UNIVERSITE MOULOUD MAMMERI DE TIZI-OUZOU
FACULTE DES LETTRES ET DES LANGUES
DEPARTEMENT D'ANGLAIS

جامعة مولود معمري- تيزي وزو
كلية الآداب واللغات
قسم الإنجليزية

Domaine : Lettres et Langues Etrangères

Filière : Langue Anglaise

Spécialité : Langue et Communication

Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Master in English
Title:

Investigating the Reliability of Speaking Tests/Exams in the Department of English – MMUTO :
The case of Third-Year Classes.

Presented by: **Supervised by:**

Ms. AOUINE Soraya   Dr. GADA Saïd

Ms. BAKHOUCHE Dehia   Mr. HAMI Hamid

Board of Examiners:

Chair: Mme Adem Karima, MAA, Department of English, Mouloud Mammeri University of Tizi-Ouzou

Supervisor: Dr. Gada Said, MCB, Department of English, Mouloud Mammeri University of Tizi-Ouzou

Co-supervisor: Mr. Hami Hamid, EA, Department of English, Mouloud Mammeri University of Tizi-Ouzou

Examiner: Mme Benaissa Amel, MAA, Department of English, Mouloud Mammeri University of Tizi-Ouzou

June 2018

№ D'Ordre:

№ De série:

Dedications

To my beloved parents

To my precious brothers and sisters

To my lovable nephew and nieces

To my dear Ahmed and his family

To all my friends, with whom I shared precious moments.

Aouine Soraya

....

To my father and mother

To my beloved family

To my lovable brothers and aunts

To my dear Hamza

To all my friends.

Bakhouche Dehia

Acknowledgements

Abstract

The present research work aims at exploring the reliability of speaking assessment tools used in the English Department at Mouloud Mammeri Univesity of Tizi-ouzou. The overall objective is to check the reliability of 'speaking' tests, starting from the principle that the more accurately a score represents a student's true language ability, the more reliable a test will be. The study is conducted relying on Lyle Bachman's theory referred to as 'Communicative Language Ability' (1990), which is the combination of both communicative competence and language proficiency. For this, data are collected through a questionnaire administered to twenty teachers of Speaking/Listening subject of study, and an analysis of speaking exams. The research is conducted using a mixed method approach to analyze both qualitative and quantitative results. The findings reported in this paper indicate that the speaking tests are reliable, since students who pass these tests can cope with similar communication situations in real life. Moreover, it is deduced that many factors can affect negatively on students' performance. To conclude, some recommendations are suggested at the end of this dissertation.

Keywords: Speaking exams, reliability, Communicative Language Ability, real-life situations, communicative competence.

List of Abbreviations

CC : Communicative Competence

CLA : Communicative Language Ability

FL: Foreign Language

IELTS : Inernational English Language Testing System

L2: Second Language

MAA : 'Maître Assistant classe A'

MAB : 'Maître Assistant classe B'

MCA : 'Maître de Conférences classe A'

MCB : 'Maître de Conférences classe B'

OC : Organisational Competence

QDA : Qualitative Discourse Analysis

SPSS : Statistical Package of Social Sciences

List of Diagrams

List of Figures

Table of contents

General Introduction

Chapter 1: Review of the Literature

Theoretical framework

Chapter 2: Research Design

Chapter 3: Presentation of the Findings

Chapter 4: Discussion of the Findings

GENERAL
INTRODUCTION

General Introduction

Statement of the Problem

Traditionally, the major goal of EFL teaching is to help learners to develop their linguistic competence to master: lexis, sounds, grammar patterns and sentence structure of the target language. Nowadays however, as Frances Butter and Robin Stevens put it, *"being communicatively competent entails being able to adjust one's language to specific audiences and situations"* (1997: 216). Since teaching and testing are closely interrelated processes, our endeavor through the present research work is to explore the way students are tested in speaking skills, focusing on the reliability of the measurement tools used to assess those skills. The test/exam in the 'listening/speaking' subject at the Department of English of MMUTO constitutes a high-stakes test, as it should determine whether students are able to hold a conversation in English in real-life communication situations, academic and professional contexts.

Assessment should enable the evaluation of the extent to which students have learned and the extent to which they can demonstrate that at learning (Brown and Smith, 1998). Accordingly, with the rise of the communicative approach, the role of the speaking ability has become more prominent in language teaching. As a result, performance testing, especially testing the speaking abilities, has become one of the most important issues in language testing. Oral assessment, then, provides rich opportunities for student's engagement as well as for enhanced judgment about students' achievement. Huxham *et al* state that *"the use of oral assessment motivates students to practice and improve their English speaking skills"* (2012: 45); that is to say, oral tests are considered as a type of performance tests to assess language proficiency through interaction and productivity as Wigglesworth has put it : *"in the context of language testing and assessment, performance assessment has become increasingly important*

over the last two decades, and has been the focus of substantial empirical investigation" (2008: 111).

Reliability has to do with the extent to which scores are consistent (Brown and Hudson, 2002; Henning, 1987; Luoma, 2004). It is important because if the results are dependable, we cannot rely on them in decision making, and if test scores are not reliable they can have negative impact on the students' learning process, motivation and future, such as wrong placement, unjustified promotions, or undeserved high or low grades (Luoma, 2004). Reliable assessment tools produce dependable, repeatable, and consistent information about learners. In order to meaningfully interpret test scores and make useful learning-related decisions, teachers need reliable tools.

According to Brown & Yule *"a listener's task performance may be unreliable for a number of reasons. We have only a very limited understanding of how we could determine what it is that listening comprehension entails"* (1983: 104). There are, of course, factors that can threaten the reliability of a test. Some of them are variation in administration, quality of the test, differences in test forms, changes in test takers over time, differences in scoring and differences in scorers (Fulcher, 2010). Furthermore, Lado (quoted in Fulcher, 2010) claims that no test is a perfect measure. For example the problem of how to choose what to be tested, since everything cannot be tested in one test. Moreover, he notes that if a test item includes very different things, it would reduce its reliability. Finally, he highlights that unreliability can be caused by the scoring (Fulcher, 2010). These elements that can threaten or influence reliability are often referred to as measurement errors, and are sources of inconsistency in test scores that will affect their consistency (Bachman, 2004).

Reliability may be considered as the major criterion to determine the quality of a language assessment tool. It refers to how dependably or consistently a test measures a

characteristic. If a student takes the test again, will he get a similar test score, or a much different score? A test that yields similar scores for a student who repeats the test is said to measure a characteristic reliably. Reliability is the focus of the present study, which seeks to scrutinize this very aspect of the tests set for 3rd-year students at the English Department of MMUTO to measure their oral communicative competence. However, our parameters do not lie in the test/re-test experiments, but they are related to real life communication situations. In other words, we consider a test as reliable when the student's abilities shown in tests are likely to be repeated in communication situations in social and professional contexts.

To our knowledge, none of the previous research works bearing on assessment, in general, and reliability, in particular, has endeavored to investigate speaking tests set to 3rd-year students at the English Department of MMUTO. Thus, any student's ability to perform certain tasks in the real world is then estimated by drawing inferences from their performance in the test, in other words, by generalizing it.

Aims and Significance of the Study

The present research work attempts to check whether the speaking tests in the English Department at MMUTO are built in a way to give consists, since reliability is the unavoidable criterion, an aspect that needs consideration when designing a test. The main objective of the present research work is to assess the acceptability of the level of reliability of these tests on the basis of four major factors: Students' temporary psychological or physical state; environmental factors; test form and scorers (teachers). Yet, the peculiarity of the present research work lies in the fact that we consider a speaking test as reliable when its results are consistent with real-life communication situations, not with the results that the test takers would get if they take the same test again sometime later.

When considering reliability in testing, we believe that it is the crucial facet of assessment, because without it the test results cannot be valid, since reliability and validity are closely interconnected: a test cannot be valid if it is not reliable. The tests that are low in reliability can lead to incorrect diagnoses and false conclusions. Accordingly, the significance of the present study lies in its impact on the assessment of the speaking subject at the English Department of MMUTO. Indeed, it seeks to determine whether there is a need to use more reliable tests to measure MA students' ability to tackle, for instance, a real oral presentation for the requirements of the Masters degree, in which they are expected to speak around 20 minutes along, and listen to and answer jury members' questions.

Research Questions and Hypotheses

the present study intends to address the following research questions.

Q1. Are the speaking tests under scrutiny reliable?

Q2. Can students who pass these tests cope with similar communication situations in real life?

Q3. What are the factors that affect students' performance during speaking exams?

The following hypotheses try to predict the possible answers:

H1. The speaking tests are reliable.

H2. The speaking tests are not reliable.

H3. Students who pass these tests can cope with similar communication situations in real life.

H4. Students who pass these tests cannot cope with similar communication situations in real life.

H5. There are several factors that affect student's performance during speaking exams.

Research Techniques and Methodology

Most of the analytic categories used in this research work are borrowed from Lyle F. Bachman's theory on language testing displayed in his *Fundamental Considerations in Language Testing* (1990). Besides, in an attempt to reach the aims stated above, we have adopted a mixed-method approach combining both quantitative and qualitative methods for data collection and data analysis. On the one hand, statistical data have been collected through a set of tests given to 3rd-year students in previous academic years to check their degree of reliability. On the other hand, a questionnaire has been designed for and administered to speaking teachers of the Department to identify their perception of the importance of reliability in test design.

For the data analysis, a descriptive statistical method relying on SPSS (Statistical Package for Social Sciences) program is used in order to make the analysis of the data easier. In addition, a Qualitative Content Analysis is adopted in order to interpret and explain the results obtained from the corpus and the open-ended questions of the questionnaire.

Structure of the Dissertation

The research paper is designed and structured following the traditional simple style of dissertation format. It comprises a 'General Introduction' and four main chapters: 'Review of the Literature', 'Research Design and Methodology', 'Presentation of the Findings', 'Discussion of the Findings', and a 'General Conclusion'.

The General Introduction is devoted to topic generalization and claiming centrality by displaying the aims and significance of the work. It also puts forward the research questions and hypotheses. The first chapter, named 'Review of the Literature', includes some mainstream definitions of the key concepts of the research topic, relying on the works of different scholars. The second chapter, entitled 'Research Design and

Methodology', displays the procedures used in data collection and data analysis. The third chapter, 'Presentation of Findings', displays the results of the investigation. As for the fourth and the last chapter, labeled 'Discussion of the Findings', it is devoted to our main contribution in the present research work, since it comprises our own interpretation and discussion of the findings of the study. Finally, in the 'General Conclusion', we shall summarize the main points of the study and suggest designing techniques to make oral assessment as reliable as possible.

LITERATURE REVIEW

Review of the Literature

Introduction

This chapter reviews the existing literature relevant to the present study and provides a justification for conducting the study, since it has gained the interest of many researchers. It involves four major sections. The first section reviews definitions about oral assessment, in general, and speaking skills, in particular. The second section presents the term of reliability and the factors that may affect test scores. The third section is entitled communicative competence and communicative language testing; it includes a historical overview of communicative competence relying on other researchers. The last section deals with the theoretical framework focusing on the Communicative Language Ability model suggested by Lyle F. Bachman (1990).

1. Oral Assessment

1.1 Definition

Assessment is used as a general umbrella term to cover methods of testing and evaluation. It is the process of collecting and organizing information from purposeful activities such as tests on performance or learning to draw inference about teaching and learning. Thus, assessment is the systematic use of information about educational programs in order to improve learning and development (Palomba and Banta 1999).

1.2. Types of Assessment

William (2008:59) asserts that assessment in general accounts for *"supporting learning (formative), and certifying the achievement or potential of individuals (summative)"*. Hence, formative assessment is an ongoing process whereby teachers use feedback over the course to meet students' need and diagnose their progress (black and William, 1998); in other words, it refers to the kind of assessment that provides detailed

and specific information about students' abilities and struggles, that will help teachers to adapt their learning and teaching methods (Hyghes,2003; Stoynoff,2005). Therefore, summative assessment is to test students' competence at the end of the teaching and learning process. It helps teachers to make judgments about students' achievement at certain relevant points in the learning process, as Brown states *"it deals with the exams which take place at the end of a semester or a year of study" (2003:5).*

1.3. **Criteria of Oral Assessment**

There are several aspects of speaking that influence how a person's oral proficiency can be evaluated, that are pronunciation, grammar, vocabulary, in addition to slips and errors.

Pronunciation is the ability to produce individual sounds and to link words together, as well as using stress and intonation to communicate meaning (Thornbury, 2005:128-129). According to Luoma (2004:10-11) speech sound is difficult to assess because there are a number of dialects and variations in one language especially English. So, students are able to learn clear and understandable pronunciation, without sounding at all like a native speaker. Therefore, she proposes two elements that can be in focus; the accuracy of pronunciation and expressiveness of the speaker's use of voice.

Grammar is the building blocks in a language, which constructs the language form, Bachman and Palmer (1996) state that grammatical knowledge contains vocabulary, syntax, and morphology. Besides, Grammar is an element that influences the way the scorer evaluates spoken production. Accordingly, Spoken language is different from written language, thus spoken language differs greatly in different contexts, depending on whether the situations of speech are planned or unplanned, formal or informal (Luoma, 1996).

Vocabulary is the ability *"to express oneself precisely and providing evidence of the richness of one's lexicon"* (Luoma, 2004:16). In test situations where vocabulary is a criterion for oral assessment both knowledge of, and the use of vocabulary is assessed (Bachman and palmer, 1996:276), hence lexical competence and accuracy of vocabulary are required.

Additionally, Slips and errors are natural, that there are errors in spoken language such as mispronunciations, the use of wrong words or mixed sounds. However, when EFL learners speak with such errors, this is often considered lack of competence; Assessors, then, should not take in consideration all the mistakes they hear, since this is a natural part of spoken language even for native speakers.

2. Communicative Competence and Communicative Language Testing

The history of language testing has been divided by various researchers. According to Spolsky (1989) language tests involve measuring subject's knowledge, and proficiency in the use of a language, he identifies three eras of language testing: pre-scientific era, psychometric era, and psycholinguistic-sociolinguistic era.

2.1. Pre-scientific Era: it is a period where testing is considered as a traditional consideration of test on phonetics, morphology and grammar, it is characterized by *"the use of essays, open-ended examinations or oral examining with the result determined intuitively by an authorized and authoritarian examiner"* (Spolsky 1995:353). So testing during this period does not rely on linguistic theory and reliability was not important.

2.2. Psychometric-structuralism Era: this era tries to join the structural approach, which identifies specific elements of language to be tested, and the psychometric approach, which produces objective and reliable methods of testing these elements. Thus, each test item is intended to give information about student's ability, since the focus has shifted from the language system itself into the language ability of candidates.

2.3. Psycholinguistic-sociolinguistic Era: testing aims to measure the candidate's competence rather than his/her performance. Dell Hymes (1972) introduces a theory of communicative competence to register his dissatisfaction with Chomsky's (1965) terms: competence and performance; he defines it as *"a knowledge of the rules for understanding and producing both the referential and social meaning of language"* (1972, p.270). In other words, communicative competence describes the knowledge and ability of individuals for appropriate use in the communicative events and concrete situation; hence it includes both linguistic competence and explicit knowledge of grammar and contextual background of the rules. Hymes, then, views CC as having the following four types: what is formally possible, what is feasible, what is the social meaning or value of a given sentence, and what actually occurs to fulfill a successful communication.

