

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY OF MOHAMED BOUDIAF-MSILA

Faculty of letters and Foreign Languages

Department of Letters and English language STREAM:

N°



DOMAIN: FOREIGN LANGUAGES

ENGLISH LANGUAGE

OPTION: LINGUISTICS

**Correlation between Subjective and Objective
Assessment of English Oral Proficiency
The Case of Third Year Students at Msila
University English Department**

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of
Master in Linguistics

Submitted by:

Hamdaoui Rafik

Bibi TakiEddine

Supervised by:

Mr. Tayoub Abdelmadjid

Board of Examiners

Dr. Imane CHERIET	University of M'sila	Chairperson
Mr. Abdelmajid TAYOUB	University of M'sila	Supervisor
Mrs. Dalal Omri	University of M'sila	Examiner

2023/2024

Dedication 1

To my dear mother's soul, to my father, may God protect him, to my wonderful wife, to my beloved little girl, Balkis, to my brothers, sisters, and friends, to my colleague in this work, Abdo Bibi, to the college administration, to my professors, whom we all deeply respect, to my supervisor and professor, the honorable Mr. Tayoub Abdelmadjid, who guided us in completing this work. To all of them, I dedicate this work.

Hamdaoui Rafik

Dedication 2

To my dear family

To my precious grandfather God's mercy on him

To my partner in this work and this university trip

To my friends and colleagues

Thank you for believing in me and supporting me during this journey

I gift this success for all of you

Bibi Taki Eddine

Acknowledgements

First, we would like to thank our supervisor Mr. Abdelmadjid Tayoub for being there for us when needed. A special thanks for the jury members for their time and effort in examining our work.

We would like to thank the English department staff of M'sila University. Their insightful guidance and constant support will always be remembered and acknowledged.

Abstract

This study investigates the correlation between subjective and objective assessments of English oral proficiency among third-year students at Msila University's English Department. The research aims to determine the extent of alignment between these two assessment methods and identify potential areas of discrepancy. Specifically, it examines the relationship between subjective evaluations of overall oral proficiency and subjective assessments of accuracy, fluency, and pronunciation constructs. Findings indicate strong positive correlations between subjective assessments of oral proficiency and the constructs of accuracy and fluency, with accuracy having the highest correlation. Pronunciation, while still significantly related, exhibited a slightly weaker correlation. The study also revealed that factors such as pause time and speech rate play significant roles in teachers' evaluations of fluency. These results underscore the importance of integrating both subjective and objective measures to achieve a comprehensive assessment of oral proficiency.

Key words: Accuracy, English oral proficiency, Fluency, Objective assessment, Pronunciation, Subjective assessment

List of Abbreviations

AR: Articulation rate

CBA: Competency – Based Approach

EFL: English as a Foreign Language

LMP: Mean length of pause

LP: Length of Pauses

MLR: Mean length of run

TNP: Total number of pauses in the sample

NP: Number of pauses

TNR: Total number of fluent runs in the sample

NR: Number of runs

NS: Total number of syllables in the sample

PAM: Proportion of accuracy mistakes

PPM: Proportion of pronunciation mistakes

PT: Pause Time

SR: Speech rate

T+P: Time including pause

T-P: Total time of the sample excluding pause time

List of tables

Table 1: Teacher 1 Evaluation of Students' Voice Recordings.....	48
Table 2: Teacher 2 Evaluation of Students' Voice Recordings.....	50
Table 3: Teacher 3 Evaluation of Students' Voice Recordings.....	52
Table 4: Teacher 4 Evaluation of Students' Voice Recordings.....	54
Table 5: Teacher 5 Evaluation of Students' Voice Recordings.....	55
Table 5 A: The correlation coefficients of the 3 criteria across the 5 teachers.....	57
Table 6: Evaluation of Student 1 Voice Recording.....	58
Table 7: Evaluation of Student 2 Voice Recording.....	58
Table 8: Evaluation of Student 3 Voice Recording.....	58
Table 9: Evaluation of Student 4 Voice Recording.....	59
Table 10: Evaluation of Student 5 Voice Recording.....	59
Table 11 Evaluation of Student 6 Voice Recording.....	59
Table 12: Combination of Teachers Evaluation with Students Objective Results.....	60
Table 13: Number of students' mistakes in the speech recoding.....	63

Table of contents

Dedication 1.....	II
Dedication 2.....	III
Acknowledgement.....	IV
Abstract.....	v
List of abbreviations.....	VI
List of Tables	vii
Table of Contents	viii

General Introduction

1. Introduction.....	2
2. Statement of the Problem.....	2
3. Aims of the Study.....	2
4. Research Questions and Hypotheses.....	3
5. Objectives of the Study.....	4
6. Significance of the Study.....	4
7. Literature Review.....	5
8. Structure of the Dissertation.....	6

