of selection
Teacher of English
Working at a public high school
Giving English lessons

6. Instruments and procedures

The data collection forms in qualitative research are data collecting, observation and document analysis. For this study, two semi structured interviews (see appendix 4 and 5), a checklist (see appendix 2) and a rating scale were utilized (see appendix 3).

The first instrument was the semi structured interview. This type of interview has open-ended questions that can be asked to a person or to a group of people. These questions could be explored into detail by the interviewer (Jamshed, 2014). Merriam (2009) adds that the semi structured interview has a list of questions. However, if the researcher requires it, he or she can exchange the order or change the words. These interviews were conducted in Spanish. One semi-structured interview was applied to the teachers and one to a sample of students. The teachers' interview (instrument 1) had the aim of discovering how teachers assess speaking. It had 17 open questions which were answered orally and recorded. The other semi structured interview was applied to the students (instrument 2). Its objective was to discover the students' perception about how teachers assess speaking in class. It considers 8 open questions.

Instrument 3 was a checklist. According to Gottlieb (2006), checklists are scales which have two options in which it is marked if skills, competencies, strategies or language functions are present or not. They do not provide descriptions. The aim of the checklist was to determine how the teacher assesses speaking and how it influences on students' performance. It had 10 aspects to be observed during a class which were related to the activity, assessing instrument, assessed criteria and marks. These aspects were defined based on the theory.

Instrument 4 was a rating scale. Gottlieb (2006) affirms that rating scales indicate the degree of how target skills, competences, strategies, or language functions are performed. In order to represent the competence or skill that is present at the moment of assessment, the rating scale as a range of frequency or quality. The objective of this instrument was to identify the test specifications that teachers use to assess speaking. It had 21 aspects taken from Luoma's test specifications (2011) and the descriptors were: adequate, partially adequate, minimally adequate, inadequate and not observed. The observation was made in an English class.

6.1 Categories

In this research, the categories to analyze were presented in the instruments. They were pre-determined categories and they were presented in the interview, checklist and in the rating scale. The rating scale categories were specifically taken from Luoma's test specifications (2011).

6.1.1 Categories to identify the instruments and procedures that teachers of English apply to assess speaking (Instrument 3).

- **Assessment activities:** These activities allow the student to use the language in order to show his or her interlanguages development and also to communicate in a fluent and effective way (Ellis, 2009).
- **Assessment instrument:** it is a tool used to get information to know if the learner achieves every outcome in a class, unit or year (Gallaudet University, 2016).
- **Assessment Criteria:** The aspects assess. Moreover, criterion levels of oral language proficiency should be based on the classroom objectives (O'Malley and Valdez, 1996).
- **Assessment techniques:** Techniques used to have a valid picture of students' performance in a speaking assessment (Coombe, Folse, Hubley, 2010).
- **Marks ranges:** A way to assess learners' performance and individual learning. It is usually done by teachers. They could be presented as letters, percentages, and numbers, among others (Concordia University, 2013).

6.1.2 Categories to characterize the tasks that teachers of English and students declare to use to assess speaking (Instrument 1 and 2)

Teachers' interview

- Purpose to assess oral production: why teachers assess students (Gottlieb, 2006).
- Frequency of assessment: number of times a teacher has assessments planned in the year (Gottlieb, 2006).
- Dimensions assessed: the speaking activities and instruments used (Coombe, Folse, Hubley, 2010).
- Scoring instruments: it is a document that teachers use to document students' oral production through observation or recording in order to interpret it (Gottlieb, 2006).
- Factors that affect oral assessment: the aspects that can interfere in speaking exams (Coombe, Folse, Hubley, 2010).

Students' interview

- Frequency: the assessments planned in the year (Gottlieb, 2006).
- Types of evaluative tasks: activities which allow students to use the language with the aim of reaching a specific objective in a certain situation (Bachman and Palme, 1997).
- Scoring instruments: it is a document that teachers use to document students' oral production through observation or recording in order to interpret it (Gottlieb, 2006).
- Dimensions of assessment: the criteria to assess which could be: grammar, vocabulary, intonation and stress, pronunciation, fluency and accuracy (Coombe, Folse, Hubley, 2010).
- Feedback: information provided to students about how he or she performs in a task with the aim of improving the student's performance (Ur, 2006).

6.1.3 Categories to identify the test specifications that teachers of English utilize to assess speaking (Instrument 4).

Pre-determined category by Luoma (2011)

The following categories were taken from the test specifications mentioned by Luoma. The author affirms that specifications for speaking assessment indicates in what type of speaking will be assessed , how it is going to be assessed and how it is going to be scored (2011). The specifications are:

- The test's purpose.
- Description of the examinees.
- Test level
- Definition of construct (theoretical framework for the test)
- Description of suitable language course or textbook
- Number of sections/papers
- Time for each section/paper
- Weighting for each section/paper
- Target language situation
- Text-types
- Text length
- Language skills to be tested
- Language elements to be tested
- Test tasks
- Test methods
- Rubrics
- Criteria for marking
- Descriptions of typical performance at each level
- Description of what candidates at each level can do in the real world
- Sample papers

- Samples of students' performance on task

7. Data Analysis procedure

Before applying the instruments, they were revised and validated by four university professors specialized in Education.

To analyze category 1: to identify the instruments and procedures that teachers of English apply to assess speaking, the researcher went to the classroom to observe the teachers using the checklist. In the class, the teachers were assessing speaking. In second category 2: to characterize the tasks that teachers of English and students declare to use to assess speaking, the researcher interviewed to the participant teachers and participant students individually and recorded them. After that, the researcher transcribed the answers which were written in a chart to compare them. Finally, to study category 3: to identify the test specifications that teachers of English utilize to assess speaking, the researcher observed a class where the teachers were assessing speaking. The researcher used a rating scale based on Luoma's test specifications (2011). After having all the information required, it was separated per category and compared. Then, it was taken the information which was similar and the information that was different to the rest.

CHAPTER IV

DATA ANALYSIS AND INTERPRETATIONS

In this chapter, the analysis of data collected will be carried out and this analysis will be presented by each specific objective in order to analyse the results in detail. To initiate this analysis, specific objective 1 will be studied to identify the instruments and procedures that teachers of English apply to assess speaking.

1. Specific objective 1: To identify the instruments and procedures that teachers of English apply to assess speaking.

With the purpose of knowing the instruments and procedures that teachers use during the speaking assessment, they were observed and the information was collected through a checklist. This instrument includes 10 different items in order to know the tasks, instruments and criteria that teachers applied to assess speaking. The numbers in the checklist represent if the aspects were present (1) or not (0). The idea is to equal the score and, in this way, identify the teacher who has more aspects of the checklist present.

1.1 Category 1: Assessment activities

To start with the analysis, the *assessment activities* will be commented. These activities allow the student to use the language in order to show his or her interlanguages development and also to communicate in a fluent and effective way (Ellis, 2009).

Table IV.1: Assessment activities

Aspects to assess	TEACHER 1			TEACHER 2		
	YES (1)	NO (0)	Observation	YES (1)	NO (0)	Observation
1. S/he develops one or more defined assessing activities.	1		Presentation about different natural disasters	1		Dialogue
2. The activities developed in		0			0	Pronunciation Memory

classes are different among them.						
---	--	--	--	--	--	--

It is observed that teachers use one activity to assess the task. In relation to the activities, T1 uses an oral report about natural disasters and students have to describe a picture, and T2 utilizes a dialogue in which friends are talking about a party. Both activities were taken from the English book provided by the Ministry of Education without including any adjustment.

From the observation, it can be stated that the teachers use activities to promote oral production. Even though, they are not enough to assess the student's speaking ability. Coombe, False and Hubley (2010) explain that teachers should offer the students different assessment activities to demonstrate their speaking ability. As it was mentioned before, T1 includes an oral presentation as an assessment. According to Harmer (2003), the oral presentation belongs to prepared talks since students have to investigate about a topic and they present the information they found in front of the class. The other activity, the dialogue, prepared by T2, is considered by the author as a possibility of acting from a script. However, in this case, the only requirement the teacher made was to memorize the dialogue without any acting. Therefore, we may conclude that this is not a communicative activity because it includes only memorization.

1.2 Category 2: Assessment instrument

The next aspect observed is the *assessment instruments* that teachers used.

The assessment instrument is a tool used to get information to know if the learner achieves every outcome in a class, unit or year (Gallaudet University, n.d).

Table IV. 2: Assessment instrument

Aspects to assess	TEACHER 1			TEACHER 2		
	YES (1)	NO (0)	Observation	YES (1)	NO (0)	Observation
3. S/he uses assessment instruments:						
a) Checklist.		0			0	
b) Rating Scale.	1		Taken from the 3rd grade course book.		0	
c) Rubric.		0		1		Taken from the 2nd grade course book.
d) Others.		0			0	

Referring to the assessment instruments that teachers employed, they were a rating scale, T1, and an analytic rubric, T2. A rating scale indicates the degree of how target skills, competences, strategies, or language functions are performed which is represented with a range of frequency or quality which shows the competence or skill that the learner has at the moment of assessment (Gottlieb, 2006). A rubric is the instrument which locates students' performance in a specific point on a rating scale with the help of the descriptors (Fulcher and Davidson, 2007).

Both instruments were taken from the English book provided by the Chilean Ministry of Education. These assessment instruments were not adapted to the assessment task. Hence, there were some missing criteria. For example, T1 did not assess the completion of the task and pronunciation, these aspects are considered relevant in speaking assessment task (Luoma, 2011) and T2 did not include memory in the rubric but she assesses it in her copybook. To continue with the analysis, the criteria that teachers use are presented.

1.3 Category 3: Assessment criteria

Table IV.3: Assessment Criteria

Aspects to assess	TEACHER 1			TEACHER 2		
	YES (1)	NO (0)	Observation	YES (1)	NO (0)	Observation
4. S/he uses assessing criteria.						
a) Pronunciation.		0		1		
b) Fluency.	1			1		
c) Grammar.	1			1		
d) Linguistics skills		0			0	
e) Contents or ideas		0			0	
f) Other criteria.	1		Vocabulary	1		Interaction, task and memory. However, "memory" is not present in the rubric. The teacher takes notes on her copybook and then she adds the score below the criteria in the rubric.

As mentioned before, the assessment instruments were not adapted to the task and it could be observed from the criteria the teachers used to assess the students. O'Malley and Valdez (1996) suggest that the criterion levels of oral language proficiency should be based on the classroom objectives. T1 assessed fluency, grammar and vocabulary but there was not assessment of the achievement of the task. Otherwise, T2 assessed interaction, pronunciation, fluency, grammar, task and memorization. However, memorization was not

included in the rubric; it was added to the student's score in the teacher's copybook. To continue with the analysis, the assessment arrangement will be presented.

1.4 Category 4: Assessment techniques

In this category, it will be studied when and how the teacher organizes the classroom and the students for the assessment.

Table IV.4: Assessment techniques.