3. Speaking Skill

Speaking is described by Fulcher (2003) as the ability that makes us human by producing language. In learning process, it is important that learners of a foreign Language produce spontaneous and automatic speech in the target language; however speaking a foreign language is difficult for learners, because effective oral communication depends on learner's fluency and language use in social interaction. Therefore, it is quite possible for learners to make mistakes when speaking can be failed with hesitations, false-starts, grammatical inaccuracies, and limited vocabulary (Hughes, 2002:77).

3.1. Characteristics of Speaking Skill

According to Mazouzi (2013) learners' activities should be designed based on equivalence between fluency and accuracy achievement, which are both important elements of the communicative approach. The first characteristic of speaking performance is fluency, which is the main aim of instructors in teaching speaking

skills. According to Hughes (2002), fluency is the learners' ability to speak in an understandable way in order not to break down communication because listeners may lose their interest. Hedge (2000) claims that fluency is the ability to answer coherently by connecting words and phrases, pronouncing sounds clearly, and using stress and intonation. The second characteristic of speaking performance is accuracy, where learners should pay enough attention to the exactness and completeness of a language form when speaking a foreign language, such as grammatical structure, vocabulary, and pronunciation (Leong and Seydeh, 2013). Hence, in order to speak English accurately, learners should master its phonological rules and be aware of the various sounds and their pronunciation.

3.2. Factors Influencing Learner's Speaking Competence

Students speaking competence can be affected by the factors that come from competence conditions, since communicative competence includes not only linguistic competence but also a range of socio-linguistics and conversational skills.

3.2.1 Cognitive Factors: according to Levelt (1989) speaking processes include conceptualization, formulation and articulation. Conceptualization deals with the way meaning is expressed using specific information; formulation deals with the choice of words used in appropriate grammatical structure; articulation is the production of speech with articulation organs.

3.2.2 Linguistic Factors: the correct use of a foreign language needs the accurate pronunciation, vocabulary and grammar. Pronunciation includes the ability to produce sounds to link words as well as using stress and intonation to communicate meaning. Luoma (2004) states that people tend to status of native and non-native speaker based on their pronunciation. Therefore, grammar contains the structure of any language, as it is difficult for learners to transfer the correct grammar when

speaking, so for assessing grammatical accuracy in oral communication it is relevant to establish to what extent students manage to apply the rules such as the correct form of verbs, adjectives and plural of nouns. In addition, Vocabulary is a broad lexical repertoire, which involves idiomatic expressions; accordingly, Simensen (1998) argues that the luck of sufficient vocabulary is the main obstacle to using language.

3.2.3 Affective Factors: Oxford Rugby states that *"the affective side of the learner is probably one of the most important influences on language learning success or failure"* (1990:140). Hence, Affective factors include emotions, self-esteem and anxiety; and worrying being *"wrong, stupid or incomprehensible"* (brown 2001:269) affects learners speaking performance and fluency by causing tongue-tied and absolutely influences their achievement in speaking tests.

4. Reliability

4.1 Definition: Reliability refers to how dependably or consistently a test measures characteristics, so it is the consistency of the score and rank order of test takers from a test when it is administered in different situations (Bachman and Palmer 1996:19-20). However, Bachman argues that in order to measure test takers' language ability in similar language use situations, the design of the test tasks should be made as similar as possible. Besides, to ensure test reliability, the test designer controls and compares learners' performance and their topical knowledge relying to the existing accordance between performing tasks and speaking in a real-life situation (Bachman and Palmer, 1996).

4.2.Factors Affecting Language Test Scores

A number of studies have been conducted in different contexts to investigate the factors which may affect test candidates' performance on language tests. Bachman designs a diagram to represent the four categories that affect language testing scores.
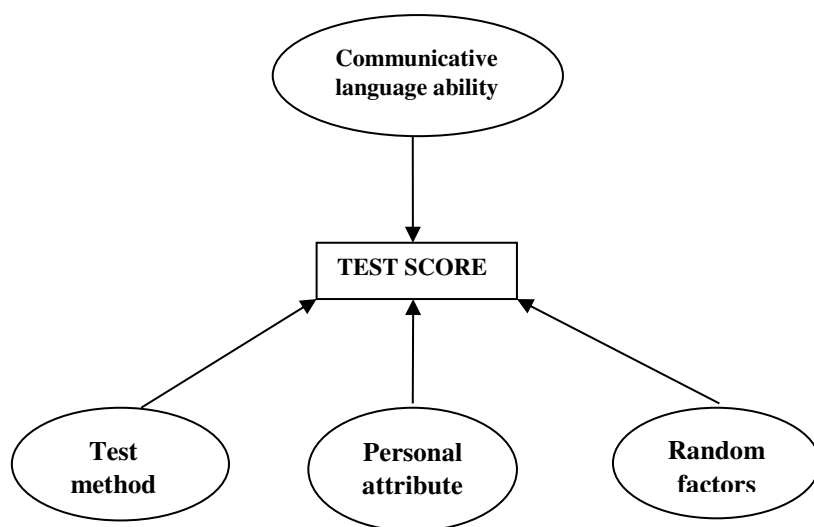
Figure1: factors that affect language test scores (Bachman, 1990:165)

4.2.1. Communicative Language Ability: the basic factor which comprises of language competence, strategic competence, and psycho-physiological mechanism, where scores are decided by the language ability of the candidate; in other words, if the candidates take the same test many times and their performance is not affected, their language ability is estimated through their true scores.

4.2.2. Test Method Facets: refer to the characteristics of the test's tasks used to obtain test performance and test score variation. This factor includes the test environment (physical conditions and time of testing); test rubric (task construction); task input (language used in the input); the expected response (nature of language used in the input); and the interaction between the input and the response.

4.2.3. Personal Attributes (test-taker characteristics): contain background characteristics (age, gender, ethnic identity and educational background); socio-cultural characteristics (attitudes and motivation); psychological and physical state (anxiety, fatigue and self-esteem); cognitive characteristics (aptitude and learning strategies).

13

4.2.4. Random factors: include interactions among components of communicative language ability test method facts and personal characteristics, as well as measurement errors. Hence, random factors refer to unpredictable and temporary conditions such as, students' mental and emotional state, in addition to uncontrolled differences in test method facets, like changes in the test environment from one day to the next.

4.3. Feedback: it provides information about candidates' performance after an assessment. Thus, it describes students' achievement, strengths and weaknesses. In this respect, Fulcher (2010) states that feedback enables the students to develop their skills in the target language by making them notice their mistakes and lead them towards the correct features of a language.

Theoretical Framework

The purpose of the present study is to check the reliability of listening/speaking tests and try to find the factors that affect the students' performance. Therefore, to attain our objectives, we rely on Bachman's theory called 'Communicative Language Ability' (1990).

1. Definition of Communicative Language Ability

The American linguist Bachman has developed an influential theory called 'Communicative Language Ability' (CLA) in 1990, and subsequently refunds it with Palmer. The perspectives of CLA are built on the previous theoretical models, for instance (Canale and Swain, 1980; Hymes, 1972; Munby, 1978; Widdowson, 1978; Saignon, 1972), as their primary concern is to explain factors affecting L2 performance on language tests. Consequently, CLA is described as *"consisting of both knowledge, or competence, and the capacity for implementing or executing that*

competence in appropriate, contextualized communicative language use" (Bachman, 1990:84).

2. The Components of CLA

The construction of CLA is based on three core components: language competence, strategic competence, psycho-physiological mechanism, all coordinating with the knowledge structure in situation context to depict an overall picture of CC. Thus, Bachman describes knowledge structure as the language user's socio-cultural knowledge in real-world.

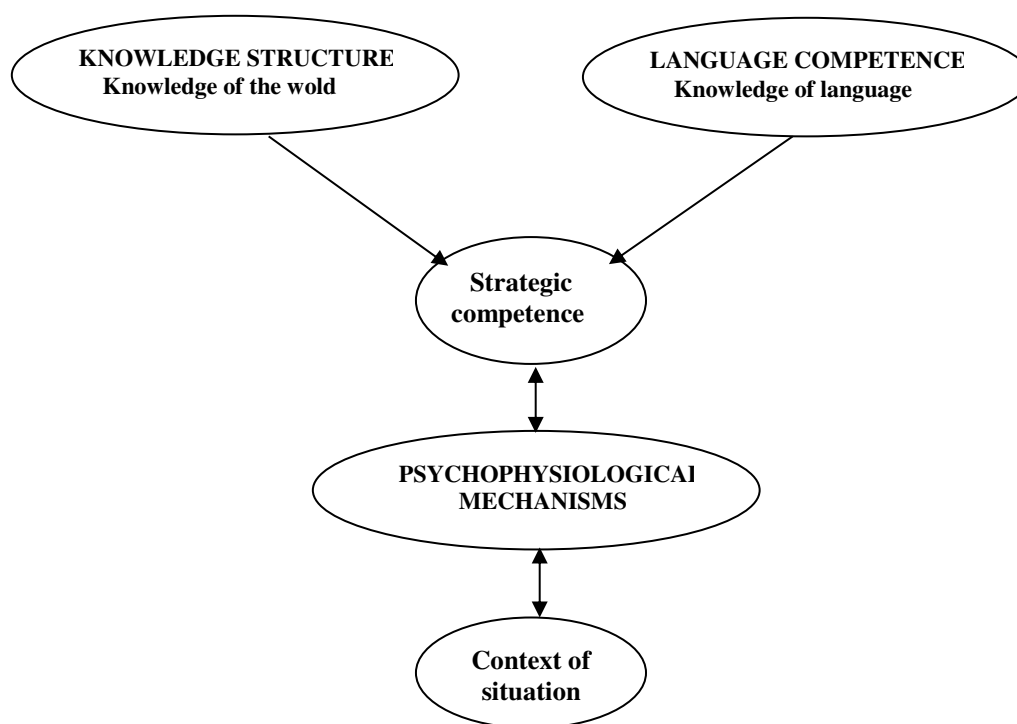**Figure2: Components of CLA in communicative language use   (Bachman, 1990:85)**

2.1.   Language Competence: Bachman (1990) categorizes language competence into organizational competence and pragmatic competence.

2.1.1. Organizational Competence: it involves the abilities in controlling the structure of language, by producing correct sentences to form coherent texts. OC comprises

15

grammatical competence, which is related to the capability of mastering grammatical ranging from vocabulary, morphemes, syntax, phonology/graphology, and textual competence that refers to cohesion and rhetorical organization.

2.1.2. Pragmatic Competence: it is concerned with the use and organization of linguistic signals in a particular context; hence it relates the language users and communicative situation. Pragmatic competence is categorized into two main subcomponents:

2.1.2.1 Illocutionary Competence: it is derived from the theory of speech act of Austin (1962) and Searl (1969), which is the basis for performing language functions. It refers to the intention of the language users in producing specific type of utterances. Illocutionary competence is further classified into four functions: ideational function (expressing meaning relying on language users knowledge in real-world situations); manipulating functions (the objective is to affect others by making suggestions, requests and commands); heuristic functions (occurs when the knowledge is shared such as teaching and learning); imaginative functions (used for creating language function for humorous or aesthetic purposes such as jokes, metaphors and writing stories).

2.1.2.2 Sociolinguistic Competence: refers to the ability to use the appropriate language in a specific communication, it is defined by Bachman as *"the sensitivity to, or control of the conversation of language use that are determined by the features of the use context"* (1990:94). He introduces sociolinguistic competence referring to four major abilities: differences in dialect or variety (means variation in language use which is due to different social groups); differences in register (a terminology used by language users to achieve specific communicative goal); naturalness (production and comprehension of language correctly); and interpretation of cultural references and figures of speech such as metaphors and similes.

Accordingly, Bachman presents strategic competence as having the central role in his modal; it characterizes the mental capacity of the user including memories, past experiences or topical knowledge.

The following figure represents the components of language competence:

```
                        ┌──────────────┐
                        │  Language    │
                        │  competence  │
                        └──────────────┘
                         ↙            ↘
            ┌──────────────┐      ┌──────────────┐
            │ Organizational│      │  Pragmatic   │
            │  competence  │      │  competence  │
            └──────────────┘      └──────────────┘
             ↙          ↘          ↙            ↘
```

Grammatical competence	Textual competence	Illocutionary competence	Sociolinguistic competence
-Vocabulary - Morpheme - Syntax -Phonology/ Graphology	-Cohesion -Rhetorical Organisation	-Ideational functions -manipulative functions -regulatory functions – interactional func	-Sensitivity to differences in dialect or variety -sensitivity to differences in register -sensitivity to naturalness -ability to interpret cultural references and figures of speech

Figure3: Subcomponents of language competence in the CLA model (Bachman 1990:87)

2.2. Strategic Competence: it is defined as a set of meta-cognitive strategies that explain the interaction of the knowledge and affective components of language use. As Bachman states *"the mental capacity for implementing the components of language competence in a contextualized communicative language use"* (1990:84); hence it is affected by the knowledge structures (world knowledge) of language users.

2.2.1 Assessment Components: integrate the ability to select a particular language to be used for realizing communicative effectiveness in a given context, and evaluating the achievement of successful communication. According to Bachman, assessment components provide a means by which individuals relate their topical knowledge and language knowledge to the language use, setting and tasks (1990).

2.2.2 Planning Components: enable communicators to formulate a plan in realization of communicative purpose with certain language knowledge. Planning, then, is a key meta-cognitive strategy for second language acquisition, involves in directing the course of reception and production.

2.2.3 Execution Components: Bachman states that execution components draw on the relevant psycho-physiological mechanism, to implement the plan in the modality and channel appropriate to the communicative goal and the context (1990).

2.3. Psycho-physiological Mechanisms: involve the neurological and physiological processes responsible for the actual execution of language use, which make language users perceive the mental production of utterances before they are performed. Therefore the processes include channel (visual/auditory) and mode (receptive/productive). In receptive mode, the language user uses his/her auditory and visual skills to get instructions and non-linguistic information. In productive mode, language user shows his/her articulatory skill to pronounce words and utterances with correct stress and intonation; hence psycho-physiological mechanisms comprise the factors that make language competence and strategic competence into practice to achieve a specific communicative goal in a specific context.

Bachman's theoretical framework proposes that CLA is a combination of language knowledge and the ability to put that knowledge in communicative situations (Bachman 1990).

Conclusion

In this chapter, we have presented an overview of concepts related to the present study. Thus, previous research studies on the issue of reliability of oral assessment have been highlighted. In the first part, a review of oral assessment and speaking skills has been covered. Then, the term of reliability and the factors that affect language tests scores have

been discussed. Next, an overview of the communicative competence notion has been attempted, relying on models of some scholars, since our specific concern in the present study is to measure the consistency of tests in relation to real communication situations. As for the last part of this chapter, it has been devoted to the theoretical framework of the research work.