Chapter One: Theoretical Framework

Introduction.....	9
1.1. Definition of Assessment.....	9

1.2. Types of Classroom Assessment.....	10
1.2.1. Formative versus Summative Assessment.....	10
1.2.2. Informal Versus Formal Assessment.....	11
1.3. Teaching, Learning and Assessment.....	11
1.4. Purpose of Assessment.....	13
1.5. The Function of Feedback.....	13
1.6. Definition of Test.....	14
1.7. Approaches to Language Testing.....	15
1.7.1. The Essay-Translation Approach.....	15
1.7.2. The Structuralist Approach.....	15
1.7.3. The Integrative Approach.....	16
1.7.4. The Communicative Approach.....	16
1.8. Types of Test.....	17
1.8.1. Achievement versus Proficiency Tests.....	17
1.8.2. Aptitude Tests.....	17
1.8.3. Diagnostic Tests.....	18
1.8.4. Norm-Referenced versus Criterion-Referenced Tests.....	18
1.8.5. Subjective Versus Objective Tests.....	18
1.9. Criteria of a Good Test.....	19

1.9.1. Reliability.....	19
1.9.1.1. Student-Related Reliability.....	20
1.9.1.2. Rater Reliability.....	20
1.9.1.3. Test Administration Reliability.....	20
1.9.2.1. Face Validity.....	21
1.9.2.2. Content Validity.....	21
1.9.2.3. Consequential Validity.....	21
1.9.2.4. Predictive Validity.....	22
1.9.2.5. Construct Validity.....	22
1.10. Practicality.....	22
1.11. Authenticity.....	23
1.12. Washback.....	23
1.13. Scoring Rubrics.....	25
1.13.1. Analytic Rubrics.....	25
1.13.2. Holistic Rubrics.....	25
1.14. Language Oral Proficiency and Speaking.....	27
1.14. Language Speaking Skill.....	28
1.14.1. Background and Conceptualization of Language Speaking Skill.....	28
1.14.2. The Components of Language Speaking Proficiency.....	30

1.14.2.1. Grammar.....	31
1.14.2.2. Vocabulary.....	31
1.14.2.3. Pronunciation.....	32
1.14.2.4. Fluency.....	33
1.14.2.5. Comprehension.....	33
1.14.3. The Significance of Enhancing L2 Speaking Skill.....	34
1.15. Communicative Language Teaching CLT.....	34
1.15.1. Background Overview of the CLT Approach.....	35
1.15.2. The Main Objective of the CLT Approach.....	35
1.15.3. Principles Underlying the Communicative Language Teaching Approach (CLT).....	36
1.15.4. CLT as a Basis to Develop EFL Speaking Skills.....	37
1.16. Types of CLT Speaking Activities.....	38
1.17. Teaching Techniques and Activities in CLT.....	40
1.17.1. Role-Playing Activities.....	40
1.17.2. Problem-Solving Activities.....	41
1.17.3. Interviewing Activities.....	41
1.17.4. Using Authentic Materials.....	42
1.18. Assessment of Oral Proficiency.....	42
Conclusion.....	44

Chapter Two: Methodology and Data Analysis

2.1. Research Methodology and Data Collection.....	46
2.1. 1.Research Design.....	46
2.1. 2. Participants.....	46
2.1.3. Data Collection.....	46
2.1.4. Data Analysis.....	47
2.1.5. Instruments.....	48
2.2. Data Analysis, Findings, Discussion and Recommendations.....	48
Conclusion.....	67
General conclusion.....	69
Implications of the Study.....	71
Limitations.....	74
References.....	76
ملخص.....	83

General Introduction

1. Introduction

English language learners, particularly those enrolled in English as a Foreign Language (EFL) programs, undergo regular assessments to gauge their oral proficiency. These assessments typically involve a combination of subjective and objective measures. Subjective assessments rely on the judgment of human raters, while objective assessments utilize standardized criteria or scoring rubrics.

This study explores the correlation between subjective and objective assessments in evaluating EFL learners' oral proficiency. Understanding this relationship is crucial for developing effective assessment practices that accurately reflect students' abilities.

2. Statement of the Problem

EFL instructors often face challenges in ensuring consistency and fairness during oral proficiency assessments. Subjective evaluations, while valuable in capturing a learner's overall communication skills, can be susceptible to rater bias. Conversely, objective assessments may lack nuance and fail to capture the full range of a learner's spoken English abilities.

The lack of a strong correlation between subjective and objective measures creates a dilemma for instructors. This study aims to investigate the extent to which these two assessment methods align and identify potential areas of discrepancy.

3. Aims of the Study

This study aims to achieve the following objectives:

- Investigate the correlation between subjective and objective assessments in measuring EFL learners' oral proficiency.

- Investigate the correlation between subjective assessments of oral proficiency and subjective assessments of accuracy, fluency and pronunciation constructs?
- Analyze the strengths and weaknesses of each assessment method.
- Provide recommendations for developing a more comprehensive and reliable oral proficiency assessment approach.

4. Research Questions and Hypotheses

To achieve the outlined aims, this study will address the following research questions:

- To what extent do subjective and objective assessments correlate in evaluating EFL learners' oral proficiency?
- To what extent do subjective assessments of oral proficiency correlate with subjective assessments of accuracy, fluency and pronunciation constructs?
- How do subjective and objective assessments differ in their ability to capture various aspects of spoken English proficiency?