Aspects to assess	TEACHER 1			TEACHER 2		
	YES (1)	NO (0)	Observation	YES (1)	NO (0)	Observation
5. S/he assesses in the English class.	1			1		
6. S/he assesses individually.	1				0	
7. S/he assess in group.		0		1		But the mark is individual
8. S/he uses materials as a support for assessing.	1		Pictures		0	
9. Just the teacher assesses.	1			1		

From the observation, it can be deduced that the teachers assess in the English class and even if the activity is individual or in groups, the mark is given individually. It could be because the pronunciation and fluency performance might vary between the learners so it would be unfair to give the same score to all students. Besides, the teacher is the only person in charge of grading. This could

be since the teacher thinks that he or she is the only who has the experience and competences to assess. Coombe, Folse and Hubley (2010) mentioned that speaking has low reliability or it is subjective so the authors suggest that there should be two or more raters in order to discuss about the student's performance and agree with his or her mark. In relation to the use of supporting materials, T1 used pictures and T2 did not use any. Maybe, it happened because T1 assessed a presentation about natural disasters so students needed pictures to explain better the topic. Nevertheless, T2 did not use any supporting material since she assessed a dialogue taken from the English course book. Hence, the use of materials depends on the kind of activity.

1.5 Category 5: Mark ranges

In the next section, it would be studied the impact that the way of assessing has on students' marks. Marks or grades are going to be understood as a way to assess learners' performance and learning individually. It is usually done by teachers. They could be presented as letters, percentages, and numbers, among others (Concordia University, 2013).

Table IV.5: Marks ranges

	TEACHER 1		TEACHER 2	
Aspects to assess	Nº of students		Nº of students	
10. Most of the marks are in the following range.				
a) 6,5 al 7,0	2	12,5%	31	79,5%
b) 6,0 al 6,4	2	12,5%	5	12,8%
c) 5,5 al 5,9	1	6,3%	0	0%
d) 5,0 al 5,4	6	37,4%	2	5,1%
e) 4,5 al 4,9	1	6,3%	0	0%
f) 4,0 al 4,4	2	12,5%	1	2,6%

g) 1,0 al 3,9	2	12.5%	0	
Total	16		39	

According to table IV.5, it can be stated that the marks vary according to the teacher who assesses and the task assessed. In the case of T1, most of the marks are in the range of 5,0 - 5,4 (6 students) and 11 students are over 5.0, that is between 5.0 and 7.0 and 2 students are below the passing mark. Most of the marks of T2's students range from 6,5 to 7,0 (31). It is interesting to notice that there are no students below the passing mark. This difference might happen because T1's activity involved more communicative aspects and it was more demanding in terms of linguistics abilities and T2's activity consisted in learning a dialogue by heart and it was taken from the course book so the students had the opportunity to rehearse before the assessment. Another reason could be that T1's students have 2 hours of English per week and T2's students have 4 so T1's learner has less opportunities to practice. Besides, T2 gives a whole class to practice. Moreover, the rater could influence since one could put more emphasis on one aspect and the other not. T1 emphasizes the communicative aspects and T2 emphasizes memory. Coombe, Folse and Hubley (2010) mention that teacher could use a scoring sheet with clear criteria to make the process of grading easier. As a consequence, it can be stated that the teacher, activity, scoring instrument and criteria to assess influence on student's mark.

2. Specific objective 2: To characterize the tasks that teachers of English and students declare to use to assess speaking.

In relation to know which tasks teachers declare to use, a semi-structured interview was applied to teachers and a sample of their students. For that reason, seventeen questions for teachers and eight for learners were designed. These were in their mother tongue, Spanish, as the purpose was to know students and teachers' points of view instead of measuring their level of English.

The analysis for teachers' interview is separated into two cases (Teacher 1 and Teacher 2) which are described in the methodology chapter on pages 28 and 29. The analysis will be carried out by each categories of each semi-structured interview. The categories of the teachers' interview are: *purpose to assess oral production, frequency of assessment, dimensions assessed, scoring instruments used, factors that affect oral assessment and assessment of speaking v/s assessment of abilities of the language.*

The students' interview categories are: *frequency, types of tasks and performance, scoring instruments, dimensions of assessment, and feedback.* Each analysis will be illustrated through tables below. The analysis will start with the teachers' view about why to assess speaking.

Category 1: Purpose to assess oral production

This category is related to the importance that the oral production has for the participant teachers. Also, it is going to be analyzed if speaking has the same weighting in comparison to the other abilities (reading, listening and writing). According to Gottlieb (2006), the purpose of an assessment represents why teachers assess students and there are five reasons: identification and placement to determine eligibility for support services, monitoring progress of English language proficiency and academic achievement, accountability for English language proficiency and academic achievement, reclassification within or transition from support services and program evaluation to ascertain effectiveness of support services. The evidence of teachers' opinion is presented in the table below.

Table IV.6: Purpose to assess oral production

Questions	Teacher 1	Teacher 2
¿Evalúa la expresión oral en clases?	<i>Sí</i>	<i>Sí</i>
¿Por qué evalúa la expresión oral?	<i>... nuestra asignatura, inglés eemmm...debe ser impartida o enseñada o tratar de que los alumnos</i>	<i>Porque es una de las cuatro habilidades que tenemos que evaluar en la enseñanza del idioma</i>

	<i>logren eh...habilidades y una de esas habilidades es expresarse en forma oral.</i>	<i>inglés.</i>
¿Qué peso tiene la evaluación de la expresión oral en relación a las otras habilidades comunicativas?	<i>Harto porque los idiomas se hablan, después se escriben, así que yo creo que es harto, harto el peso que tiene.</i>	<i>Tiene la misma, el mismo, la misma equivalencia. Es igual que el resto. Porque es una de las cuatro habilidades y cada una tiene la misma validez.</i>

After analyzing the teachers' answers, it can be stated that oral production does not have the same weighting for them since T 1 declares that speaking is the most important skill as languages are spoken first and then comes written production. T2 manifests that speaking is as important and as valid as other skills. Therefore, for T2 oral productions should have the same weighting as the other abilities.

From the evidence, it can be stated that both of them declare to assess speaking. However, they differ in the purpose for assessing this skill. T1 refers to the communicative aspect of speaking because she declares that teachers have to make students develop the abilities to achieve communication in the second language. Nevertheless, T2 manifests that she assesses speaking because she has to. This may mean that she does not have or does not declare her own purpose for assessing ability but instead, she follows the requirement of the Ministry of Education in Chile. The Ministry of Education (2015) states that the assessment should be based on the expected learnings, which are descriptors of what a student must know or must be able to do in a certain time, in order to know how well the learner achieves it. One of the expected learnings is speaking. For that reason, speaking ought to be assessed. Besides, the purpose of this ability is to achieve that students can communicate orally. In addition, it can be stated that it does not have the same weighting for them since T1 declares that speaking is the most important skill as languages are spoken and then written. Otherwise, T2 manifests that speaking is equal to the other skills. Therefore, it can be concluded

that speaking has at least the same weighting than the other abilities or is the most important among them.

In the next category, it will be studied the assessment frequency.

Category 2: Frequency of assessment

The purpose of this category is to know how many times teachers assess each ability in a year. According to Gottlieb (2006), teachers should plan the assessments of the year. The aim of this analysis is to get to know if the abilities are assessed in a balanced way or some of them predominate. The information below refers to this category.

Table IV.7: Frequency of assessment

Question	Teacher 1	Teacher 2
¿Con qué frecuencia evalúa la comprensión de lectura en el año?	<i>... con más frecuencia diría yo que la expresión oral ... depende del grupo curso en realidad. Hay cursos que tienen facilidades para comprender el idioma, otros que tienen más facilidades para expresarse en forma oral y creo que lo que se nos hace más fácil evaluar es la comprensión de lectura... evaluó unas dos veces por unidad la comprensión de lectura... Cuatro, en cuarto.</i>	<i>En cada unidad. Cuatro.</i>
¿Con qué frecuencia evalúa la expresión oral?	<i>Una vez por cada unidad, o sea, cuatro veces</i>	<i>En cada unidad igual.</i>
¿Con qué frecuencia evalúa la expresión escrita durante el año?	<i>... cada vez la evaluó menos. Me enfoco más en la comprensión lectora, en la comprensión de lectura y emmm ... en la expresión oral. Eso es lo que hago, sí. También... eeehh. Escrito pero muy poco... Una vez por</i>	<i>Igual.</i>

	<i>semestre por lo menos, una vez por semestre.</i>	
¿Con qué frecuencia evalúa la comprensión auditiva en el año?	<i>... menos veces por muchas razones.... Igual trato de evaluarla una vez por unidad. Cuatro, pero ehhh creo que entre la expresión oral y la evaluación de la habilidad de la audición, andan por ahí en realidad.</i>	<i>También en cada unidad.</i>

According to the answers given by the teachers, both of them assess reading and listening comprehension, writing and speaking. According to the Chilean curriculum, four units should be taught in secondary school where the four communicative skills should be developed (Ministry of Education, 2015). In relation to reading comprehension, T1 assesses this skill twice per unit. However, T2 assesses this ability once per unit. In speaking, both teachers manifested that they assess it once per unit likewise reading comprehension. Nevertheless, the teachers differ in the frequency of assessing writing since T1 affirmed that she assesses writing twice a year and T2 stated that she assesses this ability once per unit. However, the Ministry of Education (2015) suggests that the assessments in a unit should be done according to the expected learnings which are related to each skill, in other words, there is one expected learning per skill. For this reason, the four skills should be assessed once per unit at least. Hence, it can be interpreted that the teachers assess all the skills in a unit in general even though T1 affirmed that she assesses writing twice a year.

The following category is related to dimensions assessed.

Category 3: Dimensions assessed

In this category, it is going to be analyzed what teachers take in consideration when assessing speaking, which instruments, activities and criteria to assess use. A rating scale indicates the degree of how target skills,

competences, strategies, or language functions are performed which is represented with a range of frequency or quality which shows the competence or skill that the learner has at the moment of assessment (Gottlieb, 2006). A rubric is the instrument which locates students' performance in a specific point on a rating scale with the help of the descriptors (Fulcher and Davidson, 2007). These instruments contain criteria and Luoma (2011) mentions that a speaking assessment instrument should have between four and six criteria. Otherwise, Luoma (2011) does not consider the concept of activity; she considers it as a task. The author defines specifically the speaking tasks as activities where learners use the language in order to achieve a specific objective in a certain speaking situation.

Table IV.8: Dimensions assessed

Question	Teacher 1	Teacher 2
¿Cómo evalúa la expresión oral?	<i>Primero, practicamos con los chicos lo que corresponda a la unidad que estamos viendo, le voy explicando las cosas que son importantes, la entonación que deben usar en tal o cual oración. Intento hacer una rúbrica, ... A veces, también he pedido cooperación, ellos hacen una coevaluación con una rúbrica puede llamarse también eso, qué, cuántos puntos de vamos a poner a esto, qué sé yo a la entonación del compañero, a los contenidos que usó, a la fluidez que usó, etcétera.</i>	<i>Mediante rúbricas.</i>
Mencione algunos ejemplos de la evaluación de expresión oral	<i>La presentación oral ...</i>	<i>Evalúo en diálogos, en presentaciones.</i>
¿Qué aspectos o criterios	<i>Los contenidos, la</i>	<i>El dominio del contenido,</i>

evalúa en la expresión oral?	<i>entonación, la pronunciación, no tanto la fluidez en realidad.</i>	<i>la pronunciación ... que mencionen ciertas expresiones que se están estudiando... la pronunciación ya la dije, en algunos caso memorización</i>
------------------------------	---	--

From the analysis of the teachers' answers, it can be deduced that both teachers use rubrics. T1 describes how she constructs the rubric and she expresses that she asks for help to the students in some occasions and asks them to assess their classmates. In relation to the activities, both teachers declare to assess oral presentations which belong to a speaking task because students can use the oral language in order to a show how well they achieved the specific aim. However, one teacher adds that she assesses dialogues too. Besides, the criteria they assess is somewhat similar, both of them assess content and pronunciation. The difference between them is that T1, apart from the criteria mentioned before, assesses intonation, and T2 assesses memorization. As a result, it can be stated that the teachers utilize rubrics at the moment of assessing speaking, that they have some common criteria like content and pronunciation and that the activity that they most use is oral presentation. Therefore, the teachers agree with Coombe, Folse and Hubley (2010) who establish that the teacher has to choose which criteria to assess. Nevertheless, the teachers do not include the criteria that the authors recommend which are: accuracy, vocabulary, linguistic ability, fluency, content or ideas.