RESEARCH DESIGN

Research Design

Introduction

This chapter depicts the collected methods and materials that have been used during the research process. It shows the research techniques used to investigate the reliability of speaking tests/exams. Moreover, it is devoted into two main sections that deal with the data collection and analysis. The first part deals with the description of the research methods and tools under the investigation; the second section explains the procedures of data analysis, by using the descriptive method that involves quantitative and qualitative approaches in order to link the research questions and objectives.

1. Procedures of Data Collection

1.1. Research Methods

This study is designed to examine the reliability of speaking tests implemented in the English department at MMUTO. This work adopts both descriptive and explanatory methods. The descriptive research is defined as the gathering of information about prevailing conditions or situations for the purpose of description and interpretation (Agrawal, 2008); it refers to the type of research design and data analysis that is applied in this topic. Furthermore, explanatory research is intended to explain and interpret the results obtained about the phenomenon being studied. Accordingly, a mixed method is used in order to link the research questions and hypothesis with the data collected. In this context, Johnson and Onwvegbuzie state that, '*mixed research involves mix[ing] or combine[ing] qualitative and quantitative research, techniques, methods, approaches, concepts, or language in a single study'* (2004:17).

1.1.1. Participants and Context of the Study

This study is carried out in a realistic context; that is, the English department at MMUTO. The participants of the present study consist of twenty teachers of speaking/listening module. The selection of such sample is based on the consideration that the teachers of speaking provide the research with more reliable data, since they teach students how to develop speaking skills. Hence, the target population includes the teachers of 1st, 2nd, and 3rd years. Moreover, to gather clear information about the participants' characteristics, they are asked to respond to questions related to their professional position, and teaching experience.

1.2. Data Collection Tools

In order to collect data for the current research, an analysis of speaking exams is highlighted, and a questionnaire is administered to teachers of speaking/listening module; it is a distinctive research technique and a flexible tool for data collection in getting insight information from the subject.

1.2.1. Description of the Teachers' Questionnaire

A questionnaire is a structured form of a formalized set of questions, designed to collect information from specific respondents. This data collection technique can be classified as both quantitative and qualitative methods, depending on the designed questions. Furthermore, a questionnaire is a research instrument which guaranties the anonymous of the participants.

For the sake of gathering reliable data, a questionnaire is distributed to 26 teachers of listening/speaking module, yet 20 have been retrieved. The instrument used is made of 20 questions which are divided into close-ended and open-ended questions; meanwhile, it

is composed of three sections, the first section includes personal background which aims at determining teachers profile (teaching experience and qualification); the second section involves foreign language assessment and speaking tests design; the third section includes factors affecting student's speaking performance and speaking problems.

1.2.2. Corpus

Linguistic corpus is the study of language based on collections of its use in real-life situations; it is used to investigate the questions related to the speaking tests. So, twenty three speaking tests have been chosen as a corpus to be analyzed in the present study. Hence, teachers design different topics to test their students by following the syllabus, so that, the models of questions include: the explanation of the tenses, asking for advice, permission, request…, or providing the right meaning of idioms through giving concrete examples. Accordingly, the students may be also asked to speak about a specific phenomenon such as globalization, women's leadership, and ethics as free topics; by using sociolinguistic background.

2. Procedures of Data Analysis

The reliability of a study can be maintained through manners of analyzing and presenting the collected data (Dörnyei; 2008). Though, in the present study, the answers of the questionnaire are followed up by the analysis of speaking tests/exams, in order to provide the research with more reliable results.

2.1. Statistical Analysis

In the questionnaire, different types of questions are included (close-ended and open-ended questions); hence, the data collected are analyzed by using quantitative and

qualitative methods. Besides, the close-ended questions (numerical data) are analyzed using computer program named SPSS.

SPSS, then, is an acronym of Statistical Package for the Social Scientists, it is a data management and statistical analysis tool with a very versatile data processing capacity; "*it is the comprehensive system for analyzing data. SPSS can take data from almost any type of file and use them to generate tabulated reports, charts, plots of distribution and trends descriptive statistic and complex statistical analysis*" (based users' guide). Accordingly, SPSS is a Windows program that can be used to perform data entry and analysis to create tables and graphs. Furthermore, the findings of the research are highlighted by means of histograms and pie-charts.

The statistical SPSS software, is applied to calculate the collected data; following the guidelines of that process. First, the data are imported into SPSS after completing the variable view and the data view; second, specific commands are served to the software then the results are given efficiently and accurately by retrieving them; finally, the graphs and charts are analyzed in order to postulate conclusions.

2.2. Qualitative Content Analysis

The qualitative content analysis is defined as, *"an approach of empirical, methodological, controlled analysis of texts within their context of communication, following content analytic rules and step by step models, without rash quantification"* (Mayring, 2000:2). That is to say, QCA consists of extracting objective content from texts to examine and interpret meaning and themes, it allows researchers to understand social reality to support valid and reliable inferences. QCA involves a set of systematic procedures for processing data, beginning with preparing the data and proceeding through

writing up findings of a report. In the present study, QCA is applied to examine and interpret the open ended questions of the questionnaire and the speaking tests.

Conclusion

This chapter has described the methodological procedures used in the study. It has presented specific methods for collecting and analyzing data gathered from the research tools (a questionnaire and an analysis of speaking tests). Indeed, the computer software SPSS is used as a statistical technique to interpret the quantitative data obtained from the questionnaire; furthermore, qualitative content analysis is adopted to analyze the open-ended questions of the questionnaire and speaking tests/exams. Hence, the data collection and analysis permit to confirm or disconfirm the research questions and hypothesis presented in the general introduction.

PRESENTATION OF THE FINDINGS

Presentation of the Findings

Introduction

The present chapter is devoted to the presentation of the findings obtained by means of a questionnaire addressed to teachers and an analysis of the tests designed to measure students' performance in speaking. The results derived from the questionnaire are presented in percentages and displayed in histograms and pie-charts, supplemented by comments and explanations. For that, this chapter is divided into two main sections, containing quantitative and qualitative results. The quantitative section shows the findings of the questionnaire, and the qualitative section deals with the presentation of the results obtained from the analysis of speaking tests/exams.

Teachers' Questionnaire

In the present research, teachers of Speaking/Listening module are handed a questionnaire. The results are based on the answers of 20 participants; thus, the material used consists of a total of 19 questions: 2 are open-ended questions, 5 are yes/no questions, and 12 are close-ended ones. In fact, the results obtained from the questionnaire are reported bellow:

Section one: Personal background

Q1: Professional position

☐ MAB   ☐ MAA   ☐ MCB   ☐ MCA   ☐ prof
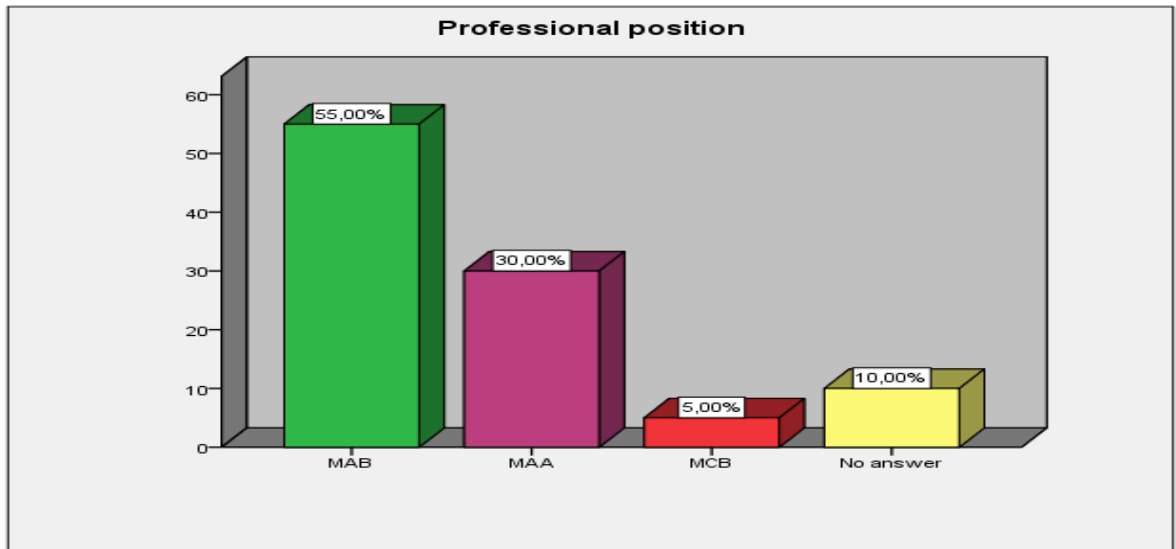
Diagram 1: Professional position

This graph presents the professional position of teachers of speaking/Listening module, where the majority of them are classified in the first category (MAB) with the percentage of 55, and 30% of the respondents are rated in the second category (MAA). As one teacher has the MCB grade; whereas, two participants did not answer the question.

Q 2: Teaching experience (in years)

☐   5   ☐   6-10   ☐   More than 10

Diagram 2: Teaching experience (in years)

As displayed in this diagram, the experience of Speaking/Listening teachers is arranged from two (2) years to more than ten (10) years. Yet, two teachers do not answer the question.
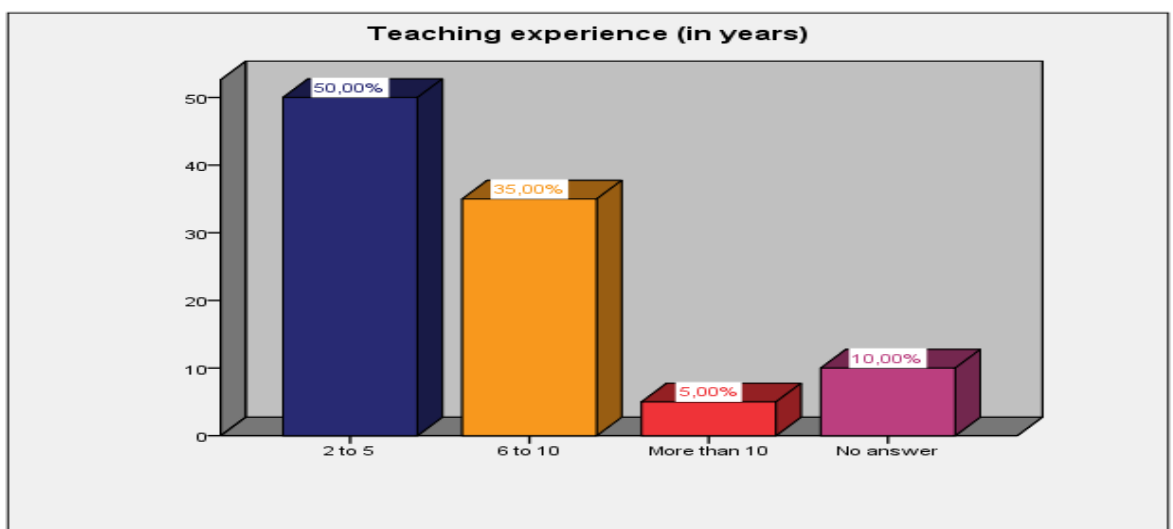
Section Two: Foreign language assessment and speaking test design

Q 3: to which approach would you relate your teaching techniques?

☐ Communicative Approach   ☐ Grammar-translation Approach

☐ Competency-based Approach   ☐ Eclectic Approach

Diagram 3: Approaches used in teaching process
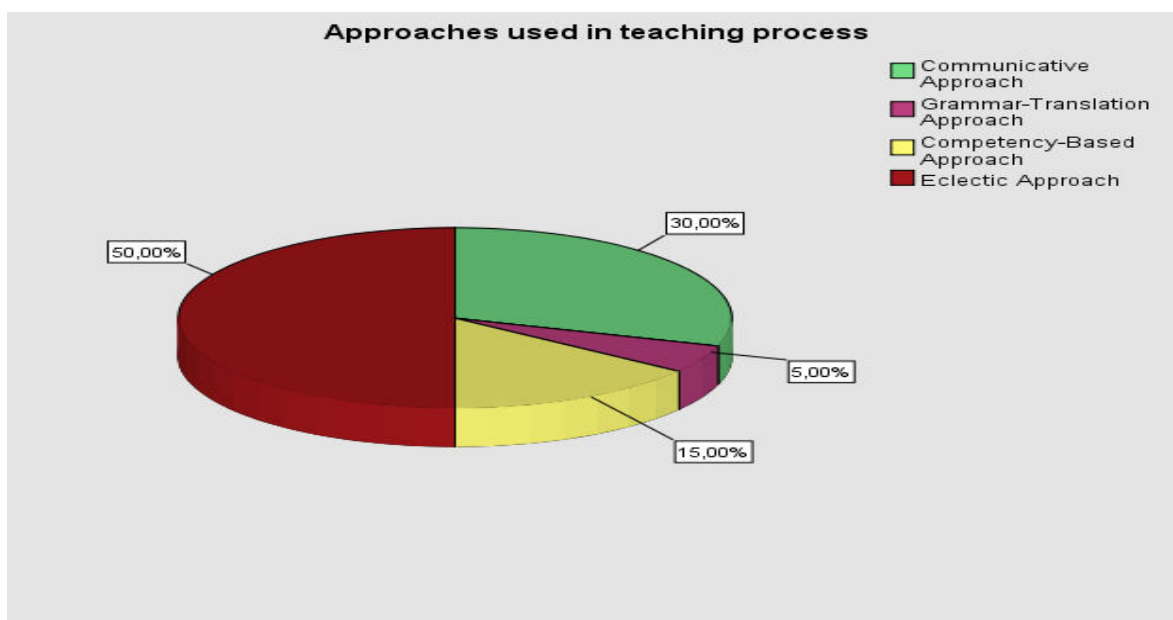
This pie-chart indicates the different approaches used by the teachers of the Listening/ Speaking subject, where 50% of them use the eclectic approach, which combines the three approaches: communicative, grammar-translation, and competency-based.

Q 4: Is the assessment of your student's speaking performance?