Hypotheses

This study proposes the following hypotheses:

- H1: There is a positive correlation between subjective and objective assessments of EFL learners' oral proficiency.
- H2: There is a positive correlation between subjective assessments of oral proficiency and subjective assessments of accuracy, fluency and pronunciation constructs.
- H3: Subjective assessments will better capture a learner's overall fluency, while objective assessments will be more effective in measuring specific grammatical accuracy, fluency and pronunciation features.

5. Objectives of the Study

This study aims to investigate the correlation between subjective and objective assessments in order to determine the extent of alignment between subjective and objective assessments in measuring EFL learners' oral proficiency. It also examines the correlation between subjective assessments of overall oral proficiency and subjective assessments of accuracy, fluency, and pronunciation constructs. The third objective is to evaluate the respective strengths and limitations of subjective and objective assessments in capturing various aspects of spoken English proficiency. The ultimate objective of the study is thus to develop and propose recommendations for creating a more comprehensive and reliable approach to assessing oral proficiency, incorporating both subjective and objective measures. These objectives aim to enhance the understanding of how different assessment methods align and to improve the effectiveness of oral proficiency evaluations in EFL contexts.

6. Significance of the Study

The significance of this study lies in its potential to enhance the understanding and practice of oral proficiency assessment in EFL contexts. By elucidating the relationship between subjective and objective evaluations, the study provides valuable insights into the strengths and limitations of each method. This has several important implications:

- **Improvement of Teaching Practices:** The findings offer evidence-based guidance for EFL instructors on how to design and implement more effective assessment strategies.
- **Development of Comprehensive Assessment Tools:** The study highlights the need for assessment tools that integrate both subjective and objective measures. Such tools can provide a more rounded and fair evaluation of students' oral proficiency, capturing both the nuanced aspects of spoken language and specific linguistic features.

- The study could be particularly useful in improving rater reliability and reducing biases in subjective assessments.
- **Support for Student Learning:** By identifying the key constructs that influence oral proficiency, the study offers insights for students aiming to improve their spoken English.
- **Contribution to Academic Research:** The study adds to the body of knowledge in the field of language assessment.

Overall, this study aims to foster more reliable, fair, and comprehensive assessment practices, ultimately contributing to better learning outcomes for EFL students.

7. Literature Review

This review will focus on studies investigating the correlation between subjective and objective measures of speech, the strengths and weaknesses of each approach, and factors influencing the assessment process.

The two primary methods for assessing speaking skills are: holistic scoring and analytic scoring. Holistic scoring, also termed impressionistic or global scale, provides an overall score based on the entire performance, while analytic scoring breaks down the performance into specific criteria and evaluates each independently (Taylor & Galaczi, 2011; Pan, 2016). Holistic scoring is noted for its simplicity and efficiency compared to analytic scoring but offers less detailed information about learners' language abilities. In contrast, analytic scoring offers more nuanced insights into learners' strengths and weaknesses, leading to increased rating accuracy and reliability (Kondo-Brown, 2002). Analytic scoring criteria typically include dimensions such as fluency, vocabulary, accuracy, interaction, coherence, pronunciation, and intonation. The number of criteria used can vary, with suggestions ranging from three to seven categories, although there is ongoing debate about the optimal number (Pan, 2016; Council of Europe, 2001). In summary, while holistic scoring is less time-consuming and complex, analytic scoring

provides deeper insights into speaking proficiency, although it requires more effort and training for assessors (Taylor & Galaczi, 2011; Kondo-Brown, 2002).

The analytic approach in speaking assessment involves examining various aspects of the test separately and assigning scores to each aspect individually (Richards & Schmidt, 2013). This method provides valuable diagnostic information about a learner's speaking abilities (Tuan, 2012). It increases consistency and reliability (Jonsson & Svingby, 2007), and contributes to a more objective and dependable evaluation of learner work (Finson, Ormsbee, & Jensen, 2011). However, its drawbacks include being time-consuming and tedious (Aleksandrak, 2011), and needs examiners' training (Vafaei & Yaghmaeyan, 2015). Despite these drawbacks, its advantages outweigh the disadvantages, making it a suitable method for assessing speaking performance.

From a linguistic perspective, Fillmore (1979) provided a logical framework for understanding first language (L1) fluency phenomena, connecting these to broader conceptualizations of proficiency. He emphasized the significance of recognizing formulaic expressions in everyday assessments of a speaker's fluency. Fillmore identified four abilities indicative of fluency in a native language: first, the capacity to speak at length with minimal pauses and to fill time with speech; second, and the ability to communicate in coherent, reasoned, and semantically dense sentences. We will also explore the differences between subjective evaluations by teachers' objective measurements by Praat software.

8. Structure of the Dissertation

This study is organized into two chapters and an introductory section. The first section introduces the research problem, question and hypotheses and the methodology adopted. Chapter one introduces the theoretical groundwork, identifying, testing, oral proficiency and its constructs. The chapter then establishes a framework for key concepts testing the speaking

construct. Chapter three presents the field study through calculating the correlation coefficients between the variable interpreting and drawing conclusions

The last section presents the answers to the questions and hypotheses tests and closes with recommendations and a call for further research.

Chapter One:

Theoretical Framework

Introduction

The term "test" and "assessment" are often used interchangeably. But in educational research the two terms are used separately because they connote two different meanings. In fact, tests are one of the most familiar methods of assessment that teachers rely on to assess their students' achievements and tasks (Brown, 2004).