To continue, the scoring instrument used will be presented.

Category 4: Scoring instruments used

In this category, it will be seen which rubrics, checklist or rating scales teachers use, the criteria and their weighting, the advantages and disadvantages of scoring instruments according to the teachers. To clarify, checklists are scales of two options in which teachers can mark if skills, competencies, strategies or

language functions are present or not and they do not detailed description (Gottlieb, 2006).

Table IV.9: Scoring instruments used

Question	Teacher 1	Teacher 2
Ordene los criterios que utiliza para evaluar la expresión oral, antes mencionados, del más importante al menos importante.	<i>el uso del contenido ... Luego la pronunciación y finalmente la entonación</i>	<i>Dominio del contenido ... pronunciación, ... memorización y las expresiones vistas en clases.</i>
¿Cómo asigna puntaje a cada uno de los criterios mencionados? Explique por qué.	<i>Con notas generalmente ... Logrado, no logrado. El contenido tendría que estar entre el seis y el siete, que para mí es lo más importante. Y luego ya iría ... descendiendo con los puntajes</i>	<i>De mayor a menor. Cinco puntos. Un punto.</i>
¿Utiliza rúbricas? ¿Cuáles utiliza, holísticas o analíticas?	<i>Pocas... Holística.</i>	<i>Sí. Mayor cantidad, analíticas.</i>
¿Qué beneficios o ventajas reconoce en la utilización de las rúbricas para evaluar la expresión oral?	<i>Bueno, más claro para los alumnos, a qué cosas tienen que ponerles más atención porque se supone que uno la conversa con ellos antes y para uno también. Da más claridad y objetividad, claro, en el momento de calificar.</i>	<i>El poder constatar el dominio de los contenidos, de las expresiones o de lo que se está evaluando.</i>
¿Qué desventajas tiene la utilización de las rúbricas para evaluar la expresión oral?	<i>... se demora mucho en hacerla</i>	<i>Podría ser que en alguna ocasión haya algunas cosas que no se observen.</i>
¿Cómo da retroalimentación con rúbrica?	<i>No doy retroalimentación con rúbrica.</i>	<i>La retroalimentación es siempre corrigiendo los errores que tienen los alumnos. Que ellos se den cuenta en qué</i>

		<i>fallaron, qué es lo que tienen que mejorar.</i>
--	--	--

In relation to teachers' answers, the most important aspects in assessing speaking are content and pronunciation. In the third place is intonation for T1 and memorization for T2.

When scoring, both teachers categorize the score from higher to lower being the highest score the one which shows the most success performance and the lowest represents the least success performance. Besides, T1 uses the concepts of achieved and not achieved. As mentioned in the previous category, both use rubrics. Nevertheless, T1 uses holistic rubrics and T2 analytic rubrics.

Referring to the rubrics advantages, on one hand, T1 specifies that they are clear and objective. On the other hand, T2 states that they are useful because they verify what is being assessed. On the contrary, the rubrics disadvantages for teachers are the following: T1 affirms that they are time - consuming, T2 is worried about the aspects that are not included in the rubrics and that are assessed in the task. In relation to the feedback given with the rubric, T1 manifests that she does not give feedback with this instrument but T2 declares that she does it correcting the mistakes in order to let them know the aspect that need to be improved. Therefore, it can be inferred that T1 gives assessment feedback which involves giving a mark to let the student know how his or her performance was in the task (Ur, 2006) and T2 provides feedback as a correction which implies to give certain information about what the learner did suggesting other ways to do it. However, it also should include information about what the student did well (Ur, 2006).

In the next category, it will be seen the factors which affect the oral assessment.

Category 5: Factors that affect oral assessment

In the following category, it will be analyzed the issues that disturb the speaking assessment in the class. According to Gottlieb (2006), it is complicated to

get reliable information in large scale classes and this is mostly because of the time.

Table IV.10: Factors that affect oral assessment

Question	Teacher 1	Teacher 2
¿Con qué dificultades se encuentra cuando evalúa la expresión oral en sus cursos?	<i>Creo que la más común es que a los niños les da vergüenza. Ellos no se atreven a expresarse oralmente. El ruido que hay en la sala, no son muy respetuosos al momento de escucharse unos a otros. El tiempo también porque tenemos cursos numerosos y... bueno..., y si hay ruido y no se respetan y no se escuchan más y más tiempo y pasan y pasan las horas</i>	<i>El grupo curso no apoya mucho estas actividades porque se necesita estar en silencio para poder captar bien la pronunciación o las expresiones, y las dificultades con los alumnos es en que les cuesta aprender bien las expresiones, la pronunciación, etc.</i>

According to the teachers' points of view, the factor that affects the most when assessing speaking is the students' noise. They express that learners are not able to be in silence so this behavior does not allow them to listen to the students when assessing. Gottlieb (2006) mentions that in assessing speaking there should be a direct interaction between the teacher and the student so, the noise affects this interaction. T1 mentions that this action is disrespectful because they do not listen to each other. Furthermore, T1 declares that another factor that affects speaking assessment is students' shyness since they feel embarrassed when speaking in English in front of the class and Ur (2006) recognizes shyness as problem that could appear in speaking activities. The author states that it could happen because learners are worried about making mistakes and receiving critics. T1 adds that the time is another issue since they teach large classrooms. Gottlieb (2006) affirms that time is a great challenge in large-scale classrooms therefore, the author suggests to use some strategies like to adapt the task according to the student's level of English and to know previously the rubric.

In the following section the analysis of students' interview will be presented. The categories are going to be the following: *frequency, types of evaluative tasks, scoring instruments, dimensions of assessment and feedback*. It will be presented four students' answers, two per teacher. Students from 1 to 6 correspond to T1 and students from 7 to 12 correspond to T2.

In the category presented below, it is going to be analyzed the number of times that speaking is assessed in a semester from the students point of view and when this skill is assessed.

Category 1: Frequency

According to Gottlieb (2006), teachers should plan the assessments of the year. Besides, Coombe, Folse and Hubley (2010) suggest that speaking should be assessed by teachers in class.

Table IV.11: Frequency

Question	Student 1	Student 4	Student 9	Student 12
¿Con qué frecuencia se evalúa la expresión oral? ¿Cuántas veces en el semestre?	<i>Una</i>	<i>Tres ... cuatro.</i>	<i>Unas 2 o 3 veces.</i>	<i>Como 10.</i>
Question	Student 2	Student 5	Student 7	Student 11
¿En qué momento el profesor (a) evalúa la expresión oral?	<i>Con una presentación, dentro de la clase.</i>	<i>En clases.</i>	<i>En clases.</i>	<i>En el comienzo de la clase.</i>

According to the students' answers, most students declare that speaking is assessed at least once per semester. Nevertheless, there is a great difference among the quantity they expressed as some of them affirmed that they are assessed even ten times per semester even though, there is a student who stated that she has not been assessed orally during that year. In relation to the moment

when speaking is assessed, all students agree that it is assessed during the English class which is recommended by Coombe, Folse and Hubley (2010). As a result of the previous analysis, speaking is assessed one or more times in a semester. Coombe, Folse and Hubley (2010) manifest that speaking should be assessed since it is the skill which is used the most to communicate and, in this way, teachers can notice students' progress in the development of this skill. There was one student who stated that it is assessed ten times so that demonstrates that he maybe do not know what speaking assessment is. Other point is when they are assessed, and all of them declare that it is done in classes.

In the next section, the types of assessment tasks will be discovered.

Category 2: Types of evaluative tasks

In this category, it will be analyzed the kinds of activities (which could be understood as tasks too) that teachers use to assess speaking in the English class. Bachman and Palmer define evaluative tasks as "...a language use task as an activity that involves individuals in using language for the purpose of achieving a particular goal or objective in a particular situation" (1997, p. 44). Coombe, Folse and Hubley (2010) manifest that teachers should choose which speaking sample they would use in the speaking assessment when designing the assessment.

Table IV.12: Types of evaluative tasks

Question	Student 3	Student 6	Student 10	Student 11
¿Su profesora de inglés evalúa la expresión oral en clases?	<i>Sí</i>	<i>Sí</i>	<i>Sí</i>	<i>No...poco en ese caso.</i>
Question	Student 4	Student 5	Student 7	Student 12
¿Con qué actividades su profesor(a) evalúa la expresión oral?	<i>Aprendiéndonos letras de canciones,... y el resto son diálogos.</i>	<i>Diálogos</i>	<i>Con diálogos, canciones y vocabulario.</i>	<i>Con disertación... y leer textos del libro.</i>

After analyzing students' answers, it could be interpreted that teachers assess speaking. However, there is a student who manifested that the teacher does not do it frequently. In relation to the activities that teachers utilize to assess speaking, learners mentioned that they were: learning song lyrics, dialogues, vocabulary, and oral presentations and reading aloud. Luoma (2011) categorizes the activities like: description tasks, narrative tasks, instruction tasks, comparing and contrast tasks, explaining and predicting tasks, decision tasks, roleplays and simulations, reacting in situations tasks and structured speaking tasks. Hence, teachers use structured speaking tasks and description task. The most frequent activity was the dialogue. It can be stated that teachers assess speaking and that the most frequent activity are dialogues and singing songs.

In the following category, it will be known the scoring instruments that teachers utilize according to students.

Category 3: Scoring instruments

The following category gives to know which instruments teachers use at the moment of assessing speaking from the students' point of view. Gottlieb (2006) states that observation or recording are used to document students' oral production and interpret it. To do this, teachers need to use rubrics, rating scales and checklists.

Table IV.13: Scoring instruments

Question	Students 1, 2, 4, 5, 6, 7, 8, 9 and 10	Student 3	Student 11/12
¿Su profesor utiliza algún instrumento de evaluación o pauta al evaluar la expresión oral? ¿Cuáles?	No	<i>Sí. El otro día nos hizo grabar un diálogo y ahí empezó a escuchar ella.</i>	<i>Lista de cotejo y escala de apreciación.</i>

According to learners' answers, it can be assumed that teachers hardly ever use scoring instruments. Two students affirmed that teachers utilized checklists and rating scales and the rest manifested that the teacher did not use any instrument. Besides, a student answered that the teacher listens to students and assesses them without any rubric. However, as the assessment is in charge of human beings, which makes the assessment subjective, it is important to have rating scales to ensure a reliable and valid assessment (Luoma, 2011). Besides, Gottlieb (2006) states that students' performance should be documented through direct observation or recording and it could be done with checklists or rating scales.

In the category below, it will be analyzed the dimensions teachers use to assess speaking according to students' point of view and how they think the teacher's assessing way affects their mark in oral production.

Category 4: Dimensions of assessment

Gottlieb (2006) mentions that the principal components or dimensions to assess in speaking are: grammar, vocabulary, intonation and stress, pronunciation, fluency and accuracy.