☐ Daily   ☐ Weekly   ☐ Monthly   ☐ Periodical
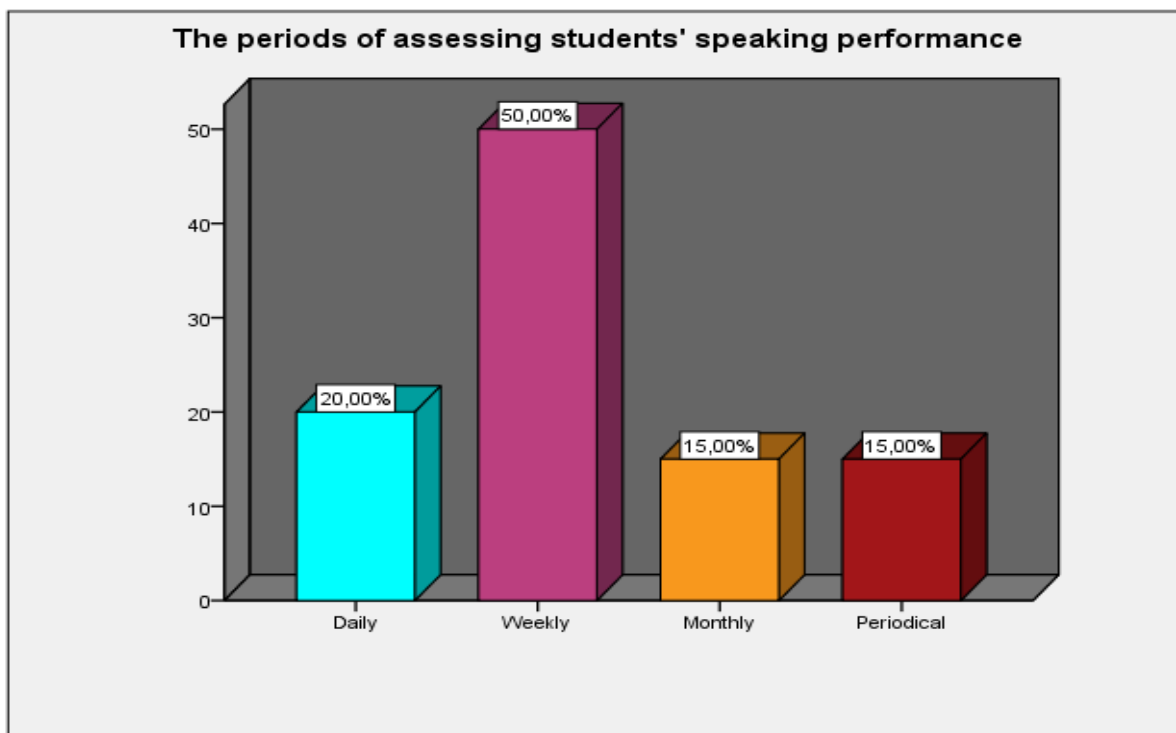
Diagram 4: The periods of assessing students' speaking performance

From the diagram above, it is clear that half of the participants assess their students weekly; and 20% of them assess learners during the lesson, while the two teachers (15%) evaluate their students' speaking performance monthly or periodical.

Q 5: What do you think is the most reliable way to assess someone's English speaking skill?

Based on the gathered data, the majority of the participants think that the most reliable way to assess someone's English speaking skills is based on their communicative performance, to get them involved in real-life situations. Whereas, some argue that direct question-answer and debate over prepared topics are the most reliable way.

Q 6: How do you select the topics to use in speaking tests/exams?

☐ I choose them myself.

☐ I ask the students what they like to talk about.

☐ I follow the course syllabus.

☐ I have another way. What is it please?

Diagram 5: Teachers perception about selecting the topics in speaking tests/exams

As underscored in this pie-chart, the majority (60%) of teachers selects the topics of speaking tests/exams following the course syllabus; and 25% of them ask their students what they like to talk about. Yet, the minority (15%) chooses the tasks themselves.

Q 7: You build speaking tests/exams…

☐ Alone   ☐ In collaboration with your colleagues   ☐ Relying on previous models
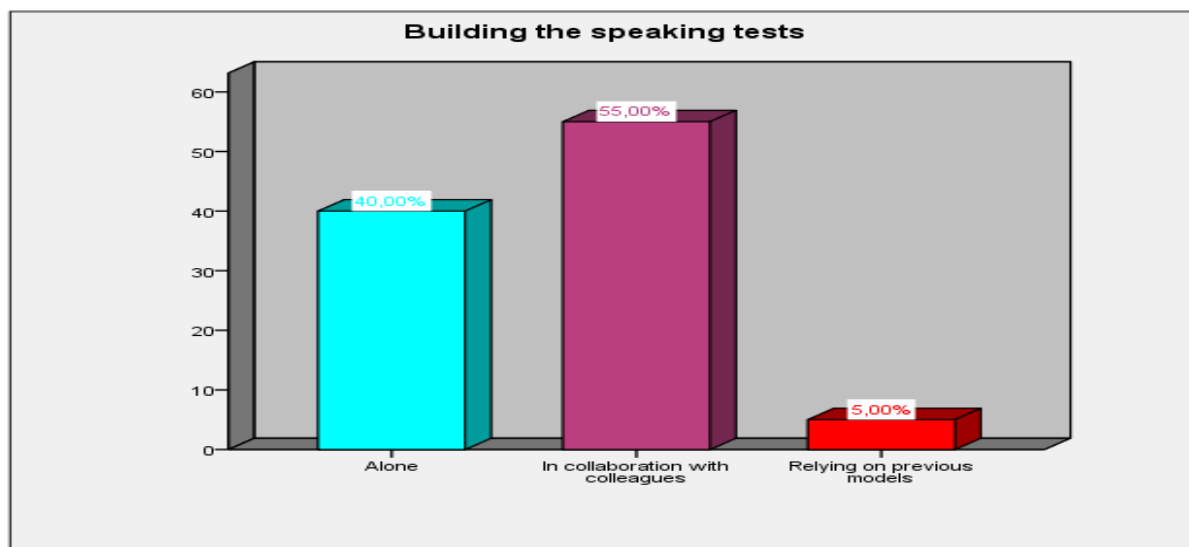
Diagram 6: building the speaking tests

The present graph indicates the way teachers build speaking tests/exams. As it is shown above 55% of teachers work in collaboration in order to design speaking tasks and 40% build them alone; yet, one of the teachers asserts that he relies on previous models.

Q 8: During the speaking tests/exams, you are…

☐ The 'Examiner'   ☐ the co-communicator   ☐ Facilitator
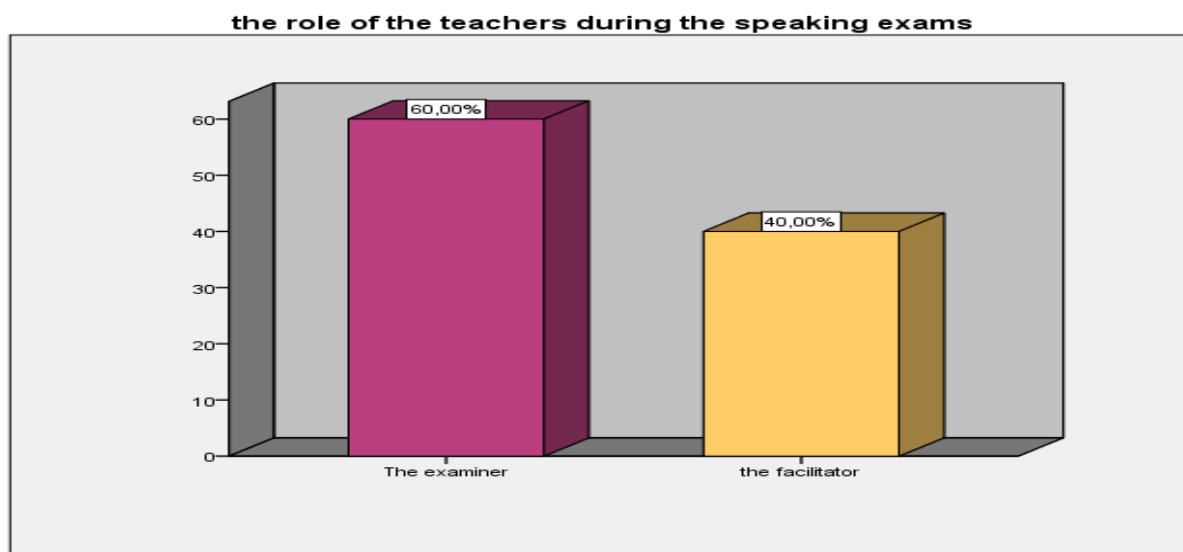
Diagram 7: The role of the teachers during the speaking exams

As shown in diagram (8), 60% of the teachers play the role of examiner during the speaking test/exam. In addition, 40% of teachers prefer to act as facilitators when they assess their students.

Q 9: What are the criteria you consider in assessing students' speaking skill?

☐ a) Pronunciation. ☐ b) Vocabulary.   ☐ c) Fluency

☐ d) Grammar.   ☐ e) All of them.

The criteria considered in assessing students'speaking skills
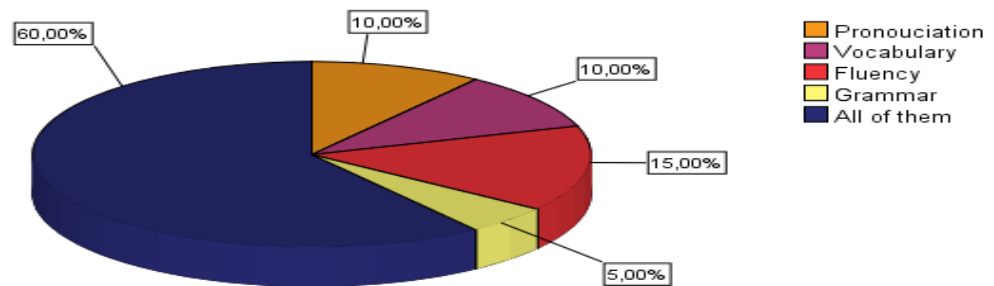
Diagram 8: Criteria considered in assessing students' speaking skill

This pie-chart denotes that the majority of teachers (60%) take into consideration all the criteria when they assess students' speaking skill. While 15% of them consider fluency as the most important criteria, since it represents students' speaking performance.

Section three: Factors affecting student's speaking performance and speaking problems

Q 10: You assess your students...

☐ Individually   ☐ In pairs   ☐ In groups

Why? Which type gives more reliable results?
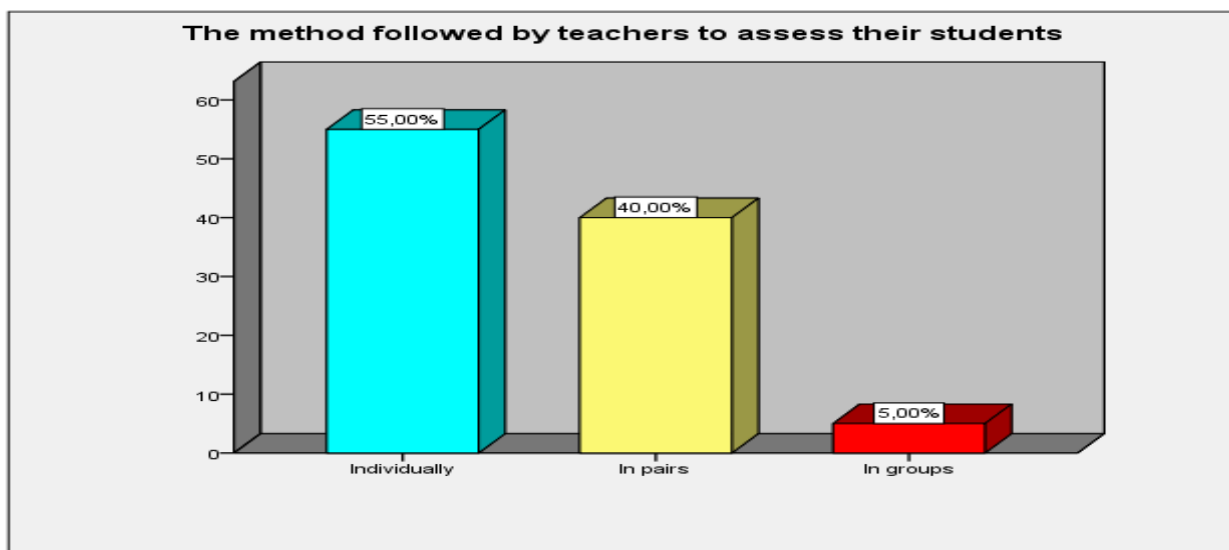
Diagram 9: The method followed by teachers to assess their students

Note: one (1) teacher did not answer the open-ended question.

The results of this question indicate that 55% of the teachers prefer to assess their students individually , because they think that the students' mastery of language differs, so assessing them individually is an appropriate way to know their personal capacities and skills which gives, in turn, more reliable results. Whereas, 40% of the teachers evaluate students' speaking performance in pairs, since it helps them to reduce their level of anxiety; and in communicative situations it is the best method to assess their interaction and communicative abilities, while, one (1) teacher chooses assessment in group to evaluate students' competence.

Q 11: During the test/exam, most of your students are…

☐ Anxious   ☐ Relaxed and self- confident

Diagram 10: students' attitudes towards speaking test/exams

As presented in this pie-chart, the majority (80%) of the teachers argue that students are anxious during the speaking test/exam; while, 20% of them state that the students are relaxed and self-confident.

Q 12: When you assess your students' speaking skill, which competence do you target the most?
☐ Linguistic   ☐ Communicative   ☐ equally, both of them
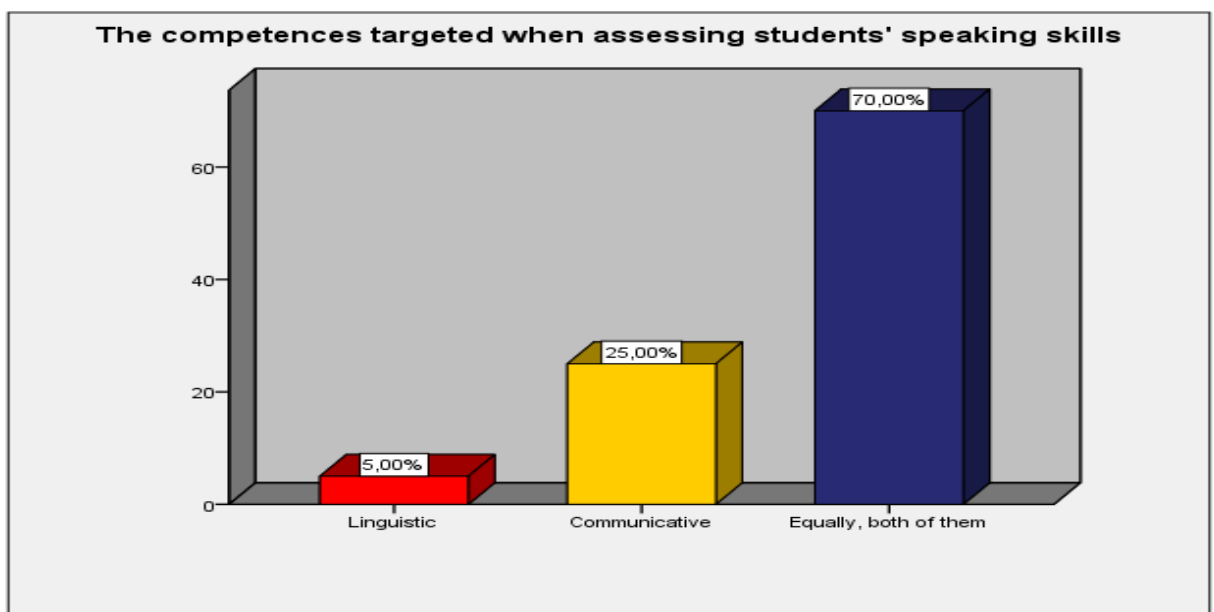
Diagram 11: The competences targeted when assessing students' speaking skills

As underscored in the diagram above, the majority (70%) of respondents target equally both linguistic and communicative competence when assessing students' speaking skill.

Q 13: Do your students prepare for the speaking task before it is performed?