Since tests are one feature of assessment, we cannot talk about testing without mentioning assessment. Thus, this chapter is divided into two sections. The first one is concerned with the basic elements of assessment where we try to know about the role played by assessment in teaching and learning process, its definition, types, purpose, and the role of feedback.

The second section of this chapter is devoted to testing. We begin by giving its definition, its approaches, its types, and the qualities of a good test. Then, we will deal with the scoring rubrics.

1.1. Definition of Assessment

It is worth mentioning that the term assessment has been used differently in current educational research. In fact, there is no agreement on a consistent definition of what assessment exactly is because each educator uses it in a specific area to reach a given goal (Bachman, 2004).

Bachman (2004, p.7) sees assessment as "the process of collecting information about a given object of interest according to procedures that are systematic and substantively grounded". By *systematic*, the author implies that assessment is done in an organized method with fixed and explicit standards so that other persons can repeat it whenever they want. In other words, assessment is conducted in clear procedures that are subjected to "public scrutiny", i.e. those procedures make a liaison between things we wish to assess and our observations. To illustrate this situation, we may describe someone's qualities

depending on a dialogue we have done with him, but this would not be an assessment because it is nearly impossible to re-describe him exactly as the first time. Also, someone would describe the same person in a different way, so there is no systematicity in the description. This systematicity in assessment is largely represented by reliability. Furthermore, assessment should be '*substantively grounded*' where it should be controlled with a clear theory about the language use, language learning, language ability, ex-studies in the domain, etc. Clearly, in the previous example, the description of the person is not based on substantive condition because people generally do not describe other people to assess their language abilities. The substantive quality in assessment is largely represented by validity (Bachman, 2004, p. 7).

The Assessment Reform Group (ARG, 2002; as cited in Gardner, 2006, p. 2), on the other hand, considers assessment as "the process of seeking and interpreting evidence for use by learners and their teachers, to identify where the learners are in their learning, where they need to go and how best to get there". So assessment is a process of gathering information that help both teachers and students to have a clear image about their current achievements as well as about their shortages to work on them in future.

1.2. Types of Classroom Assessment

Assessment can take a variety of forms. Classroom assessment is generally divided into four major categories which are: formative versus summative assessment, informal versus formal assessment. There exists other types of assessment but they are rather types of short tests. They would be presented shortly after in the coming.

1.2.1. Formative versus Summative Assessment

Formative assessment has been defined as: "the process of appraising, judging or evaluating students' work or performance and using this to shape and improve students' competence" (Bell & Cowie, 2002, p. 11). In fact, formative assessment is an ongoing assessment that allows instructors to check their learners' progress and identify the areas of

Troubles or the skills that students them and improve student learning. Also, the information gained could be used to revise the teaching method used.

Summative assessment, on the other hand, is used to classify candidates according to their performances and marks, deciding who will succeed and moving to the next year or getting the diploma and who will fail, etc. Cizek.Cizek (2010, p. 3) states that any "test or other system" of collecting data is considered as summative assessment if it fulfills the following two conditions: firstly, it is conducted at the end of the instructional term such as semester or the year. Secondly, it provides data for selection. This is why the validity and reliability of summative assessment have more importance than the other type of classroom assessment.

1.2.2. Informal Versus Formal Assessment

Informal assessment aims to find out *how* a learner can accomplish a given task. It is interested by the process of learning itself. It includes systematic observation where the teachers can assess their student's work indirectly without making use of standardized tests and scoring patterns. For instance, the teachers can comment on their students' work by telling them that they can do better, so the students would be assessed without noticing that.

On the other hand, formal assessment is a formal tool that is used to know how much a student has learnt or improved during the instructional period; it is concerned with *how well* a learner has learned concepts and developed the target skills. Formative assessment relies on standardized methods of administering the tests, and it is based on clear criteria for evaluation. Unlike the informal type, the results of the formal assessments are published (Farr & Beck, 2003).

1.3. Teaching, Learning and Assessment

In EFL/ESL classes, assessment goes hand in hand with instruction. We cannot teach without relying on assessment (Brown, 2004). Assessment is a planned process designed to

accomplish a specific educational purpose. In fact, the role of assessment now is not confined only to test the students, but it extends to assess the whole educational system.

It is worthwhile pointing out here that the role of assessment has shifted from traditional "strategies" such as standardized tests, chapter-end exams or year-end exams, etc, and which have been applied to fit the purposes of persons who function outside the classroom such as: school-masters, parents, supervisors, etc. To more overall role which is no longer limited in discriminating and ranking students but to assess syllabus and the teaching method to help students overcome of deficiencies. The assessment process, therefore, should be continuous and progressive; it should not be constrained only to the end of the semester or the term (Gitomer&Duschl, 1995). As Brown (2004, p. 4) argues "A good teacher never ceases to assess students, whether those assessments are incidental or intended".