Table IV.14: Dimensions of assessment

Question	Student 1	Student 4	Student 9	Student 12
¿El profesor (a) evalúa algunos de los siguientes aspectos o criterios en la expresión oral: contenido, entonación, pronunciación, fluidez, memorización, expresiones	<i>Contenido, entonación, pronunciación, fluidez.</i>	<i>Contenido, Pronunciación, Fluidez, Memorización, Expresiones vistas en clases.</i>	<i>Contenido, Entonación, Pronunciación, Fluidez (no tanto), Expresiones vistas en clases. También evalúa la modulación de las palabras. Y gramática,</i>	<i>Contenido, Entonación, Pronunciación, Fluidez, Memorización, Expresiones vistas en clases. Vocabulario y gramática.</i>

vistas en clases? ¿Evalúa otros aspectos? ¿Cuáles?				
Question	Student 2	Student 5	Student 7	Student 9
¿Cree que la forma de evaluar del profesor influye en la nota que obtiene? ¿Por qué?	<i>No. Porque no afecta, escucha no más, por lo que dicen.</i>	<i>No. no influye tanto, de repente. Tiene que ver la expresión de nosotros, de los alumnos no más y de lo que nos está explicando la profesora.</i>	<i>Sí porque yo considero que tengo más habilidad en expresión oral que en memorizar el vocabulario en inglés.</i>	<i>No, porque, ... ella explica lo máximo que puede y si uno no pone atención es como lógico que .. o sea... no influye la nota</i>

After analyzing students' answers, it is evident that teachers generally assess content, pronunciation, fluency and expressions learnt in class even though one student stated that fluency was not always assessed. Also, they frequently assess intonation and memorization. Apart from that, learners manifested that modulation, vocabulary and grammar are also assessed but in less frequency. In relation to students' perception about teachers' assessment methodology, most of them express that the way teachers assess impacts on their marks since the teachers sometimes assess aspects that students are not good at as memorization or singing. Moreover, students state that teachers do not use an assessment instrument or the tasks are difficult for learners. However, some of them manifest that teachers' way of assessing does not influence because they explain as much as they can and that it depends on the student's performance. It is important to notice that S7 affirmed that how teachers assess affects her since she is good at speaking but not at memorizing so it decreases her mark. As it can be seen, this

student is better at a communicative skill which is speaking and, as the teacher also assesses memorization (which is not communicative), her mark gets lower. Besides, in a speaking task, students have to use the language (Luoma, 2011). Therefore, teachers do not always assess the communicative aspects as pronunciation and fluency. The aspects that they usually assess are content, pronunciation, fluency and expression that teachers teach in class. Furthermore, students consider that the way how teachers assess influence in the mark that they obtain because they assess criteria that do not belong to speaking, they do not usually use a scoring instrument or the task is hard for them. For Luoma (2011), teachers should define the criteria and scales before the assessment. Nevertheless, there were learners who affirmed that how the teacher assessed did not impact their results since the instructions are explained in a good form and the mark depended on their performance.

In the next category, it will be commented the way teachers' provide feedback to learners according to students' point of view.

Category 5: Feedback

The category below provides information about how teachers inform the students about their mark and performance and if they express learners' strengths and weaknesses. Ur (2006) defines feedback as the information provided to students about how he or she performs in a task with the aim of improving the student's performance.

Table IV.15: Feedback

Question	Student 2	Student 4	Student 9	Student 10
¿Cómo el profesor(a) informa su resultado o nota? ¿Le dice sus fortalezas y/o debilidades?	<i>Lee las notas, en voz alta pero no dice fortalezas o debilidades.</i>	<i>Sí... cuando nos llama nos dice en qué nos equivocamos.</i>	<i>Entregando la prueba. (¿Le dice sus fortalezas y/o debilidades?) no.</i>	<i>Sí las dice, las notas las dice ... y nos dice cuándo nos va mal, o sea, nos dice al frente de todos en qué nos equivocamos y</i>

				<i>cómo deberíamos mejorar</i>
--	--	--	--	--

The students answer that teachers say the marks in general and that they communicate students' weaknesses and sometimes teachers refer to their strengths. Hence, it can be considered as one of the feedback components classified by Ur (2006). This feedback is known as assessment which is used to explicit how the students performed in the assessment. Also, it can be stated that teachers say the marks to students and that they give more negative feedback than positive feedback. Therefore, the last can be recognized as correction, the other component of feedback (Ur, 2006). It can be noticed because teachers tell students specifically what they did wrong and well. For students, all feedback given has relation to give the mark.

In the next section, it will be analyzed which test specifications the teachers use at a public high school.

3. Specific objective 3: To identify the test specifications that teachers of English apply to assess speaking.

In the following table, it will be presented how teachers accomplish Luoma's test specifications. Luoma (2011) starts suggesting that test specifications should be written at the beginning of a test design. For this author, specifications for speaking assessment indicates what type of speaking will be assessed , how it is going to be assessed and how it is going to be scored. As it is mentioned before, Luoma proposed 21 specifications (see page 18). The table below shows the average of achievement of the test specifications by each teacher. The level of achievement was distributed in a rating scale with the following level of accomplishment.

- Adequate: the test specification was completely achieved.
- Partially adequate: the test specification was partially achieved.

- Minimally adequate: the test specification was poorly achieved.
- Inadequate: the test specification was not achieved.
- Not observed: the test specification was not present.

Table IV.16: Teachers' accomplishment of test specifications

	Adequate (4)	Partially adequate (3)	Minimally adequate (2)	Inadequate (1)	Not observed (0)
T1	10/21 (47,6%)	1/21 (4,8%)	2/21 (9,5%)	0/21 (0 %)	8/21 (38,1%)
T2	9/21 (42,9%)	3/21 (14,3%)	6/21 (28,6%)	0/21 (0%)	3/21 (14,3%)

From the table, it could be observed that T1 achieved more *adequate* aspects (47,6%) from Luoma's list (2011) than T2 (42,9%) however, T1 has more *inadequate* aspects (38,1%) than T2 (14,3%). Therefore, on one hand, it could be stated that T1 considers more communicative aspects when assessing oral production but not all the aspects mentioned by Luoma. On the other hand, even though T2 has less percentage in *adequate*, she presents more test specifications than T1 but they are not completely well used. It could be since the T2's assessment task was less communicative than T1 as students had to memorize a dialogue instead of presenting orally information like in the T1's assessment task.

The table below shows the *adequate* aspects that each teacher achieved.

Table IV.17: Luoma's adequate aspects: Teacher 1 and Teacher 2

Teacher 1	Teacher 2
<ul style="list-style-type: none"> • The test's purpose • Time for each section/paper. • Target language situation. • Text length. • Language skills to be tested. • Language elements to be tested. 	<ul style="list-style-type: none"> • Test level. • Time for each section/paper. • Weighting for each section/paper. • Text length. • Test tasks.

<ul style="list-style-type: none"> • Test tasks. • Test methods. • Descriptions of typical performance at each level. • Sample papers. 	<ul style="list-style-type: none"> • Rubrics. • Descriptions of typical performance at each level. • Description of what candidates at each level can do in the real world. • Sample papers.
--	--

According to Luoma (2011), the specifications could be joined together in three modules: *construct specification*, *assessment specification* and *task specifications*. First, *construct specifications* define the connection that exists between skill and how this one is going to be performed and assessed. Second, in *assessment specifications*, the criteria that will be used is defined how they will be implemented at the moment of assessing. Finally, the *task specifications* describe the task that will be implemented when assessing and the instructions that it will have. Through the comparison of both teachers, it can be interpreted that T1's *adequate* aspects are more related to the administration and context of the assessment (*construct specifications*) and to the *task* because she has present aspects like the purpose, the time, length which are associated to the administration of a test. Also, the aspects like the language skills, test task, sample paper are related to the assessment task. Otherwise, T2's aspects are linked to the test administration and to the *assessment* instead of the task. This teacher shows an *adequate* performance in aspects as rubrics, descriptions of performance in each level and description of what candidates at each level can do in the real world. Hence, both teachers have *adequate* in aspects related to the administration of the test. However, T1 has more *adequate* aspects related to the *task* and T2 in aspects related to the *assessment*,

To continue, it will be presented the table which has the test specifications that teachers achieved in a *minimally adequate* way.

Table IV.18: Luoma’s minimally adequate aspects: Teacher 1 and Teacher 2

Teacher 1	Teacher 2
<ul style="list-style-type: none"> • Rubrics • Criteria for marking 	<ul style="list-style-type: none"> • The test’s purpose. • Description of suitable language course or textbook • Number of sections/papers • Target language situation • Language skills to be tested • Test methods

It can be seen that T1 has *minimally adequate* aspects in the *assessment specifications*. This can be explained because of the way the rating scale was designed. This scale was not made specifically for the task, instead it was taken from the coursebook. Therefore, the criteria were not totally adapted to the task. In this point, O’Malley and Valdez (1996) claim that the criterion levels of oral language proficiency should be created according to the classroom aims and tasks. Nevertheless, the assessment had criteria related to communicative aspects like fluency and the use of completed sentences. In the case of T2, the aspects which were accomplished in a *minimally adequate* way were associated to the *construct* and *task specifications*. For example, T2 did not have a clear test purpose; the number of sections and the target language situation were not specified. It may happen because she took the activity from the coursebook and she asked students to memorize a dialogue instead of asking them to create one so the task was not communicative. Besides, it is important to notice that she was assessing the speaking ability; however, the task did not allow students to demonstrate completely their speaking skill as they just had to reproduce a given dialogue. Therefore, this assessment task was not authentic enough. In this regard, O’Malley and Valdez (1996) explicit that speaking activities should give the students the opportunities to use the language for authentic purposes, in other words, to give the chance to exchange information with a purpose instead of

providing known information to the listener. In sum, T1 has *minimally adequate* aspects in *assessment* (rubrics and criteria for marking) and T2 in the specifications related to the *construct* and the *task* (the test's purpose, description of suitable language course or textbook, number of sections/papers, target language situation, language skills to be tested and test methods). It is important to note that T1 did not have *adequate* aspects in *assessment* and T2 did not have *adequate* aspects in the *task specifications* in the previous table (Table IV.17) but this teacher did have several *adequate* test specifications before.

In the following table, it will be shown the test specifications that were *not observed* in the teachers.

Table IV. 19: Luoma's *not observed* aspects: Teacher 1 and Teacher 2

Teacher 1	Teacher 2
<ul style="list-style-type: none"> • Description of the examinees. • Test level • Definition of construct (theoretical framework for the test) • Description of suitable language course or textbook • Weighting for each section/paper • Text-types • Description of what candidates at each level can do in the real world • Samples of students' performance on task 	<ul style="list-style-type: none"> • Description of the examinees. • Definition of construct (theoretical framework for the test) • Samples of students' performance on task

From the table above, it can be observed that T1 did not present specifications related to *construct*, *assessment* and *task* but most of them are

related to the *construct specifications* as the description of the examinees, test level, definition of construct among others. The aspect related to assessment was the description of what candidates at each level can do in the real world and the weighting for each section. Samples of students' performance on task are associated to the task specifications. In the case of T2, who has less *not observed* specifications, the aspects were related to *construct* and *task specifications*. In relation to the *construct*, there were not present the description the examinee and the definition of construct and in relation to the *task*, she did not show samples of the students' performance on task. It could be notice that both teachers coincide in not having the description the examinees, the definition of construct and samples of the students' performance on task. Hence, it can be stated that teachers are not used to these specifications. This has relation with interactiveness, one of the assessment qualities, which is related to how the test taker's individual characteristics are involved to achieve a test task (Bachman and Palmer, 1997).