☐ Yes   ☐ No
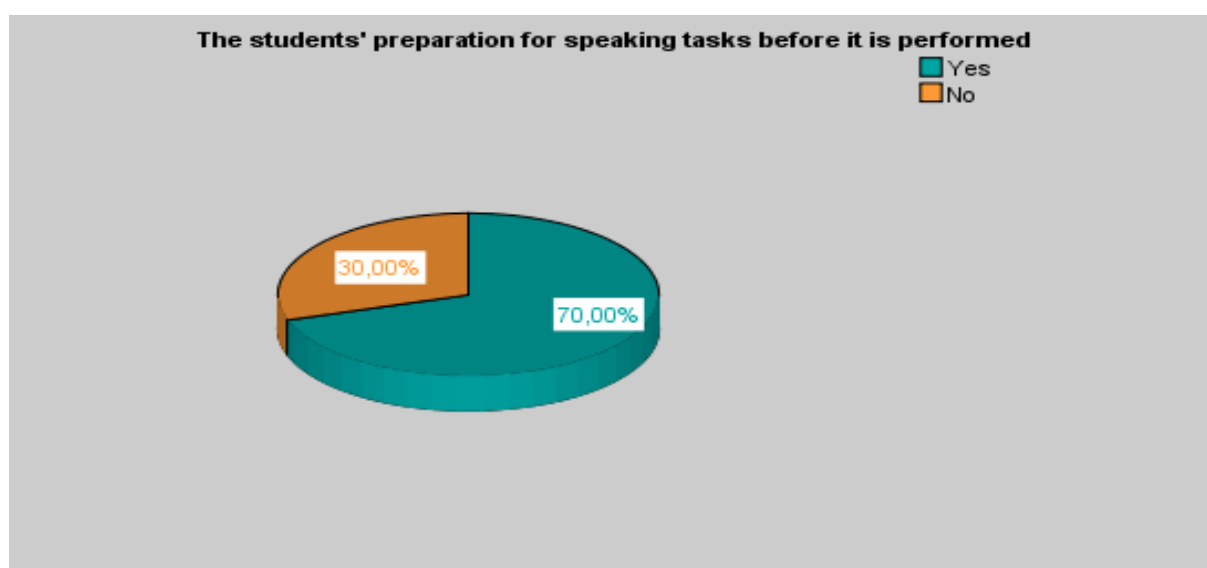
Diagram 12: The students' perception for speaking tasks before it is performed

As shown in this diagram, 70% of the teachers say that the students prepare for the speaking tasks before it is performed.

Q 14: Are your students given enough time to perform a speaking task?

☐ Yes   ☐ No
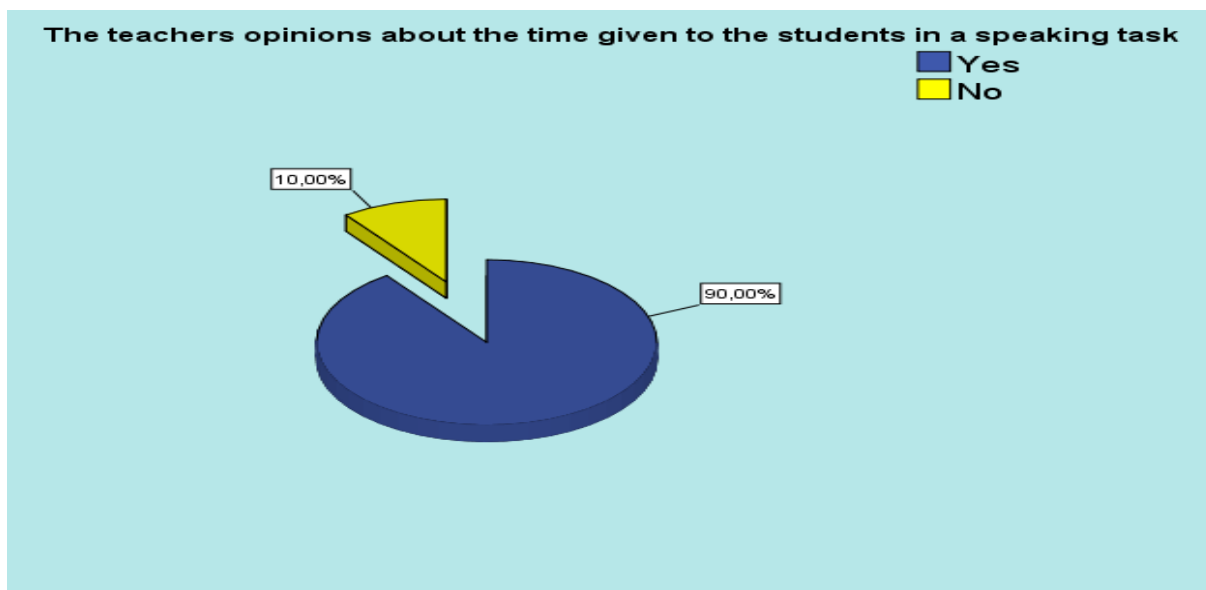
Diagram 13: The teachers' opinions about the time given to the students in speaking tasks

This pie-chart demonstrates that 90% of the teachers think that the students are given enough time to perform the speaking tasks; whereas 10% say no.

Q 15: how do you score them?

Remark: six (6) teachers did not answer this question

The data gathered from this open-ended question reveal that the majority of the teachers score their students relying on the previous criteria (fluency, pronunciation, vocabulary, and grammar). In addition, one (1) teacher adds cultural literacy (the use of idioms and cultural signs) as another criterion.

Q 16: In your opinion, which of the following factors may have a negative impact on your students' speaking performance during a speaking test/exam and scoring?

☐ Lack of time for preparation   ☐ Lack of self-confidence

☐ Poor background knowledge   ☐ Foreign-language anxiety

☐ Examination anxiety   ☐ Misunderstanding of test instructions
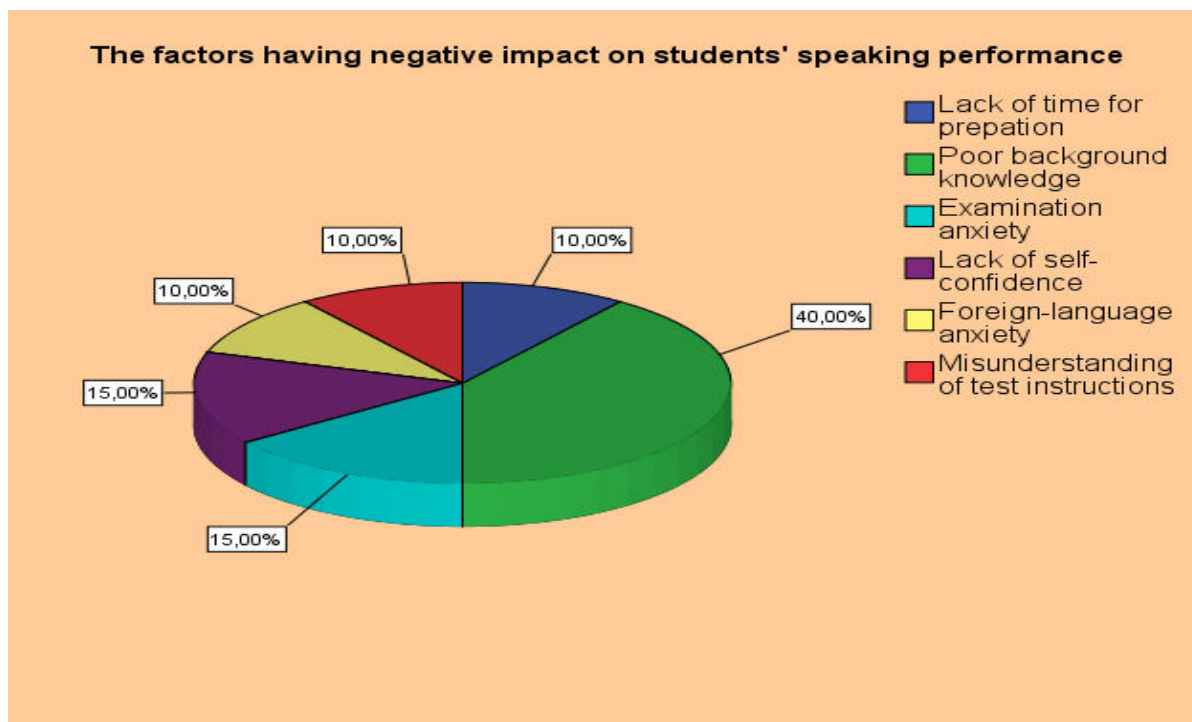
Diagram 14: The factors having negative impact on students' speaking performance

The results clearly demonstrate that poor background knowledge is the main negative factor that has an impact on students speaking performance with the percentage of 40% representing seven (7) teachers. In addition, examination anxiety and lack of self-confidence are equally selected by 30% of teachers as negative factors; moreover, the percentage of 10% to each of: lack of time for preparation, foreign language anxiety, and misunderstanding of test instructions.

Q 17: Do you think that the students' scores reflect their communicative competence? Why?

☐ Yes   ☐ Somehow   ☐ No
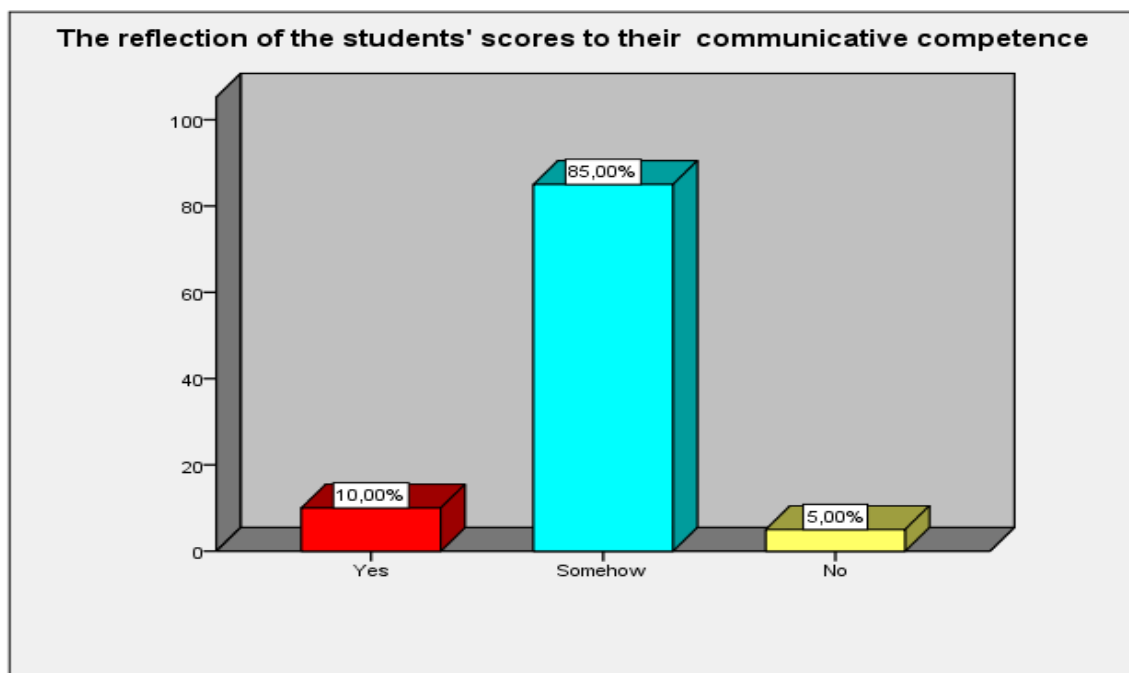
Diagram 15: The reflection of the students' scores to their communicative competence

As shown in the diagram (16), the majority of the teachers (85%) assert that students' scores do not really reflect their communicative competence because of shyness and anxiety, and 10% of the respondents see that students' scores reflect their capacities during the exam, since they master basics of language so they can express and communicate with people fluently, while only one (1) teacher has said NO.

Q 18: Has it ever happened to you to set the same test to the same students after a relatively short period of time?

☐ Yes   ☐ No

Teachers perception about assessing the same test to the same students

15,00%

85,00%

Yes
No
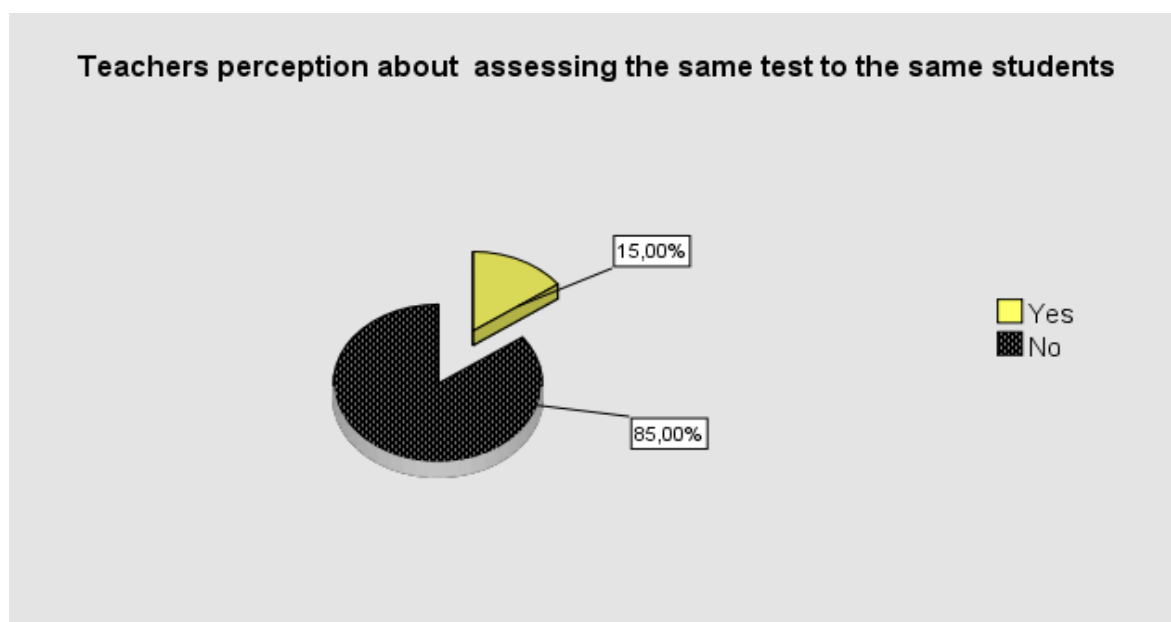
Diagram 16: teachers' perception about assessing the same test to the same students

The results of this diagram show that the majority of the teachers (85%) have not set the same test to the same students after a relatively short period of time, and only three (3) teachers have said YES.

If your answer is YES, have your students got the same scores?