In order to understand more the relationship between assessment and the learning/teaching process, James and Gipps (1998) state that there are four ways in which assessment affects learning. Firstly, assessment is a good tool for motivating students. That is, it gives them the sense of accomplishment and success, so they feel confident. Secondly, assessment provides an opportunity for students to determine what to learn by spotlighting their weaknesses. Thirdly, assessment helps students to know how to "learn" by knowing the learning style that fits their needs. It affects, also, their choice to the learning strategies used. In addition, it teaches students how to manage and control their skills and helps them to develop their knowledge, skills and their understanding in general. Finally, assessment affects learning by giving students the chance to evaluate their

progress by assessing their current level, boosting their "existing learning", or by encouraging the new one.

1.4. Purpose of Assessment

There are several purposes behind assessment. The National Research Council, in the USA, (1999; as cited in Bell & Cowie, 2002) summarized these purposes under three points. First of all, it has the aims of *summative assessment* as it enables instructors, administration, parents, and government, etc, to check learners' development and to make judgments about a student's level of performance, and certify which students may pass to the next level of their studies. Summative assessment is very important since it urges government to introduce reforms and make decisions about staffing and resourcing. The second objective of assessment is named *formative* where it is used to assess both instructors and learners to identify the area of weakness to build on in order to improve the quality of learning. The last purpose is called *accountability assessment*. In this case, assessment information can be used to undertake policy review and development at a national level, so the government intervention is targeted appropriately to change the global educational policy of the whole state.

1.5. The Function of Feedback

Feedback is recognized to be an important part of the learning cycle. Feedback is defined as "information that is given to the learner about his or her performance of a learning task, usually with the objective of improving this performance"(Ur, 1996, p. 242). So, it is the comments given by the teacher to tell their students how well or bad they perform.

Feedback affects learning/ teaching in two aspects: cognitive and motivational. On the one hand, feedback influences learning cognitively by making the students more aware about their progress, and about what is necessary to perform better in the coming stages.

On the other hand, feedback can impact students' motivation by developing the sense of accomplishment as they have a clear image about their learning.

Feedback plays a significant role in promoting learning and achievement if it respects some strategies namely: timing, amount, mode and audience. Firstly, by timing, teachers should know when and how to submit feedback. It should be immediate so that the students find the opportunity to use it at the right time before they forget something about the work done. Secondly, the amount of feedback should be considered where the teachers have to select the required information to give, and the level of students where some students need to receive much more feedback than the others. Thirdly, the teachers have to choose the suitable mode of feedback either orally, written or visually. For example, it is advisable to supply the students with oral feedback in interaction works. Finally, feedback is affected by the audience which can be single student, a group, or the whole class (Brookhart, 2008).

1.6. Definition of Test

Many definitions have been given to the term test. One of the overall definitions is provided by Brown (2004, p. 3) where he defines it as: "a method of measuring a person's ability, knowledge, or performance in a given domain". By method the writer means that test is composed of particular "techniques, and procedures, or items" that should be clear, obvious and organized. For instance, if we use multiple choices questions, we ought to enclose them with a list of potential correct responses. In addition, test is an instrument of measuring. The measurement can be general where the focus is on evaluating general abilities such as writing texts and essays or specific when the focus is on measuring specific skills such as the appropriate use of articles in writing texts. A test is designed to evaluate a learner's ability, knowledge, or performance. That's why test-makers have to consider the test-takers 'background, the validity of the test, and the examinees' Interpretation of their scores. Testing examinees 'performances display the level of their

competence. Generally speaking, tests are designed to evaluate one capacity of performing the language such as to speak, to read, to write, or to listen of the target language. In fact, the "Performance-based tests" are used to survey test-takers' use of the language, and from this survey the examiners can deduce the level of their general competence. To illustrate, reading comprehension tests that are constituted of several passages followed by quizzes about the general understanding may use to deduce the level of the reading competence. Next, the test targets a given domain. For example, in vocabulary test, examiners have to cover a list of items or words studied in a given unit, and the phonetic test might deal with minimal pairs.

1.7. Approaches to Language Testing

Language testing as a crucial component of language teaching has been influenced by many other fields and sciences such as: psychology, psycho-pedagogy, applied linguistics, sociology, statistics, etc. In fact, language testing has undergone four major stages: the essay-translation approach, the structuralist approach, the integrative approach, and the communicative approach (Heaton, 1988).

1.7.1. The Essay-Translation Approach

This stage is considered to be pre-scientific era of language testing because the tests are designed intuitively without depending on scientific principles and procedures. As its name indicates, the tests of this approach focus on writing essays, translation, and grammar. It should be noted that the tests of such an approach are characterized by high extent of bias to culture and literature, as well as, the judgment was described to be intuitive, subjective and unreliable.

1.7.2. The Structuralist Approach

As a consequence to the failure of the previous era to fit the aim of both language teachers and learners and as a result to the development in the field of linguistics and psychology, a new approach has emerged and which is called the structuralist approach, in

other books psychometric- structuralist approach. Actually, this approach has been influenced by the structural linguistics and by the behaviourism theory. This approach based on the claim that each language component (grammar, phonology, and vocabulary) as well as each language skill (speaking, listening, reading, and writing) should be tested at a time to ensure the effectiveness of measurement. The structuralist tests are more valid, reliable as they are statistics-based exams.