In the following section, there will be studied the connection that exists among assessment tasks, assessment instruments and test specifications with the students' performance in speaking.

4. Specific objective 4: To analyze the relation among assessment tasks, assessment instruments and test specifications with the students' performance in speaking.

In this section, it will be analyzed what the relation is among assessment tasks, assessment instruments and test specifications with the students' performance in speaking because the learners' performance will be used since it is wanted to know how these elements influenced on students' marks. Even though, in the teachers' interviews were asked which assessment tasks they used when assessing speaking, in this objective it will be taken in consideration what was observed in class using the checklist and the rating scale.

In the next table, it will be examined what the relation was in T1.

Table IV.20: T1's assessment task, assessment instrument and test specifications and students' performance in speaking.

	Assessment task	Assessment instruments	Test specifications	Students performance
TEACHER 1	Oral presentation about natural disasters	Rating scale taken from the coursebook (fluency, grammar and vocabulary)	Adequate	1.0-3.9 (12.5%)
			Test's purpose.	4.0-4.4 (12.5%)
			Time for each section/paper explicit.	4.5-4.9 (6.3 %)
			Target language situation explicit.	5.0- 5.4 (37.5%)
			Text length.	5.5-5.9 (6.3%)
			Language skills to be tasted.	6.0-6.4 (12.5%)
			Language elements to be tested.	6.5-7.0 (12.5%)
			Test tasks.	
			Test methods.	
			Descriptions of typical performance at each level.	
Sample papers.				
Minimally adequate				
Rubrics.				
Criteria for marking.				

It could be noticed that the teacher used an oral presentation as a task and a rating scale to assess it. The oral presentation is part of prepared talks where students choose a topic, research about it and then present it in front of the class and they can use notes.

The specifications which are *adequate* are related to *construct* and *task*. Most of the marks are between 5.0-5.4 (37.5%). Therefore, it could be inferred that the task and the way of assessing it could have an impact on the students' mark. The students were assessed through a communicative task so they had to apply their speaking abilities in an authentic context as O'Malley and Valdez (1996) state. This way of assessing and the students' performance (69% over 5.0 and 88% over the passing mark) proves that student participants are able to communicate in L2 in context of authentic tasks. However, as it was mentioned before, the rating used by T1 did not assess content and pronunciation as O'Malley and Valdez (1996) suggest. But the range of marks is in the normal Gaussian distribution, just a 12,5% failed.

It will be presented the table of T2 where it can be seen the relation among the assessment task, assessment instrument and test specifications and students' performance in speaking.

Table IV.21: T2's assessment task, assessment instrument and test specifications and students' performance in speaking.

	Assessment task	Assessment instruments	Test specifications	Students performance
TEACHER 2	Dialogue previously memorized	Analytic Rubric (pronunciation, fluency, grammar, interaction, task and memory)	Adequate: Test level. Time for each section/paper. Weighting for each section/paper. Text length. Test tasks. Rubrics. Descriptions of typical performance	1.0-3.9 (0%) 4.0-4.4 (2.5%) 4.5-4.9 (0 %) 5.0- 5.4 (5.1%) 5.5-5.9 (0%) 6.0-6.4 (12.8%) 6.5-7.0 (79.4%)

			<p>at each level.</p> <p>Description of what candidates at each level can do in the real world.</p> <p>Sample papers.</p> <p>Minimally adequate:</p> <p>The test's purpose.</p> <p>Description of suitable language course or textbook</p> <p>Number of sections/papers</p> <p>Target language situation</p> <p>Language skills to be tested</p> <p>Test methods</p>	
--	--	--	--	--

As the table above evidences, T2 used a memorized dialogue as an assessment task that just allowed students to give known information to the listener instead of giving the chance of having a conversation as O'Malley and Valdez (1996) note. Besides, the students had a whole class to practice and learn the dialogue to be assessed. It can be seen that T2 has strengths in *construct* and *assessment* specifications and that most of the marks are between 6.5 and 7.0 (79.4%). Hence, it can be stated that students got good marks because they had the time to rehearse in class and clarify their doubts in pronunciation. Furthermore, the activity was not communicative as learners did not create the dialogue; it was taken from the coursebook and their task was to learn it by heart and reproduce it.

Thus, memory was the only aspect assessed and not the students' communicative competence.

Therefore, through the analysis, it can be deduced that the students' performance depends on the assessment task. It is relevant to remind that the tasks have to give learners the chance to utilize the language to know students' development and how they can communicate in a fluent and effective way as Ellis (2009) affirms. This is what T1 tries to implement. Another issue that influences is the assessment instrument because, if it does not assess all the aspects that the task involves or that represent the class outcomes as O'Malley and Valdez (1996) state, the mark could not represent the student's performance. This is what happened with T2's task that only assessed memory instead of communication so the marks did not show the real learners' ability to communicate in English as a Foreign Language. For instance, if the assessment instrument does not assess pronunciation in a speaking activity, the mark would not show the real learner's performance and this happened with T1's students.

5. Discussion

This research project was conducted in order to investigate which test specifications, instruments and tasks are used by teachers of English of a public high school to assess speaking and its effect on secondary students' performance. As a result, it was discovered that how teachers assess impacts on students marks.

First, teachers must use assessment instruments. This idea is supported by Gottlieb (2006) who affirms that teachers observe or record in order to interpret students' oral performance. To achieve this, teachers need to occupy assessment instruments like rubrics, rating scales and checklists. However, there are instruments that are not adapted to the tasks so there are some aspects that are not assessed. Coombe, Folse and Hubley (2010) give the responsibility of choosing the criteria to assess to teachers. Nevertheless, O'Malley and Valdez

(1996) advise that the criteria to assess speaking should be based on the classroom aims. In a speaking task, it is important to assess task and pronunciation (Luoma 2011). Furthermore, there should be added accuracy, vocabulary, linguistic ability, fluency, content or ideas (Coombe, False and Hublely, 2010).

Second, most of the activities that teachers use to assess speaking are communicative. Nonetheless, the aspects or criteria that they assessed are not totally communicative because teachers assess more linguistics aspects, like grammar and vocabulary, than the communicative ones, like pronunciation, fluency and interaction. As mentioned before, teachers should include the following criteria: accuracy, vocabulary, linguistic ability, fluency and content or ideas (Coombe, False and Hublely, 2010) and they should be according to the objectives that the teacher has (O'Malley and Valdez, 1996). Otherwise, to have more communicative aspects to assess, teachers should follow Luoma's test specifications (2011). The author declares that specifications for speaking indicate what type of speaking will be assessed, how it is going to be assessed and how it is going to be scored. Among the test specifications suggested are: language elements to be tested, test tasks, test methods, rubrics, criteria for marking and descriptions of typical performance at each level. The teachers use some of the specifications but they do it in an unconscious way. Teachers could use them consciously but they should know and understand them first of all. Also, teachers should have the time to work with them.

Third, when teachers assess, they are the people in charge of choosing the task, the assessing instruments and the criteria which are going to assess the students' marks. If the teacher chooses an activity that is not communicative, like memorize something, students are going to get better grades than having an activity which is more communicative because, in the last one, they would have to use the language. Ellis (2009) declares that a task is an activity where students have to use the language in order to achieve an objective in a specific situation.

O'Malley and Valdez (1996) add speaking activities should provide the change to give the chance to exchange information with a purpose instead of providing known information to the listener. Furthermore, it is important to select the suitable assessment instrument and the criteria. If the criteria do not represent what the activity involves, the mark will not represent the students' performance in the activity. For this reason, it is not a good idea to use the rubrics present in the coursebook because they are general. They could be used as a template to create one specifically for the activity and students' level of English.

Finally, teachers should have workshops or courses at schools that allow teachers to know that there are instruments, criteria, test specifications and tasks specifically for speaking. This would permit that teachers assess properly the oral productions and that marks represent students' performance,

CHAPTER V
CONCLUSIONS

1. Conclusions

Throughout the data analysis, this research has tried to find out which test specifications, instruments and tasks are used by teachers of English of a public high school to assess speaking and its effect on secondary students' performance. With the purpose of having a logical sequence of conclusions, this section will be organized by specific objectives and their assumptions.

1.1 To identify the instruments and procedures that teachers of English apply to assess speaking.

According to what was observed in a class using the checklist, teachers use oral activities to assess speaking like oral report and dialogues. Therefore, they allow students to show their ability to speak (Coombe, False and Hubley, 2010). In relation to the assessment instruments, teachers assess speaking utilizing analytic rubrics and rating scales. However, they take these instruments from the English book and the teachers did not adapt these instruments to the activity. For that reason, there were some aspects which were not considered like task and pronunciation which are important when assessing speaking (Luoma, 2011). One of the teachers added memorization outside the assessment speaking. Otherwise, at the moment of assessing speaking, teachers prefer to do it in class, it may happen because it is the only occasion when teachers can meet all the students. Besides, teachers are in charge of doing it and he or she is the only responsible of assessing. The assessment could be individual or in groups but the mark depends on each student's performance. The positive aspect of it is that every student knows how they performed in that activity. Another aspect is the materials students use at the moment of speaking. One of the teachers used pictures and the other one did not let students to have any material but they had the dialogue to practice and learn it by heart. Otherwise, it can be stated that the way teachers assess has an impact on students' marks, especially the activity that the teacher uses to assess speaking. It was evidenced because the teacher who utilized a memorized dialogue had better marks (most of students had between 6,5 and 7,0) than the teacher who assessed an oral report (most of students had between 5,0 and 5,4).

Another aspect that influences was the oral report which was more communicative since students had to describe a natural disaster with the support of a picture. Moreover, the teacher who got lower marks occupied an analytic rubric with communicative aspects like fluency and the other used a rating with less communicative aspects that were more linguistics like memorization. Maybe, this occurred since it is more difficult to achieve the communicative aspects as learners do not practice them in class. Nevertheless, the linguistic aspects are more practiced in class so students get good scores in this type of criteria.

1.2 To characterize the tasks that teachers of English and students declare to use to assess speaking.

The conclusions for this objective were deduced from the interview made to the teachers and their students. Hence, it can be concluded that teachers assess speaking in English class at least once per semester. The frequency depends on the teacher. Students agree that teachers assess speaking during the class once per semester at least. Nevertheless, students gave different quantities when they were asked for the frequency of speaking assessment. Hence, it can be inferred that there are some students who do not know when it is an assessment or when it is an activity.

The Ministry of Education (2015) establishes teachers should assess the expected learnings and, in the case of the English subject, they are related to the communicative skills so the teachers accomplish this requirement. However, this ability does not have the same importance for the teachers since one assesses this skill because it is necessary for communication; the other does it because it is required by the Chilean Ministry of Education. Anyway, teachers do not consider that speaking is less important than the other communicative abilities.

Otherwise, the instrument that teachers declare to use is the analytic rubric which has the description of specific dimensions or characteristics of the contrast that is going to be measured (Gottlieb, 2006). Nevertheless, most of the students

answered that teachers do not use any assessment instrument. Therefore, it could be thought that students do not know the instrument when they are assessed so it could be the cause that they manifest that their teachers do not use one. In relation to the activities used to assess speaking, teachers mention principally oral presentation and one of them a dialogue so they are using the adequate activities since learners can use the language to reach an objective in a specific context (Luoma, 2011). Besides, students add that one of the activities is singing songs.

The criteria that teachers include in their rubrics are frequently: content and pronunciation. Coombe, False and Hubley (2010) propose five categories: accuracy, vocabulary, linguistic ability, fluency and content or ideas. Furthermore, students declared that teachers assess content, pronunciation, fluency and expressions learnt in class. As it can be seen, teachers used two of the five categories suggested. Therefore, teachers and students coincide in content, and pronunciation.