☐ Yes   ☐ No

The reliability of the students' scores

10,00%   5,00%

Yes
No
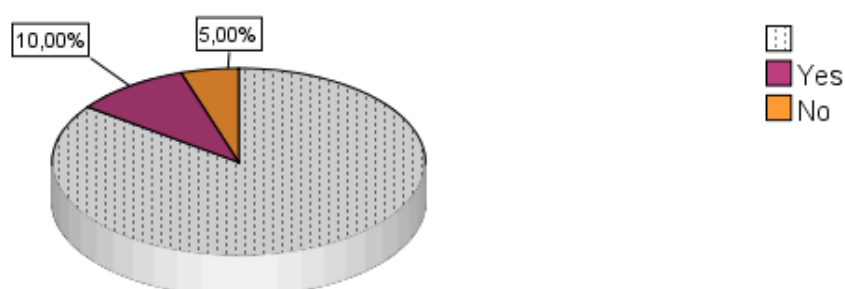
Diagram 17: The reliability of the students' scores

As illustrated in this diagram, two (2) teachers claim that the students have got the same scores after setting the same test; that is, one (1) teacher has said that he did not get reliable results.

English Speaking Tests Analysis

Through analyzing the structure of the examination subject, we find that there is a pattern of different models, which comprises free topics, IELTS models (the International English Language Testing System), or asking questions related to the syllabus. During the speaking exam, the examiner manages the test and evaluates the candidates, either individually, in pairs or in groups.

Generally, the test is divided into two parts; where the candidates draw a topic and get 3 minutes to prepare and brush up on their knowledge, and 2 minutes to perform the task. The candidate is asked to speak about a personal experience such as a memorable day, significant person, or everyday familiar topics (free topics). Furthermore, the students and the examiner will have a discussion relaying to the subject area, where the examiner asks the candidates to do more complicated tasks, such as evaluate, justify, give opinions, make predications and express preferences (IELTS models); whereas, some examiners prefer to relate the topics into the syllabus such as explaining the meaning of idioms or making conversations including modalities.

Conclusion

As a conclusion, one may say that this chapter has presented the results obtained from the questionnaire informed by teachers of the Listening/Speaking subject in the Department of English at MMUTO, and the analysis of a number of speaking tests/exams. In sum, one may assume that several extrinsic factors constitute an obstacle to the realization of an ideal

reliability of speaking exams. Together with the above results, these factors are to be discussed in detail in the next chapter.

DISCUSSION OF THE FINDINGS

Discussion of the Findings

Introduction

This chapter displays of the discussion of the results presented in the previous chapter, in an attempt to account for the research questions and hypotheses that we set earlier. The findings are discussed relying on analytic categories borrowed from Bachman's theory put forward in his *The Fundamental Consideration of Language Testing* (1990). Our discussion starts from the assumption that, in language testing, claiming that a test is valid means being able to establish a reasonable link between a student's performance and their actual language ability, or else their ability to show the same performance in real communication situations. It will revolve around the results we have obtained from our investigation of reliability in the 'speaking' tests administered to 3rd-year students at the Department of English of MMUTO.

1. Teachers' Questionnaire

1.1. Informants' Profile

The results of the first section of the questionnaire reveal that the majority of teachers are classified in the first category: Assistant Professors (MAB) or young part time teachers (Vacataires). Their professional experiences differ from less to more experienced teachers. Yet, the majority of them are less experienced. In fact, *"experience is not significantly related to achievement following the initial years in the profession"* (Rivkin *et al*, 2005) since experience does not guarantee excellence. Accordingly, individual teachers tend to improve with experience, but *"not all teachers begin their careers with the same skills or rise to the same level"* (Xu et al, 2005). In the same vein, Sass *et al* (2010) point out the fact that *"some less-*

experienced teachers are more effective than teachers with more experience", meaning that students are less inhibited and more likely to participate in oral classes when the instructor's age and standard of English are closer to theirs.

1.2. Foreign language assessment and speaking test design

After analysing the answers of the teachers of the Listening/Speaking subject, we have noted that half of the informants use the eclectic approach - the combination of different techniques - in their teaching process. Indeed, the use of an eclectic approach may be considered as a judicious choice, since a single method has a narrow theoretical basis and a limited set of activities. In addition, eclecticism brings diversity into the classroom, as it provides students with different ways of learning. Kumar (2013:2) states that *"the purpose of advocating eclectic methods is to connect life experiences to the ideas presented in the learning of the language. The types of learning activities teachers select are often directly related to their experiences in the real world"*, and this helps learners to look at learning as an extension of social, academic and professional environments

Some of our informants use the communicative approach to teach speaking skills taking it for granted that it helps students to develop their communicative ability in the learning process and real-life situations. They imply that communicative competence refers to grammatical and social knowledge about using utterances appropriately in a specific context. This goes in accordance with Bachman's framework CLA (1990), which specifies a set of essential components to create an interactive system; including language and strategic competences as well as psycho-physiological mechanisms. Accordingly, for consistency, any assessment tool should derive from this approach and take into consideration all the elements that constitute a communicative act.

It is also worth mentioning that half of the informants assess their students' speaking skills each week. They intend to improve their students' performance by giving oral or gestural feedback, which provides, in turn, information to both learners and teachers about the students' current level of understanding, performance and desired goal. In this respect, Anderson (2015) claims that *"good teachers plan specific objectives for each lesson and for each week,"* adding that *"exam preparation should not take the place of learning"*. He suggests that the two should be combined in a balanced curriculum with new vocabulary and grammar each week as well as each assessment operation. This would be quite acceptable if only speaking teachers, as test designers, accounted for communicative competence and considered that language is just a means, not an end in itself.

Concerning daily assessment, 20% of our informants think that this kind of assessment gives students more opportunity to develop their English speaking skills. This assumption fits with what Katheryn and Geralyn say *"many types of discussion take place in a classroom on a daily basis, and teachers might evaluate these to determine if students are provided ample opportunities to refine their oral language skills"* (2007: 199). It is also argued that to know the real performance of the students in a foreign language, the teacher should accomplish daily tasks to observe learners' achievement (Bachman, 1990). Yet, the students' ability to cope with real-life communication situations cannot be assessed through mere simulations initiated by the teacher aiming to check whether a given language point is part of students' linguistic competence, forgetting all about the appropriateness of those language forms in communicative contexts.

As for the teachers' viewpoint on the most reliable way to assess someone's English speaking skill, it may be summarized as follows: *"the most consistent way is*

42

to test the ability of the students focusing on their speaking performance". One of our informants claims that: *"It is to give learners the opportunity to talk and express their ideas in their interaction with one another in different situations"*. Bachman (1990: 61) states that *"in order to make a decision about individuals or inferences about their language ability, the tester has to be able to show how test performance corresponds to language use in other situations"*. In the same vein, Weir (1990:12) points out that *"in assessing the ability really we should try to reflect the interactive natures of normal spoken discourse and attempt to ensure that reciprocity is allowed for in the test tasks included"*. This includes the nature of language and the ability to use it to negotiate meaning and cope with other elements that make up strategic competence in the Bachman's model of communication, where he simply asserts that *"strategic competence performs assessment, planning, and execution functions in determining the most effective means of achieving a communicative goal"* (1990: 108). Therefore, these results confirm the first and the second hypotheses, since the speaking tests under scrutiny are reliable, and students who pass these tests can cope with similar communication situations in real life.

As regards the issue of the selection of the topics to devise in speaking tests/exams, the survey reveals that the majority (60%) of teachers follow the course syllabus, since they build speaking tests in relation to the program, with some coordination with colleagues. This way of doing is well in accordance with Bachman's approach (1990: 06), assuming that *"such tests are equally important for use in making decisions about language competency [whether] in the context of evaluating learners achievement in language programs"*. He adds that *"performance in oral interview tests might be sufficient information to determine whether the program has been successful in term of students' mastery of the program's*

objectives" (ibid). However, the question that arises here is whether the objectives of the course programme are formulated in a way to meet students' needs in terms of communication in real, academic or professional life. Another question is whether each lesson objective is specific, attainable and measurable; otherwise when teachers come to assess their students' speaking competence, there may be high risks of unreliability.

Some of our informants have a tendency to build the speaking tests alone, and most of them give their students the opportunity to choose a topic to speak about. Bachman states that *"In some situations, where incoming students may vary widely in their backgrounds, it might be more appropriate to base the placement tests on a general theory of language proficiency, while still basing the progress tests on the objectives of the course"* (ibid). It goes without saying that it is useful to let students speak about topics of their choice when they take a test, especially for motivation reasons, but the assessment of speaking skills should be considered as a first step towards learners' communicative competence. Bachman advises teachers to test students on their general language proficiency when education institutions, especially private language schools, come to place new students in different classes with different language syllabuses to follow. Still here, we are concerned with students' speaking ability as a pre-requisite to follow a given syllabus designed to develop communicative competence.

Another point worth mentioning is that 60% of the teachers who kindly accepted to inform our questionnaire claiming that they play the role of examiner during the speaking tests/exams. The examiner usually takes notes to remember the learners' mistakes. However, most often the assessor considers not only the weaknesses of the test takers, but they also note down the strengths of the learners'

speech to make considerable decisions with reliable judgment. In this sense, Bachman says that language teachers regularly use tests to *"diagnose student strengths and weaknesses, to assess students process, and to assist in evaluating student achievement"* (1990: 03). So the assessor listens to the learners without interviewing them, and by doing so, they can fully concentrate on the performance of the learners. Underhill, on his part, defines the assessor as *"a person who listens to a learner in an oral test, and makes an evaluative judgment on what he/she hears"* (1987:07). Yet, their judgment would be unreliable if, at least, they do not jot down notes on an evaluation grid designed for this purpose. The evaluation grid should obviously be criterion-based, with communicative indicators.

On the other hand, 40% of our informants prefer to act as facilitators when assessing their students' speaking skills, their aim being to create an atmosphere of relaxation to enhance an exchange of ideas and thus encourage students to explore, identify and practice specific speaking skills. They also give them more responsibility and freedom during the task performance to display their speaking skills. According to Billings and Terry (2013: 80), *"the facilitator gives time and space for students to consider their own communication skills, both individually and collectively. The facilitator helps the assessment process by keeping the data organised so that it can be retrieved for future use"*. This assertion is quite acceptable, but even with facilitation; the test might still remain unreliable if the examiner is not an inspirer.

As stated in the responses to the questionnaire, there are several aspects of speaking that influence the way teachers evaluate their students' speaking proficiency. The aspects include: pronunciation, comprehension, vocabulary, fluency, grammar and interaction. Thus, 60% of our informants do take into

consideration all these criteria which, in turn, form a language as a whole. According to Bachman (1990), in language use, components are not seen as separate parts, but they are all related to one another in real-life situations. Hence, communicative language use is an interlocutor between all the competences and the actual context under two main competences: organizational and pragmatic. However, some elements are typically considered more important by teachers: fluency, pronunciation, vocabulary, and grammar. This may inevitably lead to high risks of unreliability, as one examiner, for instance, may focus on pronunciation accuracy while another may find vocabulary the most salient feature.

Nicolosi *et al* define fluency as *"smoothness with which sounds, syllables, words and phrases are joined together during oral language; lack of hesitation or repetition in speaking"*. In this regard, 15% of our informants have a tendency to adopt only fluency as the basic criterion in speaking assessment, in which they evaluate the candidates' ability in linking utterances to form coherent speech without undue hesitation. The utterances should be relevant to the task and should be arranged logically to develop the themes or arguments required by the task. This goes hand in hand with what Bachman asserts: *"In testing, testers should be interested in quantifying mental attributes and abilities such as motivation, fluency in speaking and intelligence"* (1990: 17).

The same number of participants (15%) prefers pronunciation and vocabulary as the basic criteria during the speaking tests/exams. 10% of them think that vocabulary is an essential building block of language, and so it helps to measure learners' oral standard of the English language. They assume that assessing vocabulary should be given prominence in any attempt to set criteria for assessment. In this respect, Semensen (1998) asserts that the lack of sufficient vocabulary is an obstacle in using language. Indeed, during the test teachers tend to emphasise the use of the lexis that is related to

the topic being discussed. It is true to say that vocabulary poses a real problem to foreign language learners when they come to communicate, and the solution resides in training them to develop some communicative strategies, such as paraphrase, transfer, and avoidance. Since teaching and evaluation are closely connected, it would advisable to design tests that are likely to assess students' ability to use communicative strategies in their struggle to convey meaning. In tests based on student conversation, for instance, the examiner should focus on the communicative strategies adopted by the interlocutors to overcome the problem of poor vocabulary mastery.

On the other hand, 10% of the participants tend to assess their students focusing more on the overall impact of pronunciation, which refers to the candidates' ability to produce intelligible utterances to fulfil the task requirements. This includes stress, intonation as well as individual sounds. Furthermore, teachers consider that for students to speak English accurately, they should master its phonological rules and be aware about the different sounds and their articulations. Ultimately, 5% of teachers consider grammar as the most important element to evaluate student's speaking performance which refers to the accurate and appropriate use of a range of grammatical forms. Luoma (1993) argues that grammar is an element that influences the way we evaluate spoken production; it is an efficient way to measure student's mastery of language. Indeed, it constitutes a reliable element in testing learners' linguistic competence, but it should not be the main focus in speaking assessment. A communicative strategy, called cooperative strategy, can be used by one of the interlocutors. If they cannot find the right grammatical structure to use, for instance, they signal the problems to the interlocutor and attempt to get the problem solved on a cooperative basis.

1.3. Factors Affecting Student's Speaking Performance

This section is concerned with the factors which affect students speaking performance and speaking problems. Thus, the results displayed previously concerning the methods followed by the teachers to assess their students, show that half of the respondents (55%) prefer to assess their students individually since their mastery of language differs, so assessing each candidate alone reflects their individual speaking skills, as one of the teachers claims: *"I prefer to assess my students individually because they are given more opportunity to be fluent and discuss the topic at ease. So, I ask them questions about a given topic which they should answer at length to demonstrate their abilities at speaking English"*. Bachman rightly sustains that *"our primary concern is whether an individual's test performance can be interpreted as an indication of his competence, or ability to use language appropriately and effectively in non-test contexts"* (1990:11). In other words, if the student cannot adopt the same communicative behaviour in a "non-test context", the test is doomed to be unreliable.

In the same vein, 40% of the participants use pair assessment to test their students speaking performance, and teachers affirm that the motivation for paired testing is based on the interaction between candidates when speaking to each other in pair-test tasks, where students feel more relax and self-confident during the speaking exam. One of the respondents claims:

> *"Most often, I assess my students in pairs to get them involved in a specific communicative situation, so, in my opinion, pair assessment is the best method to assess my students interaction and communicative abilities to prepare them interacting in real-life situation outside the academic scene".*

In regard to what has been mentioned above, Finch *et al* argue that pair work in speaking tasks increases students' motivation and lowers the inhibitions of learners who are unwilling to talk in front of the full class. Besides, only one teacher

evaluates his students in group, which is, for him, an effective way to show their abilities when communicating with others by adding external factors of learning such as gestures and facial expressions. Richard *et al* have point at the efficiency of this testing technique, claiming that *"in an attempt to make assessments as realistic as possible, teachers often have students work in a group"* (1998: 223). It is indeed advisable to use it, as a group conversation constitutes one of the best opportunities for the teacher to get an exhaustive evaluation of their students' communicative competence; all features of communication are provided in group discussions, be they social, sociolinguistic, psychological, etc.