1.7.3. The Integrative Approach

This approach involves testing the language skills and components integrally, i.e., tests designed under the basics of such approach combine two or more skills at a time with the emphasis on the meaning and context. Integrative tests commonly use cloze testing, dictation, oral interview, translation and essay writing (Heaton, 1988).

Cloze testing is " a reading passage (perhaps 150 to 300 words) in which roughly every sixth or seventh word has been deleted; the test-taker is required to supply words that fit into those blanks" (Brown, 2004, p. 8). These tests reflect the testees' competence such as: "knowledge of vocabulary, grammatical structure, discourse structure, reading skills and strategies, and an internalized "expectancy" grammar".

Another major type of integrative test is dictation where the teacher reads a passage of hundred words and learners write it down. Dictation requires the use of many language abilities as the examinees are going to receive the information, "analyse and synthesize" them. Moreover, both listening and writing skills are stimulated (Aslam, 1992).

1.7.4. The Communicative Approach

The need for tests that are constructed of real language use which simulates real life context makes the new perspective of language testing. This new trend has produced the so-called communicative approach. The crucial aim of communicative tests is testing the learners' abilities to communicate effectively. This approach is based on the idea that Communicative competence could be tested through examinees' communicative

performance. By communicative competence it is meant both grammatical competence which presents the language usage and the socio-linguistic and strategic competence which presents the language use (Aslam, 1992).

1.8. Types of Test

The language tests can be embodied in several types. The distinction between test types generally is based on different criteria, purpose, the way of response, and the scores interpretations, etc. Accordingly, we distinguish achievement versus proficiency tests, diagnostic tests, Direct versus Indirect tests, Norm-Referenced and Criterion-Referenced tests, and Subjective versus objective tests.

1.8.1. Achievement versus Proficiency Tests

The achievement test, as its name indicates, is a method which attempts to measure how much students earned or achieved linguistic knowledge and skills during an educational unit. The tasks used mirror what the learners' have studied. Also, it is used to check the efficacy of teaching method used (Davies, 1999; Sharma, 2002).

The proficiency test; however, is defined by Allison (1999, p. 80) as tests which are used to "measure a person's language ability, irrespective of how this ability has come about". In other words, this type does not adhere to a particular course of study or a special book; rather it measures general skills, whatever its source. Actually, there are several examples of this type, including the TOEFL, and the Cambridge Proficiency Examination (CPE).

1.8.2. Aptitude Tests

Aptitude test is usually applied to give a look at the general linguistic capacities and skills the examinees have. It specifies whether a student has "any special aptitude "to learn

the target language in future or not (Saraswati, 2004).

1.8.3. Diagnostic Tests

Diagnostic tests are prepared to indicate and reveal any difficulties the learners' have. Using this test allows instructors to know their learners' "strengths and weaknesses".

Subsequently, based on those results, the teachers review their syllabi, or teaching method to remedy this shortage (Saraswati, 2004).

1.8.4. Norm-Referenced versus Criterion-Referenced Tests

In fact, the norm-referenced tests are designed to classify the examinees according to their level. The results obtained from this kind of test are used to "compare" between the test-takers' performances. Then, "discrimination" is set to "rank" students' work from the best to the least (Allison, 1999; Brown, 2004). The Scholastic Aptitude Test (SAT) and TOEFL are considered to be norm-referenced tests (Brown, 2004).

The criterion-referenced tests, on the other side, are used to "determine how an individual performs with respect to some "criterion" (Allison, 1999, p. 82). So the ultimate aim is to know if the examinee has developed the required skill or not compared to a specific criterion. According to Brown (2004) this type of tests allows test developers to supply the examinees with precise "feedback" presented in marks.

1.8.5. Subjective Versus Objective Tests

In the case of subjective tests, the testers are going to correct the testees' answers and supplying scores depending on their own estimation as there is no one fixed correct answer to the tasks given. In other words, scorers rely on subjective judgment in correcting the examiners' responses. In fact, a well known example of this kind is essay test.

Whereas the objective tests determine one answer to each question, so when correcting

this test, there would be no room for any kind of subjectivity. There are various types of objective questions such as: true/false questions and multiple-choice questions, etc. Unlike the previous type, the objective tests supply consistent scores as correctors depend on uniform and precise rating scale (Taiwo, 1995).

It is worth mentioning that there other types of testing, but these types are the common used ones. In fact, the test developers can combine several kinds in the same test. For example, the testers can put together the subjective and objective tests to ensure validity as the dominant feature in subjective tests, and reliability which is also the dominant feature in objective tests.

1.9. Criteria of a Good Test

Since test results affect individuals, as well as the educational system, a great consideration should be given to its qualities. In order to judge any test to be good enough, it should respect a number of criteria. These criteria can be summarized as follows: reliability, validity, practicality, authenticity, and washback.

1.9.1. Reliability

A given test is considered to be reliable if it gives unchangeable results to test-takers. That is to say, the same learner should get the same score if the same test is repeated (Brown, 2004; Bachman & Palmer, 1996; Broughton et.al, 2003; Heaton, 1988). Bachman and Palmer (1996) state that there are many factors that may affect test reliability, so the scorers should know them in order to avoid them and guarantee the "consistency of scoring". Brown (2004), in his turn, sums up those factors under four types namely: Student-related reliability, rater reliability, test administration reliability, test reliability.