In regard to the advantages and disadvantages of the use of rubrics, teachers consider that the rubrics advantages are that they are clear, objective and they verify what is assessed. Otherwise, in relation to the disadvantages, the teachers stated that the rubrics are time-consuming and that they do not have all the criteria needed to be assessed in the task in some occasions. Nevertheless, Gottlieb (2006) affirms that rubrics are clear, observable and measurable. In relation to the feedback provided, teachers do not use the rubric to give it because one teacher does not write or say any comment and the other teacher corrects the learners' mistakes. Students manifested something similar; they affirmed that teachers communicate their marks and that they usually give corrective feedback because the teachers give comments about what students did well and wrong (Ur, 2006). Referring to the difficulties when assessing speaking, teachers mentioned that the most important factor is the noise. Other aspect that affects is the students' shyness. Both of them are related to students' behavior which influence in their speaking performance since directed interaction may be affected (Gottlieb, 2006)

1.3 To identify the test specifications that teachers of English utilize to assess speaking.

In order to obtain the conclusions for this objective, the researcher observed the teachers' classes using a rating scale with Luoma's test specifications (2011). The test specifications proposed by Luoma are presented on page 18. The descriptors of accomplishment in the rating scale were the following:

According to what it was observed in the teachers' classes, it can be concluded that T1 had more adequate aspects than T2 which could be because the task was more communicative as learners has to present a topic and, in T2's task, students had to learn a dialogue. The aspects that T1 achieved in an adequate level were related to construct specifications (related to administration and context) and task since the present aspects were related to the purpose, the time, length which are associated to the administration of a test. In the case of T2, the adequate aspects were presented in construct specifications and assessment. T2 shows an adequate performance in aspects as rubrics, descriptions of performance in each level and description of what candidates at each level can do in the real world. Thereupon, it can be noticed that teachers do not have problems with the administration of the test. Nonetheless, one of them has more adequate aspects in the task, maybe because it was more communicative, and the other in assessment, it could be since the rubric was more specific.

In relation to the minimally adequate aspects, T2 had more aspects than T1. T1's aspects were related to the assessment because the criteria and descriptors of the rating scale did not totally match with the purpose of the task. In T2, the minimally adequate aspects were related to the construct and task. It could have occurred since the task and the rubric were taken from the coursebook and she did not ask to create a dialogue, she asked just to memorize one as the task was not adapted for the unit goals.

Otherwise, referring to the not observed aspects, T1 had more not observed aspects than T2. T1 had some missing aspects in construct specifications like the description of the examinees, test level, definition of construct among others. It could happen as it is not usual for teachers to describe students or construct in a written and explicit way but they consider these before assessing. The non observed aspects in T2 were similar because they were related to the test administration but they were less.

As a conclusion, it can be manifested that the more communicative is the task, the more specifications are present. Moreover, it can be stated that teachers don't use test specifications in a conscious way. Maybe, if they do it, the task and the assessment would be achieved in a correct form by the students. Another issue that can influence on the use of test specifications is that teachers took the activity and the rubric from the coursebook so they are not adapted to unit goals and the rubric is not made specifically for the task. These issues can affect the learner' marks.

1.4 To analyze the relation among tasks, assessment instruments and test specifications with the students' performance in speaking.

The conclusions were made according to what was observed in class using the checklist and the rating scale. From the observation, it could be stated that the task impacts the learners' performance. When the task is communicative, students get lower marks. It could happen because learners have to create and then explain at the moment of speaking or maybe they do not practice these aspects enough in classes. It was the case of T1 where 37% of students got between 5.0 – 5.4. Other aspect that affected the learners' mark was how the task was assessed. T1 used a rating scale taken from the coursebook with which she assessed fluency, grammar and vocabulary. However, the teacher did not assess nor content neither pronunciation which is suggested by O'Malley and Valdez (1996). Maybe, if these aspects had been assessed, the students' marks would have been better.

On the contrary, T2 got better marks, 79,4% of students got between 6.5-7.0. It could have happened because learners have to say a dialogue by heart so they did not speak naturally. Moreover, they had time to rehearse. Therefore, as the task was not communicative and complex, the students had more opportunities to have higher marks. In addition, the analytic rubrics scales were more descriptive even though it was taken from the coursebook. This is why there were some missing criteria to assess, for instance; interaction.

After analyzing the information obtained by the observation, it could be stated that the task has an influence on the mark since the more communicative the task is, the lower their marks will be. It might occur because learners have to use the language in a fluent and effective way as Ellis (2009) manifests. Moreover, the instrument and their criteria can influence on students' mark because, if there is a criterion which is not included in the rubric, the mark will not demonstrate exactly the learners' performance. The assessment instrument should be done according to the task in order it represents the objectives that teachers want students to accomplish

2. Limitations

The current research had to deal with three limitations.

The first one is the lack of knowledge that teachers have about assessing speaking. It could be evidenced since they do not use properly the assessment instruments because they do not include the criteria that represent all the task aspects. When they utilize an instrument, they take them from the coursebook. Also, there are some tasks that are not communicative so they are not assessing oral production and they are assessing memory. Maybe, if teachers had a workshop about speaking assessment, they would improve their tasks and instruments.

The second one deals with the quantity of samples and participants for this research. The sample could have been larger and varied since not all the teachers and their students from the high school were observed. It did not happen because the other teachers did not want to participate. In addition, it could be observed more speaking activities from the same teachers but, as this skill is not assessed frequently, it could not occur. Moreover, it could have included more teachers from public high school to have a wider view.

The last one is concerning with the lack of time to explain how to assess speaking in order to make teachers improve their assessments. The researcher just had time to observe and analyze the data collected but she could not give feedback or suggestions to the teachers. Anyway, to give feedback was not an objective of the research.

3. Further research

The process of researching about how speaking is assessed will help teachers to research and improve the assessment of this communicative skill. Therefore, the tasks, instruments and test specifications will help to represent students' performance when their speaking skill is assessed.

Also, this research could be applied to in other skills like writing, reading and listening in order to know if teachers are more prepared to assess those skills. To know which activities and instruments they utilize.

In addition, the research project could be applied in different types of schools like subsidized and private to notice if there is a difference among teachers when assessing speaking. Furthermore, the teachers could be divided into ages to discover if the age influences on the way of assessing speaking.

REFERENCES

Agencia de la Calidad de la Educación. (2015). Informe Resultados de Aprendizaje Docentes y Directivos. Santiago de Chile: Agencia de la Calidad de la Educación

Assessment of Student Learning. (2016). Taken on August 29th, 2016 from <http://www.gallaudet.edu/office-of-academic-quality/assessment-office/assessment-of-student-learning-outcomes/the-hows-and-examples/assessment-tools-and-methods.html>

Bachman, L. and Palmer, A. (1997). *Language Testing in Practice*. Hong Kong: Oxford University Press.

Brown, D. (2003). *Language Assessment: Principles and Classroom Practices*. United States of America: Longman.

Burns, A. (2003). *Collaborative Action Research for English Language Teachers*. Cambridge : Cambridge University Press

Cohen, L., Manion, L. and Morrison, K. (2007). *Research Methods in Education* 6th ed. USA and Canada: Routledge

Concordia University. (2013). Grading vs. assessment: What is the difference? Taken August 21st, 2016 from <http://education.cu-portland.edu/blog/curriculum-instruction/grading-vs-assessment-whats-the-difference/>

Coombe, C., Folse, K. and Hubley, N. (2010). *A Practical Guide to Assessing English Language Learning*. United States of America: The University of Michigan Press.

Ellis, R. (2009). *Task-based Language Learning and Teaching*. China: Oxford.

Fulcher, G. *Language Testing Resources*. Retrieved on August 29th, 2017 from <http://languagetesting.info>

- Fulcher, G. and Davidson, F. (2007). *Language Testing and Assessment: An Advanced Resource Book*. Great Britain: Routledge
- Gallaudet University. (n.d). *Introductions: Assessment Tools & Methods*. Taken on October 13th, 2016 from <https://www.gallaudet.edu/office-of-academic-quality/assessment/assessment-of-student-learning/instructions-and-examples/assessment-tools-and-methods>
- García M del P and García M L. (2003). *Age and the Acquisition of English as a Foreign Language*. Great Britain: Multilingual Matters Ltd
- Gerring, J. (2004) 'What Is a Case Study and What Is It Good for?', *American Political Science Review*, 98(2), pp. 341–354. doi: 10.1017/S0003055404001182
- Gottlieb, M. (2006). *Assessing English Language Learners: Bridges from Language Proficiency to Academic Achievement*. United States of America: Corwin Press.
- Grading vs. Assessment: What's the Difference?. (2013). Taken on August 29th, 2016 from <http://education.cu-portland.edu/blog/curriculum-instruction/grading-vs-assessment-whats-the-difference/>
- Harmer, J. (2003). *The practice of English Language Teaching*. Malaysia: Longman.
- Jamshed, S. (2014). "Qualitative Research Method-interviewing and observation". *Journal of Basic and Clinical Pharmacy*, 5 (4), pp 87-88. Taken on March 3rd from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4194943/#>
- Jupp, V. (2006). *The SAGE Dictionary of Social Research Methods*. Great Britain: SAGE
- Knupfer, N., and McLellan, H. (1996). Descriptive research methodologies. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1196–1212). New York: Macmillan.

- Kumar, Y. (2007). *Research Methodology*. A. P. H Publishing Corporation: New Delhi.
- Luoma, S. (2011). *Assessing Speaking*. United Kingdom: Cambridge University Press
- McDonough, J. and McDonough, S. (2005). *Research Methods for English Language Teachers*. London: Arnold
- Merriam, S. (2009). *Qualitative Research: A Guide to Design and Implementation*. United States of America: Jossey-Bass
- Ministry of Education. (2011). Inglés: programa de Estudio para Segundo Año Nuevo. Santiago: Ministerio de Educación, República de Chile
- Ministry of Education. (2015). Inglés: Programa de Estudio Cuarto Año Medio. Santiago de Chile: Ministerio de Educación de Chile
- Nguyen, C., and Tran Thi Le. (2015). The Interconnection Between Interpretivist Paradigm and Qualitative Methods in Education. *American Journal of Educational Science*, 1(2), 24–27.
- O'Malley, J. and Valdez L. (1996) *Authentic Assessment for English language Learners: Practical Approaches for Teachers*. United states of America: Addison-Wesley Publishing Company.
- Phillipson, R. (2003). *Linguistic Imperialism*. China: Oxford University Press
- Phothongsunan, S. (2010). *Interpretative Paradigm in Educational Research*. Taken on January 24th, 2017 from <http://repository.au.edu/bitstream/handle/6623004553/13708/galaxy-iele-v2-n1-1-oct-10.pdf?sequence=1>
- Sahu. S. and Singh, T. (2016). *Research Methodology*. India: SBPD Publications

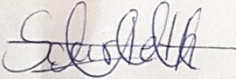
- Thomas, P. (2010). Research methodology and design. Taken on January 24th, 2017 from http://uir.unisa.ac.za/bitstream/handle/10500/4245/05Chap%204_Research%20methodology%20and%20design.pdf
- Ur, P. (2006). *A Course in Language Teaching: Practice and Theory*, United Kingdom: Cambridge University Press.
- Zainal, Z. (2007). Case study as a research method. Taken on January 24th, 2017 from http://psyking.net/htmlobj-3837/case_study_as_a_research_method.pdf

APPENDIXES

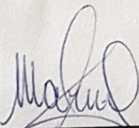
APPENDIX N°1: Sample of Informed Consent

Consentimiento informado

Yo, Paulina Acuña Guerrero declaro que voluntariamente estuve de acuerdo en participar en esta investigación dado que recibí toda la información necesaria de lo que incluirá la misma y que tuve la oportunidad de formular todas las preguntas necesarias para mi entendimiento. Además, autorizo que la información que proporcioné sea publicada por los investigadores. También, entiendo que debería tener una copia de este consentimiento.