Throughout the answers obtained from the question concerning the competence which is targeted more when assessing students' speaking skills, the majority of teachers (70%) pretend to focus more on both linguistic and communicative competencies, which are considered as part of an overall language proficiency. Lin claims that *"we should study the knowledge that people have when they communicate, just like linguistic competence tells you whether a sentence is grammatical or not, communicative competence tells you whether an utterance is appropriate or not within a situation"* (2004: 02). In other terms, both competencies are complementary, since one cannot communicate without using their grammatical knowledge, this goes hand in hand with what Faerch *et al* pretend: *"it is impossible to conceive of a person being communicatively competent without being linguistically competent"* (1984: 168). Yet, the same question arises: Do teachers design tests in a way to assess their students' communicative competence, or else do they grant linguistic competence when they come to set their test criteria.

In fact, Bachman argues that language knowledge is the combination of both: linguistic competence and communicative competence. First, OC (organizational

competence) refers to the mastery of language, as Hedge rightly puts it: *"linguistic competence involves knowledge of spelling, pronunciation, vocabulary, word formation, grammatical structure, sentence structure and linguistic semantics"* (2000: 47). For instance, the candidate should be able to produce utterances that are grammatically correct. Second, PC (pragmatic competence) deals with the use of linguistic competence in a particular context to create a communicative system; that is to say, PC includes not only the mastery of language structure but also illocutionary and sociolinguistic competence. Indeed, students should use appropriate language to achieve specific communicative goals. Furthermore, using cultural references and language functions is an effective way to perform tasks during the speaking exam. All in all, for a test to be as much reliable as possible, it should also as deep as students' sensitivity to register, language variants, idiomatic expressions, figures of speech, etc.

Bachman (1990) argues that when measuring and making decisions about individuals or inferences about their language ability, the tester should relate the test performance to language use in other situations. By "other situations", we understand outside-classroom contexts. On the contrary, the results of the present study show that teachers score their students focusing on fluency, vocabulary, pronunciation, grammar and interaction. The following respondent's testing method is quite illustrative of their speaking assessment approach: *"the student's performance is scored out of 20 points: 8 for language, 8 for fluency (the ability to interact and communicate) and 4 for cultural literacy (the use of idioms, cultural signs)"*.

Accordingly, Bachman argues that *"there are many factors other than the ability being measured that can affect performance on tests and that constitute sources of measurement error"* (1990: 24). The outcome from the questionnaire,

dealing with the factors that may have a negative impact on students' speaking performance during an oral test/exam, reveals very positive results concerning student preparation for speaking tasks; thus, the vast majority of the teachers report that students prepare for the task before it is performed. In the same vein, almost all the participants declare that students are given enough time to perform. However, a task done without planning and lack of time for preparation is more likely to lead students to anxiety and lack of self-confidence through choosing relatively undemanding language and inappropriate utterances. For that, Bachman has highlighted the importance of time for preparing and performing the task *"the amount of time allocated for the test or its parts is likely to affect test performance. In some tests, the time limit is such that not all test takers can manage to answer all the items or parts of the test"* (1990: 122).

The misunderstanding of test instructions is another factor mentioned by teachers, considering it as an obstacle during the speaking exam; hence most of the time some slow students do not know what exactly is expected of them to do to demonstrate their abilities in speaking. As for other language skills, test instructions play a crucial role in test takers' performance, since their performance depends, to a great extent, on how well they understand the conditions under which the test is taken. It is a factor which refers to characteristics of test tasks such as timing and test rubric. Besides, *"unclear or inaccurate instructions and inadequate time allocation act as a source of test anxiety, and hence, influences on test performance"* (Madsen, 1982). Most of the time, these factors are not taken into account when one comes to score students' performance.

The teachers of the Speaking/Listening subject view that anxiety prevents students from successfully learning a foreign language, while it affects negatively on

their performance. In this regard, Abu-Rabia (2004) states that *"oral communication skills are more likely to be affected by language anxiety where the learner is afraid of spontaneous communication in the FL"*. Additionally, Anxiety is a psychological aspect which affects on the mental capacity of the candidates, relatively it is argued that *"one element of general anxiety composed of cognitive intentional process that interferes with competent performance in academic or assessment situations"* (Spielberger & Vagg, 1995). To the more particular aspect in speaking a foreign language, anxiety seems to occur, which prevents candidates from producing accurate and appropriate utterances. According to Spolsky *"the learner's meta-cognitive awareness of his/her inability to speak fluently and understand the full message given in the FL may lead him/her to despair, hesitation, rejection of any communication activity"* (1989:83). In this regard, Na (2007) states that FL anxiety makes students become discouraged, and they lose confidence in their potential to learn the foreign language. In other words, for psychological considerations, an unreliable test can give rise to serious problems in terms of motivation, and this is unlikely to happen when a learner is tested on their writing skills, for example.

Contrary to being anxious because of a tense atmosphere, being relaxed and self-confident helps students to speak fluently and accurately when performing the speaking task. This idea goes in accordance with what Hanton *et al* say *"high self-confidence increase the intensity of thoughts and feelings that you can control... if you're confident you stay in control of your thoughts..."* (2003: 481); put otherwise, feeling relaxed during speaking assessment enriches and promotes student's English production whereas a low level of self-confidence, forestalls the achievement of oral performance, as foreign language learners cannot perform a task or express themselves fluently and accurately

without some degree of self-confidence (Brown, 1994). In fact, these factors are relatively derived from the test-taker characteristics.

Accordingly, lack of time for, misunderstanding of instructions, anxiety and poor background knowledge are factors that may affect student's performance and test scores. This fits with Bachman's framework:

> *"In any language testing situation, as with any non-test situation in which language use is involved, the performance of an individual will be affected by a large number of factors, such as the testing context, the type of tests, tasks required, and the time of day, as well as her mental alertness at the time of the test, and her cognitive and personality characteristics"*(ibid: 30-31)

The findings of this research show that almost all the teachers suppose that the students' scores do not really reflect their communicative competence. Besides, some teachers claim that, when students are anxious during the period of examinations, they less-perform the tasks. Hence, anxiety is an affective factor that may impact on students cognitive and linguistic capabilities by causing grammatical errors and tongue-tied in articulating and producing speech sounds. In other words, even if the students are competent, such factors (cognitive, linguistic and affective factors) might affect their scores, which in turn reflect on their communicative competence. In addition, because students are evaluated once in a semester, it is not obvious to determine whether they are competent and skilled in speaking. This point fits with what Bachman points out: *"individuals' performance may be affected by differences in testing conditions, fatigue and anxiety they may, thus, obtain scores that are inconsistent from one occasion to the next"* (1990: 24).

1. Analyzing Speaking Questions

In the English department of MMUTO, the teachers of the Listening/Speaking subject design different tasks to assess the performance of their students and measure

their oral communicative competence. Some tasks are directly related to the course syllabus and its objectives that are far from being specific, realistic and measurable. In a word, they do not respond to Bachman's suggestion: *"all language tests must be based on a clear definition of language abilities whether this derives from a language teaching syllabus [...] and must utilize some procedures for eliciting language performance "* (1990: 09);

The major problem in speaking assessment is that, though the majority of the topics selected for discussion are taken from students'personal experiences and real-life situations, still most of them may be perceived by students to be artificial and chosen for testing purposes. When the test item is unrealistic - suffering from the lack of information gap - students quickly get uninterested, and this impacts negatively on their performance even though they try to do their best in order to get a good mark. Regarding this very point, Weir indicates that:

> *"[...] in testing communicative language ability we are evaluating samples of performance, in certain specific contexts of use, created under particular test constraints for what they can tell us about a candidate's communicative capacity or language ability"* (1990: 07).

The major handicap does not reside in the selected topic only, but it has to do with the "particular test constraints"mentioned above. Students may be underscored for their performance in a test, and yet they could do better if they performed in a real-life communication situation that has nothing to do with the academic context.

The aim of speaking tests derived from social tasks is to see how well candidates would cope in a spontaneous real-life situation. For Bachman, *"there must be a relationship between the language used on tests and that used in real-life"* (1990: 356). As for Katheleen, he states that *"communicative language tests are intended to be a measure of how the testees are able to use language in a real-life situation"*

(1996). This statement invites teachers to try hard to design tests that allow them to predict their learners' ability to communicate in authentic contexts outside the educational scene, and infer the capability of the candidates to apply their skills and engage in effective communication, but in practice, it is not obvious that teachers comply with these suggestions, because whatever the quality of the test item, it cannot replace real context.

A quick study of the sample tests we have gathered at the Department of English has revealed that the majority of the teachers build test tasks that are likely to make students show their linguistic language abilities. Most of the topics are borrowed from IELTS (International English Language Testing System), such as describing a city, a movie, a museum or a historic place. According to one of the British council's Newsletters (2015), IELTS tests seem to be *"interactive and as close to a real-life situation as a test can get"*, but one may say that, when for example a student is set to describe a place, there is no place to the assessment of their communicative competence. On the other hand, topics like explaining English idioms and making conversations are more likely to give reliable results in terms of communication.

Starting from the viewpoint that the speaking tests' purpose is to identify learner's ability to use language authentically, we consider test-takers should be assessed on their aptitude to produce words appropriately in the right context and accurately with correct meaning. Accordingly, choosing the right topics in exams that give reliable information about students'communicative performance is significant, since oral tests are the most consistent way to measure and evaluate speaking proficiency with a close relationship between the selected tasks and real life in its widest sense. Hence, the findings reveal that the teachers are aware of the

importance of the reliability of speaking tests, even though most of them focus more on the students' academic achievements than on their ability to cope with communication situations in the future

Conclusion

In the light of what has been said so far, one may claim that the speaking tests set to 3rd-year students in the English Department of MMUTO are interpreted and discussed to draw conclusions about the research questions outlined previously. Consequentely, from analysing the results of the questionnaire and the corpus, the study concludes that students who pass speaking tests can cope with similar communication situations in real life, since the tests under scrutiny are reliable. In addition, many factors can affect students performance during the speaking exams which may influence, in turn, on their reliablity such as time limitation, anxiety and lack of language knowledge. Therfore, the major findings of the present research confirm the previous hypothesis about the reliability of the speaking tests adminstered in the English department at MMUTO.

GENERAL CONCLUSION

General Conclusion

The aim of the present research work has been to investigate the reliability of speaking tests devised for 3rd-Year students at the English Department of Mouloud Mammeri University at Tizi-Ouzou. It has sought to determine the extent to which students' performances are likely to be replicated in real life communication situations. Being based on Lyle Bachman's theory put forward in his *Fundamental Considerations in Language Testing* (1990), it has taken authentic communication as a parameter to evaluate the reliability of the tests under study. Accordingly, we have considered that, in order for a student to demonstrate communicative competence, their performance at the 'speaking' test should be measured on the basis of the major elements of the communicative approach as it is advocated by Bachman (1990), including linguistic competence, strategic competence, socio-cultural competence, actional competence, and discourse competence.

We have found it very useful to consider the aspects of reliability in the designing of speaking tests at the Department of English. Subsequently, our research work has made us look at testing, especially oral testing, in a more critical way, and be more aware about the need for building tests within a communicative framework to achieve acceptable reliability. It has also made us realise that there is a lot more to do to improve oral testing that we initially envisaged. We have realised that a speaking test cannot be objective if once created it is put in a 'test bank' and then taken out only at times of use.

Before conducting this research work, we often thought, particularly during the discussions held in ELT classes at the Department, that there is no need to be so thorough and precise concerning all aspects of the speaking test. Now, we have also have come to realise the tremendous difficulty our teachers used to face when they came

to build tests to assess our speaking abilities, trying hard to make them as reliable as possible. Accordingly, we feel that our experience with this study has equipped us with a better understanding of the complex dynamics of oral testing and will certainly be of good benefit for our future professional development.

As for the results we have reached, they reveal some weaknesses worth mentioning here. One of them has to do with time limitation, poor background knowledge, and psychological and physical conditions (fatigue, anxiety and lack of self-confidence), which are factors that may impact negatively on students' speaking performance and their scores, which in turn do not reflect their real language abilities. In addition, while assessing students through an interview, the role of the interviewer (scorer at the same time) remains dominant, and the student's role is restricted to responding only to the questions asked by the interviewer. Another weakness lies in the fact that the test is conducted in formal context only. In daily life, we have to speak mostly in informal context, but the speaking tests we have analyzed do not test the speaking skills in informal contexts. Moreover, in real life, ideas are not well formed in mind. They have to be generated immediately and quick responses are required, but in the tests object of our study, the candidate is given enough time to prepare their ideas.

Keeping in view the above discussion, and even though the results we have attained confirm, at some extent, Hypothesis 1 and Hypothesis 3 set in the 'General Introduction', the following three suggestions, among many, are presented for bringing improvements in speaking assessment and making it more reliable. First, 5 minutes of time is not enough to assess the oral ability of a non-native speaker. The candidate may want to expand the topic and ask supplementary questions, and they should be encouraged. It will not only be helpful to elicit more authentic communication, but it will also provide opportunity to the teacher to assess a candidate's questioning skill,

which is an important aspect of speaking skills. Second, a single task, i.e. an interview, is not sufficient to elicit the information expected. At least one more task like role play or picture description, etc. task should also be introduced. Thirdly, there should be more than one examiner. It will not only increase the reliability of assessment but will also reduce entire responsibility from a single scorer. It will also help to make the discussion more informal and reduce pressure on the candidate. There should also be some variation in the grading scale considering the social and educational background of the candidate as a non-native speaker of English.

Even though we have put ourselves in a position to give recommendations, we are still aware of limitations and shortcomings of our research. First of all, the reliability of the tests was not controlled by means of real life communication situations in an English speaking country. Second, the population of the experimental group is small; only twenty teachers might not represent the majority of the pedagogical staff of the Department. Third, since this study is not reinforced by an experimental operation, it is unavoidable that a certain degree of subjectivity can be found. Therefore, further researches can be done to investigate other aspects of reliability by using different data collection tools such as experimental methods, as well as it can pave the way to further development studies about speaking assessment.

BIBLIOGRAPHY

Bibliography

Abu-Rabia, S. (2004) *Teachers' role, learners' gender difference and FL anxiety among seventh-grade students studying English as a FL educational psychology,* vol. 24, no. 5, pp. 711-719. Available: https://doi.org/10.1080/0144341042000263006

Aggrawal, Y. R. (2008) *Statistics Education*, 2nd edition, Sterling: Delhi.

Alderson, J.C. and Hugh, A. (1981) *issues in language testing*, London: The british council, pp. 12.

Anderson, J. (2015) *Affordance, learning apportunities, and the lesson plan performance,* ELT Journal 6913, pp. 207-210.

Anivan, S. (1991) *current development in language testing*, Singapore: SEAMEO regional language centre, pp.92-102.

Bachman, L.F. *(1990) Fundamental Considerations in Language Testing,* Oxford: Oxford University Press.

Bachman, L. F. and Palmer, A. S. (1996) *Language Testing in Practice,* Oxford: Oxford University Press.

 Bachman, L. F. (2004) *Statistical Analyses for Language Assessment*, Cambridge: Cambridge University Press.