1.9.1.1. Student-Related Reliability

It is believed that the "physical or psychological" condition of students, such as sickness or test anxiety while test-taking may affect their scores and so the test reliability (Mousavi, 2002, as cited in Brown, 2004). Anxiety is defined as "an unpleasant emotion experienced as dread, scare, alarm, fright, trepidation, horror or panic". In fact, anxiety could be "facilitative" or "debilitative". A reasonable amount of anxiety may encourage the learners to reach the approval performance. Contrarily, the "excessive amounts of anxiety" debilitate and hinder the learners' performance (Lewis, 1970, as cited Birjandi, 2010, p. 44). Besides, motivation has a crucial impact on test results because less-motivated students Towards test-taking are likely to have low marks (Broughton et al., 2003)

1.9.1.2. Rater Reliability

The reliability can be affected by tester's intuitive judgment or erroneousness. In fact, Brown (2004) splits rater reliability issues into two types: Inter and intra. Inter-rater reliability takes place when the testers do not adopt a uniform scoring scale for the same tests, and this may happen due to the lack of training, focus and expertise, all these factors lead testers to score changeably the same students. The intra-rater reliability occurs during the test-taking when the scorers use an inaccurate and wobbling scoring system, and because of exhaustion, or due to their pre-judging of their learners' level.

1.9.1.3. Test Administration Reliability

Actually, the way of test administration may impact its reliability. The circumstances surrounding the test setting, such as noise, lighting quality, size of the room, and the placement of tables and seats, etc, might affect the consistency of scoring. For example, conducting oral test based on listening and comprehension of tape recorder with high

amount of noise outside would not reflect the reliability especially for students sitting near the windows (Brown, 2004).

1.9.2.1. Face Validity

Weiner and Craighead (2010, p. 637) define face validity as "the degree to which test respondents view the content of a test and its items as relevant to the context in which the test is being administered". This means that face validity is based on the test-takers' subjective opinion about the usefulness and suitability of a test.

1.9.2.2. Content Validity

To insure content validity, examiners have to adhere to the topics dealt with in the curriculum in designing their tests, so that the examinees would be more familiar with them (Cohen et.al, 2018; Black, Op.cit). It is argued that the content validity is achieved if the test assesses effectively the required behaviour (Brown, 2004). To illustrate, testing examinee's communication skills should be through involving him in a conversation that stimulates the real life situations. Otherwise, if this test is done through written form, it wouldn't reflect content validity.

1.9.2.3. Consequential Validity

Before designing and submitting a test to the learners, the examiners have to establish a clear purposes and goals to be achieved by this test. Accordingly, consequential validity means to interpret and employ test results to compass objectives set previously (Cohenet.al, 2018). In addition, that testing has a direct impact on examinees' enthusiasm, their future educational performance, and their learning strategies, so test designers should account for the "consequences" of conducting a test (Brown, 2004).

1.9.2.4. Predictive Validity

This criterion demonstrates the ability of predicting the learners' future achievement based on their current testing outcomes. In other words, based on examinees results, teachers can foresee their future gaining and accomplishment (Cohen et.al, 2018).

1.9.2.5. Construct Validity

Construct is defined to be "any theory, hypothesis, or model that attempts to explain observed phenomena in our universe of perceptions" (Brown, 2004, p. 25). Those constructs could be linguistic such as: "proficiency" and "communicative competence", or psychological like: "self-esteem" and "motivation". Construct validity conveys the degree of success of a test in measuring a given hypothetical concept. To do so, the testers should set up all the aspects of that construct. For example, if we have to assess someone's oral proficiency, we should address all its aspects, "pronunciation, fluency, grammar accuracy, vocabulary use, and socio-linguistic appropriateness". Therefore, if we drop one side, the construct validity would be affected.

1.10. Practicality

The test reflects the practicality if it takes into account many conditions which can be summarized in terms of time and accessible resources. The test is "usable" if it can be used in the allotted time without exceeding the available capabilities (Broughton et.al, 2003). To facilitate judging the "usefulness" of a test, Bachman and Palmer (1996) have provided an equation to calculate its practicality:

$$\text{Practicality} = \frac{\text{Available resources}}{\text{Required resources}}$$

If practicality ≥ 1 , the test development and use is practical

If practicality < 1 , the test development and use is not practical Figure

01: Practicality (Bachman & Palmer, 1996, p. 36)

They also divide resources into three types: "Human resources" that represent the manpower who would conduct the test as test-makers, administer, scorer, etc. "Material resources" which represent the test setting, the tools used like computers, audio visual tapes, and papers, etc. finally, time which is calculated from the beginning of the preparation and design to the last stage of evaluation and marking

1.11. Authenticity

Authenticity is a prerequisite to ensure the quality of a test. According to Bachman and Palmer (1996), authentic tests require the use of tasks and problems from real life situations. The scores of this type of tests reflect the test-takers' real achievement and linguistic capacities, so tester can "generalize" their result to predict their performance and flexibility in real life setting. This is confirmed by Sapolsky (1985, p. 31, as cited in Geranpayeh& Taylor, 2013) "lack of authenticity in the material or method used in the test weakens the generalizability of results". In addition, it is believed that authentic tests have beneficial effects on the learners' performance. The more tasks resemble reality the more test-takers are motivated to perform better.