Firma

18/11/2016.
Fecha


Firma del investigador

18/11/2016
Fecha

APPENDIX N°2: Checklist to observe how a teacher of English assesses speaking

Lista de Cotejo 1 para evaluar cómo el profesor(a) de inglés evalúa la expresión oral

Nombre del profesor(a): _____

Nombre del observador(a): _____

Establecimiento: _____

Curso en el que se observa: _____ Fecha: _____

Objetivo: Determinar cómo el profesor evalúa la expresión oral en inglés y cómo esto influye en el rendimiento de los estudiantes.

Instrucciones: Marque con una X en "SÍ" si el aspecto se encuentra presente y marque con una X en "NO" si se encuentra ausente.

Aspectos a evaluar	SÍ	NO	Observación
1. Desarrolla una(s) actividad(es) evaluativa(s) definida(s).			
2. Las actividades desarrolladas en clases son distintas entre sí.			
3. Utiliza instrumentos de evaluación:			
e) Lista de cotejo.			
f) Escala de apreciación.			
g) Rúbrica.			
h) Otros.			
4. Usa criterios para evaluar.			
g) Pronunciación.			
h) Fluidez.			
i) Gramática.			
j) Habilidades lingüísticas.			
k) Contenido o ideas			
l) Otros criterios.			

5. Evalúa en la clase de inglés.			
6. Evalúa de forma individual.			
7. Evalúa de forma grupal.			
8. Utiliza materiales de apoyo para evaluar.			
9. Sólo el profesor evalúa.			
10. La mayoría de las notas está en el siguiente rango:			
h) 6,5 al 7,0			
i) 6,0 al 6,4			
j) 5,5 al 5,9			
k) 5,0 al 5,4			
l) 4,5 al 4,9			
m) 4,0 al 4,4			
n) 1,0 al 3,9			

APPENDIX N°3: Rating scale to analyze the scoring instruments used by the teacher at the moment of assessing speaking

Escala de apreciación para analizar los instrumentos de evaluación utilizados el profesor al momento de evaluar la expresión oral en inglés

Nombre del profesor(a): _____

Nombre del observador(a): _____

Establecimiento: _____

Curso en el que se observa: _____ Fecha: _____

Objetivo: identificar las “test specifications” que los profesores emplean al evaluar la expresión oral.

La siguiente lista de cotejo está basada en las “test specifications” mencionadas y definidas por Sari Luoma en su libro *Assessing Speaking*. Según Luoma (2004), las “test specifications” se utilizan para especificar las actividades que utilizarán para evaluar la expresión oral, qué evaluación se hará y cuáles son los criterios que se usarán. Éstas establecen los contenidos que se evaluarán, las definiciones de las tareas detalladas y la clasificación de los criterios a evaluar para guiar el desarrollo de las tareas comparables y la entrega de una evaluación justa. Las “test specifications” evidencian por qué la evaluación se focaliza en ciertos contenidos y cómo las tareas y los criterios de evaluación se llevan a cabo en ellos. La autora señala la siguiente lista de “test specifications” (Luoma, 2004, p. 114):

Aspectos a evaluar Rango	Adecuado (4)	Medianamente adecuado (3)	Poco adecuado (2)	Inadecuado (1)	No observado (0)
1. The test's purpose.					
2. Description of the examinees.					
3. Test level.					
4. Definition of					

construct (theoretical framework for the test).					
5. Description of suitable language course or textbook.					
6. Number of sections/papers.					
7. Time for each section/paper.					
8. Weighting for each section/paper.					
9. Target language situation.					
10. Text-types.					
11. Text length.					
12. Language skills to be tested.					
13. Language elements to be tested.					
14. Test tasks.					
15. Test methods.					
16. Rubrics.					
17. Criteria for marking.					
18. Descriptions of typical performance at each level.					
19. Description of what candidates at each level can do in the real					

world.					
20. Sample papers.					
21. Samples of students' performance on task.					
Total de 84					

Traducción de “test specifications” planteadas por Luoma (2004, p.114):

1. Propósito de la prueba.
2. Descripción de los examinados.
3. Nivel del test.
4. Definición del constructo (marco teórico para la prueba).
5. Descripción del curso de idiomas o libro apropiado.
6. Número de secciones/prueba.
7. Tiempo para cada sección/prueba.
8. Ponderación para cada sección/prueba.
9. Situación de la lengua meta.
10. Tipos de texto.
11. Longitud del texto.
12. Habilidades comunicativas a evaluar
13. Elementos del idioma a evaluar.
14. Tareas de la prueba.
15. Métodos de evaluación.
16. Rúbricas.
17. Criterios de calificación.
18. Descripciones del desempeño típico en cada nivel.
19. Descripción de lo que los candidatos son capaces de hacer en el mundo real.
20. Ejemplos de pruebas.
21. Ejemplos del desempeño de los estudiantes en la tarea.

Validado por
María Angélica Fuica
Traductora

APPENDIX N°4: Semi-structured interview for teachers about the scoring instruments used by the teachers in the speaking assessment

Entrevista semiestructurada sobre instrumentos evaluativos utilizados en la evaluación de la expresión oral

La siguiente entrevista semi-estructurada se utilizará en una tesis de magíster que tiene como título “Análisis de procedimientos, actividades, instrumentos y “test specifications” utilizados por profesores de inglés del liceo de la comuna de Curanilahue, para evaluar la expresión oral en inglés y su efecto en el desempeño de los alumnos en esta habilidad comunicativa”. El objetivo general de la investigación es analizar los procedimientos, actividades, instrumentos y “test specifications” utilizados por los profesores de inglés para evaluar la expresión oral en inglés en un liceo municipal y su efecto que tiene en el desempeño de los alumnos.

Por lo anteriormente mencionado, me dirijo a usted para solicitar su colaboración respondiendo esta entrevista semi-estructurada. Toda la información entregada por usted será tratada de manera confidencial y su nombre jamás será revelado.

OBJETIVO: La siguiente entrevista tiene que objetivo de averiguar cómo los profesores evalúan la expresión oral para posteriormente realizar una tesis de magister.

INSTRUCCIONES: Responda las siguientes preguntas de acuerdo a su realidad en el aula o experiencia pedagógica.

1. ¿Evalúa la expresión oral en clases?
2. ¿Por qué evalúa la expresión oral?
3. ¿Con qué frecuencia evalúa la comprensión de lectura en el año?
4. ¿Con qué frecuencia evalúa la expresión oral?

5. ¿Con qué frecuencia evalúa la expresión escrita durante el año?
6. ¿Con qué frecuencia evalúa la comprensión auditiva en el año?
7. ¿Cómo evalúa la expresión oral?
8. Mencione algunos ejemplos de la evaluación de expresión oral.
9. ¿Qué aspectos o criterios evalúa en la expresión oral?
10. Ordene los criterios que utiliza para evaluar la expresión oral, antes mencionados, del más importante al menos importante. Asigne 1 al más importante.
11. ¿Cómo asigna puntaje a cada uno de los criterios mencionados? Explique por qué.
12. ¿Utiliza rúbricas? ¿Cuáles utiliza, holísticas o analíticas?
13. ¿Qué beneficios o ventajas reconoce en la utilización de las rúbricas para evaluar la expresión oral?
14. ¿Qué desventajas tiene la utilización de las rúbricas para evaluar la expresión oral?
15. ¿Cómo da retroalimentación con rúbrica?
16. ¿Con qué dificultades se encuentra cuando evalúa la expresión oral en sus cursos?
17. ¿Qué peso tiene la evaluación de la expresión oral en relación a las otras habilidades comunicativas?

APPENDIX N°5: Semi-structured interview for students about the scoring instruments used by the teachers in the speaking assessment

Entrevista semiestructurada sobre la percepción de los estudiantes acerca evaluación de la expresión oral

La siguiente entrevista semi-estructurada se utilizará en una tesis de magíster que tiene como título “Análisis de procedimientos, actividades, instrumentos y “test specifications” utilizados por profesores de inglés del liceo de la comuna de Curanilahue, para evaluar la expresión oral en inglés y su efecto en el desempeño de los alumnos en esta habilidad comunicativa”. El objetivo general de la investigación es analizar los procedimientos, actividades, instrumentos y “test specifications” utilizados por los profesores de inglés para evaluar la expresión oral en inglés en un liceo municipal y su efecto que tiene en el desempeño de los alumnos.

Por lo anteriormente mencionado, me dirijo a usted para solicitar su colaboración respondiendo esta entrevista semi-estructurada. Toda la información entregada por usted será tratada de manera confidencial y su nombre jamás será revelado.

OBJETIVO: La siguiente entrevista tiene que objetivo de averiguar la percepción que tienen los estudiantes sobre cómo los profesores evalúan la expresión oral para posteriormente realizar una tesis de magister.

INSTRUCCIONES: Responda las siguientes preguntas de acuerdo a su realidad en la asignatura de inglés.

1. ¿Su profesora de inglés evalúa la expresión oral en clases?
2. ¿Con qué frecuencia se evalúa la expresión oral?
3. ¿Con qué actividades su profesor(a) evalúa la expresión oral?
4. ¿Su profesor utiliza algún instrumento de evaluación o pauta al evaluar la expresión oral? ¿Cuáles?

5. ¿El profesor (a) evalúa algunos de los siguientes aspectos o criterios en la expresión oral: contenido, entonación, pronunciación, fluidez, memorización, expresiones vistas en clases? ¿Evalúa otros aspectos? ¿Cuáles?
6. ¿Cree que la forma de evaluar del profesor influye en la nota que obtiene? ¿Por qué?
7. ¿Cómo el profesor(a) informa su resultado o nota? ¿Le dice sus fortalezas y/o debilidades?
8. ¿En qué momento el profesor (a) evalúa la expresión oral?

APPENDIX N°6: Checklist with the information obtained from the observation made to the participant teachers

Checklist 1 to assess how the teacher of English assesses speaking

Observer's name: _____

High School: _____

Objective: to determine how the teacher assesses speaking in English and how it influences on students' performance.

Instructions: Put X in "YES" if the aspect is present and put X in "NO" if the aspect is absent.