Banta, T. W. and Palomba, C.A. (1999) *Assessment Essentials: Planning, implementing and improving assessment in Higher education.* San Francisco: Jossey-Bass publishers.

Billings, L. and Terry, R. (2013) Teaching Critical Thinking: Using Seminars for 21 Century Literacy, pp. 80

Black, P. J. and William, D. (1998) *Inside the Black Box: Raising standard through classroom assessment,* London: King's College. London School of Education.

Boldt, R.F. (1992) *Reliability of the tests of spoken English revisited*, [online], available: http://www.ets.org/research/policy-research-reports/publication/report/1992/hxjk.

Brown, A. (2003) *Interviewer Variation and the Co-construction of Speaking Proficiency: Language Testing.*

Brown, G. and Yule, G. (1983) *Discourse Analysis*, Cambridge: Cambridge University Press.

Brown, H. D. (1994) *Teaching by principles: An interactive approach to language pedagogy,* Englewood cliffs, NJ: Prentice Hall.

Brown, S. and Smith, B. (1998) *Getting to Grips with Assessment*, Birmingham: Seda publications.

Butler, A. F., Stevens, R. (1997) Oral Language Assessment in the Classroom. Theory into Practice, Vol. 36, No. 4, pp. 214-219.

Cabrera, P. M. (2014) *Choosing a speaking test in English as a foreign language for the university entrance exam*, vol. 26, pp. 377-400, [Online], available:

http://dx.doi.org/10.5209/rev-dida.2014.v26.46849 [july 2014].

Canale, M. and Swain, M. (1980) *Theoretical Basis of the Communicative Approach to Second Language Teaching and Testing.*

Carmines, E. G. and Zeller, A. R. (1979) *Reliability and Validity Assessment*, London: sage publications Beverly Hills, pp. 37.

Davies, A. (1978) *Language testing*. Part II language traching and linguistic abstract, vol. 2, pp. 215-231.

Dörnyei, Z. (2008) *Research Method in Applied Linguistics: Quantitative, Qualitative and Mixed Methodologies*, United Kingdom: Oxford University Press.

Fearch, C., Haastrup, K., and Phillipson, R. (1984) *Learner language and learning*. Clevedon, Avon: Multilingual matters.

Fulcher, G. (2003) *Testing Second Language Speaking,* London: Pearson longman Education.

Fulcher, G. (2003) *Testing for Language Teacher, 2nd edition,* Cambridge: Cambridge University Press.

Fulcher, G. and Davidson, F. (2007) *Language Testing and Assessment,* USA: Routledge. pp. 42-104.

Fulcher, G. (2010) *Practical Language Testing,* Hodder education, UK: London.

Given, L. M. (2008) *Explanatory Research in: the sage encyclopedia of qualitative research methods*, [Online], Available: http://dx.doi.org/10.4135/9781412963909.n164 [2008].

Hadley, G. and Mort, J. (1999) *An investigation of inter rater reliability in oral testing,* [online], available: http://nuis.ac.jp/hadley/publication/interrater/reliability/htm [6 sep 1999].

Haertel, H.E (2013) *Reliability and validity of inferences about teachers based on students test scores*, Educational testing service, pp.8.

Hanton, S., O'bien. M., and Mellaliew, S. D. (2003) *Individual differences: Perceived control and competitive trait anxiety*, Journal of sport behaviours, vol. 26, pp. 481.

Hedge, T. (2000) *Teaching and learning in the language classroom*, Oxford: Oxford University Press, pp. 47.

Huxham, M., Campbell, F., and Westwood, J. (2012) *Oral Versus Written Assessments: A Test of Students Performance and Attitudes.*

Hyghes, A. (2002) *Testing for Language Teacher,* Cambridge: Cambridge University Press.

Hyghes, A. (2003) *Testing for Language Teacher, 2nd edition,* Cambridge: Cambridge University Press, pp. 09.

Hymes, D. (1971) *On communicative competence*, Philadelphia: University of Pennsylvania Press, pp. 16.

Hymes, D. (1972) *On Communicative Competence.* In J. B, pride and J. Holmes (eds.), *sociolinguistics,* Harmondsworth: Penguin Books.

Johnson, R. B., Onwuergbuzie, A. J. (2004) *Mixed Method Research: A research paradigm whose time has come*, Educational Research, vol. 33, no. 7, pp. 14-26.

Joughin, G. (2010) *A short guide to oral assessment*, University Wollongong: Leeds metropolitan university, pp.5-14.

Katheleen, S., and Kenji. K. (1996) *Testing communicative competence*, The education sources information center: Kyoto Japan, pp. 2.

Kathryn, H. R., and Geralyn, A. (2007) *Chesner developing voice through the language arts*, pp. 199.

Kumar, C. P. (2013) *The eclectic method: theory and its application to the learning of English international journal of a scientific and research publication*, vol. 3, no. 6, ISSN 225-3553, pp. 1-2.

Lado, R. (1965) *Language testing: A Teacher's Book*, London: Longmans.

Lamprianou, I. and Athanasou, J.A. (2009) *A teacher's guide to educational assessment,* Rotterdam, Netherland: sense publisher, pp.19.

Leong, L. and Seydeh, M. A. *(2017) An Analysis of Factors Influencing Learners' English Speaking Skills,* International Journal of Research in English Education, vol. 2, no. 2, March , pp. 34-39.

Levelt, W. J. M. (1989) *Speaking Form Intensive to articulation,* Cambridge: MIT Press.

Lin, C. C. (2004) *Introduction to linguistic Anthropology: The Nature of language.*

Luoma, S. (2004) *Assessing Speaking,* Cambridge: Cambridge University Press.

Madsen, H. S. (1982) *Determining the debilitative impact of tests anxiety*, Language learning 32.

Mayring, PH. (2000) *Qualitative Content Analysis, Forum: qualitative social research.* vol. 1, no. 2. June.

Na, Z. (2007) *A study of high school students' English learning anxiety*, Asian EFL J, vol. 9, no. 3, pp. 17-18.

Nicolsi, L., Harrman, E., a,d Keresheck, J. (1989) *Terminology of communication discourse baltimore*, MD: william & Wilkins.

Oxford, R. L*. (1990) Language Learning Strategies: What every teacher should know,* New York: Newbury House Publisher, pp. 140.

Parpura, J. E*. (1990) Studies in Language Testing Learner Strategy use and Performance on Language Tests: A structural equation modeling approach,* UK: press syndicate of the university of Cambridge, pp.15-16.

Phelan, c. and Wren, J. (2005) *exploring reliability in academic assessment*, [online], Available: http://chfasoa.uni.edu/reliabilityandvalidity.htm. [June 2005].

Richards, J. C., Li, B., and Tang, A. (1998) *Exploring pedagogical,* New York: Cambridge University Press.

Rivkin, Hanushek, and kain. (2005) *teachers, schools and academic achievement*, Econometrica.

Sass, et al (2010) *Value Added of Teachers in High Poverty and Low Poverty Schools*, Calder Working Paper.

Shohamy, E. (1995) *Performance assessment in language testing*, vol.15, pp. 188-211, [Online], Available: http://doi.org/10-2017/s0267190500002683. [19 November 2008].

Simensen, A. M. (1998) *Teaching a foreign language principles and procedures*. 2nd edition (2007). Fagbokferlaget vigmostad & bjorke AS, Bergen.

Skeham, P. (1990) *communicative language testing, the journal of TESOL-France*, vol. 10, no. 1, spring 1990, pp. 115-127.

Spielberger, C. D., and Vagg, P. R. (1995) *Test anxiety transactional process model*, Washington, DC: Taylor & Francis.

Spolsky, B. (1989) *Communicative Competence: language proficiency and beyond, aplied psycholinguistics*, vol. 10, pp. 138-156. [Online], Available:

https://doi.org/10.1093/applin/10.2138

Stoynoff, S. and Chapelle, C. A. (2005) *ESOL Tests and testing,* TESL-EJ, vol. 9, no. 3, December, pp. 1-3.

Take IELTS, *British Council*. [Online], Retrieved from:
http://takeielts.britishcouncil.org/prepare-test/understand-test-format/speaking-test
[september 2015].

Taylor, L. (2011) *Examining speaking research and practice in assessing second language speaking,* UK: Cambridge University Press, p. 306.

Thornbury, S. (2005) *How to teach speaking*, Pearson Education limited: Pearson Longman.

Underhill, N. (1987) *Testing spoken language: A handbook of oral testing techniques,* cambridge, UK: Cambridge University Press, pp. 07.

Weidman, J. C. (2013) *Perspective on mission differentiation in higher education between teaching and research*, Saudi journal of higher education.

Weir, C. J. (1990) *Communicative language testing*, London: Prentice Hall, pp. 07-12.

Wigglesworth, G. (2008) *Tasks and Performance: Based Assessment*.

Wiki How. July 2017. *How to Analyze Data Using SPSS*, [Online], Available:
https://m.wikihow.com/Analyse-Data-Using-SPSS?amp=1 .

William, D. (2008) *Balancing Dilemmas: Traditional theories and new applications*. In havens, A and Mcdowell, L (eds.), *Balancing Dilemmas in Assessment and learning in Contemporary Education,* New York: Routledge, pp.276-281.

Xu, et al (2009) *Making differences: The effects of Teach for America in a High school Calder, Working Paper 17 National Center for Analysis of Longitude*, Data in Education Research.

APPENDICES

Questionnaire to teachers

Dear teachers,

You are kindly requested to answer a set of questions whose aim is to collect some information about the assessment of students speaking skills. This questionnaire is a data gathering tool for a research work entitled 'Investigating the reliability of speaking tests: The Case of the English Department at MMUTO'.

Please tick the right box(es) ti indicate your answer, or give a full answer when necessary .

Your cooperation will be highly appreciated.

<u>Section one: Personal background</u>

1. Personal Position:

☐ MAB ☐ MAA ☐ MCB ☐ MCA ☐ prof

2. Teaching experience (in years) :

☐ 5 ☐ 6-10 ☐ More than 10

<u>Section two: Foreign language assessment and speaking test design</u>

3. To which approach would you akin your teaching techniques?

☐ Communicative Approach ☐ Grammar-translation Approach

☐ Competency-based Approach ☐ Eclectic Approach

Other ……………………………………………………………………………

4. Is the assessment of your student's performance?

☐ Daily ☐ Weekly ☐ Monthly ☐ Periodical

5. What do you think is the most reliable way to assess someone's English speaking skills?…………………………………………………………………………………

……………………………………………………….....................................

6. How do you select the topic to use in speaking tests/exams?

☐ I choose them myself.

☐ I ask the students what they like to talk about.

☐ I follow the course syllabus.

☐ I have another way. What is it please?

………………………………………………………………………………………

………………………………………………………………………………………...

7. You build speaking tests/exams . . .

☐ Alone   ☐ In collaboration with your colleagues   ☐ Relying on previous models

8. During the speaking test/exam, you are . . .

☐ The 'Examiner'   ☐ the co-communicator   ☐ Facilitator

9. What are the criteria you consider in assessing the student' speaking skills?

☐ a) Pronunciation.   ☐ b) Vocabulary.   ☐ c) Fluency

☐ d) Grammar.   ☐ e) All of them.

Section three: factors affecting student's speaking performance and speaking problems

10. You assess your students . . .

☐ Individually   ☐ In pairs   ☐ In groups

Why? Which type gives more reliable results?

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

………………………………………………………………………………………

………………………..

11. During the test/exam, most of your students are . . .

☐ Anxious   ☐ Relax and self- confident

12. When you assess your students' speaking skills, which competence do you target more?

☐   ☐   ☐

Linguistic   Communicative   equally, both of them

13. Do your students prepare for the speaking task before it is performed?

☐ Yes   ☐ No

14. Are your students given enough time to perform a speaking task?

☐ Yes   ☐ No

15. How do you score them?

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

16. In your opinion, which of the following factors may have a negative impact on your student's speaking performance during an oral test/exam?

☐ Lack of time for preparation   ☐ Lack of self-confidence

☐ Poor background knowledge   ☐ Foreign-language anxiety

☐ Examination anxiety   ☐ Misunderstanding of test instructions

17. Do you think that the students' scores reflect their communicative competence?

☐ Yes   ☐ Somehow   ☐ No

Why?...

....

....

18. Has it ever happened to you to set the same test to the same students after a relatively short period of time?

☐ Yes   ☐ No

19. If your answer is YES, have your student got the same scores?

        ☐ Yes   ☐ No

Thank you very much

Speaking tests

❖ Describe a museum

-how often do you visit the museum?

-which type of museum id popular in your country?

-why do you want to visit the museum?

❖ Describe your favorite photograph

-Where was the photo taken?

-Who took the photo?

-What can be seen from the photo?

-Explain why it is the favorite photograph for you?

❖ An important historic place

-What is the place?

-Where is it located?

-Why do you think it is important?

❖ Describe an area of your country you know and like

-Where is it?

-What are its special features?

-What do you and other people do in this area?

-Explain why you like it?

❖ Describe a party

-What is the party?

-Why was the party held?

-Who attended the party?

-What did you do for that party?

❖ Talk about a subject you are studying

-What subject are you studying?

-What is the subject about?

-How do feel about it?

-Explain why you are interested in it?

> ❖ What is the important invention before the age of computer

-What is the invention?

-Why do you think it very important?

-What are the advantages and disadvantages of the invention?

> ❖ Describe an interesting city you like

-Where is it?

-How did you go there?

-What did you do there?

-Explain why you find it so interesting?

> ❖ Describe your favorite movie

-What is the name of the film?

-What is the theme of the film?

-Why do you like it?

> ❖ Describe one of the shopping centre you often go to

-Where is the shopping centre?

-How often do you go to the centre?

-Why do you often go to it and characteristics?

> ❖ Describe your favorite animal

-What kind of animal is it?

-Describe it briefly?

-Why do you like the animal?

> ❖ Advertisement

-What are the forms of advertisement?

-What are the functions of the advertisement?

-What is the effect of advertisement to people?

❖ Describe your holidays

-Where do you go for holidays and how long do them last?

-Who do you go with?

-Talk about any interesting thing happened during your holidays?

❖ Choose one of the idiom……………………………………………………………..

-Can't judge a book by its cover.

-Kill two birds with one stone

-Birds of a feather flock together

-I have been on clouds nine all days

-Better to be the head of a dog than a tail of a lion

-Feel on the top of the world

-A taste of one's own medicine

-Big fish in a small pond

-Have butterflies in one's stomach

-You can't make an omelet without breaking the eggs

-As easy as apple pie

❖ What is the difference between present perfect and past perfect? Explain and give concrete examples.

❖ Describe your favorite actor or singer.

❖ Why did you choose English? What are your objectives?

❖ Choose a topic then express your opinion

-Globalization, women's leaderships, ethics in medicine, healthy food…

❖ Summarize a short story or a novel you have already read.

❖ Make a conversation by giving an advice and suggest something to your mate.

❖ Make a conversation where you ask for permission and accept or refute an offer.

❖ Choose one of the following prepositional verbs. Explain it and give concrete examples.

-Take after, take down, turn off, and take over

❖ Give four phrasal verbs providing an example to each.