Aspects to assess	TEACHER 1			TEACHER 2		
	YES (1)	NO (0)	Observation	YES (1)	NO (0)	Observation
1. S/he develops one or more defined assessing activities.	1		Presentation about different natural disasters	1		Dialogue
2. The activities developed in classes are different among them.		0			0	Pronunciation Memory
3. S/he uses assessment instruments:						
i) Checklist.		0			0	
j) Rating Scale.	1		Taken from the 3rd grade course book.		0	

k) Rubric.		0		1		Taken from the 2nd grade course book.
l) Others.		0			0	
4. S/he uses assessing criteria.						
m) Pronunciation.		0		1		
n) Fluency.	1			1		
o) Grammar.	1			1		
p) Linguistics skills		0			0	
q) Contents or ideas		0			0	
r) Other criteria.	1		Vocabulary	1		Interaction, task and memory. However, "memory" is not present in the rubric. The teacher takes notes on her copybook and then she adds the score below the criteria in the rubric.
5. S/he assesses in the English class.	1			1		
6. S/he assesses individually.	1				0	
7. S/he assess in group.		0		1		But the mark is individual
8. S/he uses materials as a support	1		Pictures		0	

for assessing.						
9. Just the teacher assesses.	1			1		
10. Most of the marks are in the following range.						
o) 6,5 al 7,0				7		
p) 6,0 al 6,4						
q) 5,5 al 5,9						
r) 5,0 al 5,4	4					
s) 4,5 al 4,9						
t) 4,0 al 4,4						
u) 1,0 al 3,9						
TOTAL	T1	9		T2	10	

APPENDIX N°7: Rating Scale to analyze the assessment instruments used by the teacher when assessing speaking in English with the information obtained from both teachers

Rating Scale to analyze the assessment instruments used by the teacher when assessing speaking in English

Observer's name: _____

High School: _____

Objective: to identify the “test specifications” that teachers utilized to assess speaking.

- Adequate: the test specification was completely achieved.
- Partially adequate: the test specification was partially achieved.
- Minimally adequate: the test specification was poorly achieved.
- Inadequate: the test specification was not achieved.
- Not observed: the test specification was not present.

Aspects to assess Levels	Adequate (4)		Partially adequate (3)		Minimally adequate (2)		Inadequate (1)		Not observed (0)	
	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
1. The test's purpose.	4					2				
2. Description of the examinees.									0	0
3. Test level.		4							0	
4. Definition of construct (theoretical framework for									0	0

the test).										
5. Description of suitable language course or textbook.						2			0	
6. Number of sections/papers			3			2				
7. Time for each section/paper.	4	4								
8. Weighting for each section/paper.		4							0	
9. Target language situation.	4					2				
10. Text-types.				3					0	
11. Text length.	4	4								
12. Language skills to be tested.	4					2				
13. Language elements to be tested.	4			3						
14. Test tasks.	4	4								
15. Test methods.	4					2				
16. Rubrics.		4				2				
17. Criteria for marking.				3	2					
18. Descriptions of typical performance at each level.	4	4								

19. Description of what candidates at each level can do in the real world.		4							0	
20. Sample papers.	4	4								
21. Samples of students' performance on task.									0	0
Total of 84	Teacher 1= 47					Teacher 2= 57				

APPENDIX N°8: Transcript sample of Teacher 1's semi-structured interview

Transcripción entrevista Teacher 1

1. ¿Evalúa la expresión oral en clases?

Sí, sí la evalúo

2. ¿Por qué evalúa la expresión oral?

Bueno, se supone que ... nuestra asignatura, inglés eemmm...debe ser impartida o enseñada o tratar de que los alumnos logren ehhh...habilidades y una de esas habilidades es expresarse en forma oral.

3. ¿Con qué frecuencia evalúa la comprensión de lectura en el año?

Emmm, con más frecuencia diría yo que la expresión oral. Bueno, yo diría que unas, a ver... de pende del grupo curso en realidad. Hay cursos que tienen facilidades para o facilidad para comprender la la, el idioma, otros que tienen más facilidades para expresarse en forma oral y creo que lo que se nos hace más fácil evaluar es la comprensión de lectura, así que, evalúo unas dos veces por unidad la comprensión de lectura. Cuatro, en cuarto. Es que este año solamente tengo tercero y cuarto

4. ¿Con qué frecuencia evalúa la expresión oral?

Una vez por cada unidad, o sea, cuatro veces, espero, espero poder evaluar en...

5. ¿Con qué frecuencia evalúa la expresión escrita durante el año?

Emm... después de todos los cursos que hemos tenido creo que cada vez la evalúo menos. Me enfoco más en la comprensión lectora, en la comprensión de lectura y emmm ... en la expresión oral. Eso es lo que hago, sí. También... eeehh. Escrito pero muy poco. Y tengo claro que escribir es harto más complicado que hablar. Una vez por semestre por lo menos, una vez por semestre. Igual uno ... sí poh, si tú estás pidiendo que te contesten en inglés, ya estás evaluando la parte escrita.

6. ¿Con qué frecuencia evalúa la comprensión auditiva en el año?

Comprensión auditiva, emmmm...pensando en lo que ya he hecho emmm, menos veces por muchas razones. Porque no teníamos equipos antes, porque se perdió el equipo, porque los alumnos nooo, les dan nervios, no no se concentran, en los grupos son un poco desordenados. A mí personalmente no me gusta la parte de la audición, me cuesta más mantener el orden cuando trabajamos la parte auditiva, eso sería. Igual trato de evaluarla una vez por unidad. Cuatro, pero ehhh creo que entre la expresión oral y la evaluación de la habilidad de la audición, andan por ahí en realidad. Par mí, no resulta fácil evaluarlas (03:00)

7. ¿Cómo evalúa la expresión oral?

Bueno, expresión oral. Primero, practicamos con los chicos lo que corresponda a la unidad que estamos viendo. Ehhh, le voy explicando las cosas que son importantes, la

entonación que deben usar en tal o cual oración. Eemm, intento hacer una rúbrica, que no siempre puedo compartir con mis colegas de departamento y la comparto con ellos también y así lo hago. A veces, también he pedido cooperación, ellos hacen una coevaluación con una rúbrica puede llamarse también eso, con una rúbrica a ver qué, cuántos puntos de vamos a poner a esto, qué sé yo a la entonación del compañero, a los contenidos que usó, a la fluidez que usó, etcétera. Así lo he hecho otras veces. Este año no he evaluado expresión oral todavía.

8. Mencione algunos ejemplos de la evaluación de expresión oral.

La presentación oral, personal de cada que es una de las cosas que hacemos todos los años en los diferentes niveles y también lo hago con mis colegas a los que preparo su viaje a Alemania. Ellos hacen un intento de “personal introduction and introducing others” y ahí, bueno, con los colegas lo hago así no más, o sea, sin rúbrica, sin nada. Y con los chicos, ahí nos ponemos de acuerdo en qué cosas vamos a ponerle más atención.

9. ¿Qué aspectos o criterios evalúa en la expresión oral?

Aspectos, los contenidos, la entonación, la pronunciación, no tanto la fluidez en realidad. Más esas tres diría yo.

10. Ordene los criterios que utiliza para evaluar la expresión oral, antes mencionados, del más importante al menos importante. Asigne 1 al más importante.

Ya, espérate, antes tendría que haber dicho los contenidos ¿dije? Ya ese sería pa mí el más importante, el uso del contenido. Ese sería el uno, no es cierto. Y luego, ¿qué puntaje le voy a poner a los otros si el uno es ... primero, ese es el más importante, los contenidos que estamos viendo, luego ... dije entonación y después dije pronunciación. Luego la pronunciación y finalmente la entonación, que creo que es una de las cosas más difíciles de conseguir.

11. ¿Cómo asigna puntaje a cada uno de los criterios mencionados? Explique por qué.

Cómo significa eso que... si le pongo números, si le pongo conceptos o totalmente conseguido ¿esas cosas? Con notas generalmente si usó los contenidos, no sé, se me ocurre, en realidad no no no me acuerdo como ahora cómo lo he hecho otras veces. Logrado, no logrado. Pero no como lo tienes tú aquí, no primero, otra cosa. El contenido tendría que estar entre el seis y el siete, que para mí es lo más importante. Y luego ya iría en ... descendiendo con los puntajes, hasta llegar ... es que no sé, me confunde los puntajes con los conceptos. Tendría que usar una cierta cantidad de las palabras de las que hemos visto. No, no se me ocurre en este minuto.

12. ¿Utiliza rúbricas? ¿Cuáles utiliza, holísticas o analíticas?

Pocas. No, aquí me pierdo yo. Chuta, yo creo que usaría ésta. La global ¿cuál me dijiste tú? Holística.

13. ¿Qué beneficios o ventajas reconoce en la utilización de las rúbricas para evaluar la expresión oral?

Bueno, más claro para los alumnos, a qué cosas tienen que ponerles más atención porque se supone que uno la conversa con ellos antes y para uno también. Es más, da más claridad y, claridad y objetividad, claro, en el momento de calificar.

14. ¿Qué desventajas tiene la utilización de las rúbricas para evaluar la expresión oral?

Que uno se demora mucho en hacerla, pu. Que te quede bien, o sea, tienes compartirla, preguntar y ... hasta quedar conforme con lo que está ... con lo que has hecho.

15. ¿Cómo da retroalimentación con rúbrica?

No, no, no doy retroalimentación con rúbrica.

16. ¿Con qué dificultades se encuentra cuando evalúa la expresión oral en sus cursos?

Creo que la más común es que a los niños les da vergüenza. Ellos no se atreven a expresarse oralmente. El ruido que hay en la sala, no son muy respetuosos al momento de escucharse unos a otros. Creo que esa es la dificultad mayor que tenemos. En tiempo también porque tenemos cursos numerosos y... bueno..., y si hay ruido y no se respetan y no se escuchan más y más tiempo y pasan y pasan las horas y, chuta, qué pasa con el resto mientras tú estás evaluando o tratando de evaluar.

17. ¿Qué peso tiene la evaluación de la expresión oral en relación a las otras habilidades comunicativas?

Eehh...harto porque los idiomas se hablan, después se escriben, así que yo creo que es hartito, hartito el peso que tiene.

APPENDIX N°9: Transcript sample of Student 4's semi-structured interviewed

Entrevista Student 4

1. ¿Su profesora de inglés evalúa la expresión oral en clases?

Sí

2. ¿Con qué frecuencia se evalúa la expresión oral?

La evalúa generalmente cuando hace test, lo primero que hace es una evaluación de expresión oral. (*¿Eso cuántas veces al semestre?*) Es cada vez que tenemos evaluación ... como que va por pasos ... evaluación de expresión oral, luego el test, luego la prueba (*¿y eso cuántas veces ocurrió el primer semestre?*) Cuatro, las veces que tenemos prueba, tres ... cuatro.

3. ¿Con qué actividades su profesor(a) evalúa la expresión oral?

Aprendiéndonos letras de canciones, fue una vez eso ... y el resto son diálogos.

4. ¿Su profesor utiliza algún instrumento de evaluación o pauta al evaluar la expresión oral? ¿Cuáles?

(*Estos son los instrumentos: rúbrica, escala de apreciación, lista de cotejo*). No

5. ¿El profesor (a) evalúa algunos de los siguientes aspectos o criterios en la expresión oral: contenido, entonación, pronunciación, fluidez, memorización, expresiones vistas en clases? ¿Evalúa otros aspectos? ¿Cuáles?

Contenido: sí

Entonación: no

Pronunciación: sí

Fluidez: sí

Memorización: sí

Expresiones vistas en clases: sí

(*¿Evalúa otros aspectos?*) No.

6. ¿Cree que la forma de evaluar del profesor influye en la nota que obtiene? ¿Por qué?

No. Porque siempre nos evalúa de la misma forma, entonces ella, en vez de tener la pauta escrita, nos da las instrucciones antes entonces ... no lo hace con una pauta.

7. ¿Cómo el profesor(a) informa su resultado o nota? ¿Le dice sus fortalezas y/o debilidades?

Sí. (*¿Cómo les informa las notas? ¿se las dice no más?*) No, cuando nos llama nos dice en qué nos equivocamos.

8. ¿En qué momento el profesor (a) evalúa la expresión oral? (*¿En clases?*) Sí.