1.12. Washback

A final characteristic of a good language test is the washback, called also backwash. Washback "refers to the influence of testing on teaching and learning" (Cheng and Watanabe, 2004, p. 4). In fact, several studies have been conducted to determine its real impact and to discover how possible this impact could be exploited to improve the quality of instruction. According to these studies, tests can have positive as well as negative influence on teaching/learning.

On the one hand, tests can trigger a negative washback on the educational system. When

the instructors teach for the exam, the aim would be shifted from teaching and developing the learners' abilities and knowledge to teaching for passing exams. In addition,

Learners become more interested with the content included in tests and, usually, neglect the other sides of knowledge. Likewise, Fish (1988, as cited in Cheng & Watanabe, 2004) talks about the great psychological stress faced by teachers because of the exams, especially at the stage of results announcement. This stress is inversely proportional with testers' age and experience where the younger and less-experienced instructors usually suffer more. Finally, Smith (1991, p. 8 as cited in Cheng & Watanabe, 2004) recapitulates the other negative points where he claims that:

testing programs substantially reduce the time available for instruction, narrow curricular offerings and modes of instruction, and potentially reduce the capacities of teachers to teach content and to use methods and materials that are incompatible with standardized testing formats

On the other hand, no one can deny the positive washback the test may have. In fact, the test results not only used to remedy the weaknesses of the curriculum but also used to review the teaching method as a whole. In addition, Pearson (1988, as cited in Cheng & Watanabe, 2004) argues that the tests can have positive effects only if they succeed in realizing their intended objectives and the course objectives. Besides, Alderson and Wall (1993, as cited in Cheng & Watanabe, 2004) believe that it is worthless to talk about positive or the negative washback as it is related to the quality of the test itself. Clearly, the well considered tests are likely to have a positive washback. Otherwise, the opposite is true. Alderson and wall (1993, as cited in Cheng & Watanabe, 2004) suggest not focusing on determining if a test has positive or negative washback, rather the emphasis should be on determining and then avoiding the social and educational conditions that threaten its efficacy.

To conclude, there could be a "conflict" between two or more criteria; for example, between reliability and validity because in some cases it could be a test that is classified to be reliable but it is invalid, and vice versa. To illustrate, multiple choice questions tests are considered as reliable tests; however, they are less valid. This conflict should not make Abandonment of one criterion in the favor of another. On the contrary, the key solution is to create a "balance" which varies from one test type to another and from one situation to another (Bachman & Palmer, 1996).

1.13. Scoring Rubrics

In testing EFL and ESL learners, two broad types of scoring rubrics are distinguished, holistic and analytic rubrics. Each type of them has specific uses in accordance with the circumstances. And as such, each type has its pros and cons.

1.13.1. Analytic Rubrics

Analytic rubrics "describe work on each criterion separately"(Brookhart, 2013, p. 6). Understandably, the scorers give each component of performance an independent scores relying on clear and precise criteria. To illustrate, figure 2 presents an example of analytic rubrics. In this example, the examinee can have 3 in mechanics and 4 in word use. Analytic rubrics have several advantages. They supply the teachers with an overview about the areas of "strengths and weaknesses "of their learners 'language knowledge and abilities, as well as, they supply students with beneficial feedback. In addition, the scale is often clear and very detailed, so it allows for a rigorous analysis. However, they are time consuming as a lot of details would be analyzed (Nichols & Nichols, 2005).

1.13.2. Holistic Rubrics

A holistic rubric "gives a single score or rating for an entire product or performance

based on an overall impression of student's work" (Arter & McTighe, 2001, p. 18). Accordingly, the holistic rubric deals with the learner's work as a unit. One example about the holistic rubric is provided in figure 3 below adapted from Blaz (1999, p. 28)

1. **The response does not complete the assignment.**

- Information may be missing or inaccurate.
- There are problems with accuracy and logic.
- Overall impression: incomplete and unsatisfactory.

2. **Standard barely met.**

- Information provided is generally accurate.

There may or may not be a conclusion or an opinion.

If one or the other is offered, there may be problems with accuracy and logic.

3. **Meets standard.**

- Information is accurate.
- A logical conclusion or an opinion is offered.
- Writing is fluent but not interesting.
- The answer is lengthy rather than concise.

4. **Exceeds standard.**

- Information is accurate and writing is fluent and lively.
- Answer is concise and to the point.
- Conclusions and/or opinions are logical.
- Overall impression: complete and satisfactory.

This frame is considered to be faster and less consuming of time. On the other side, they have their downside. In fact, holistic rubric is not diagnostic in its nature, so instructors, as students, haven't a clear idea on the area of weaknesses to be remedied.

Brookhart (2008) sees the analytic rubrics as more useful in the instructional classes for both summative and formative assessment. Focusing on testing all the components of examinees' performance provides reliable results to use in discriminating the learners levels (summative purposes). Also, as mentioned before, knowing learners' weaknesses allows them and their teachers to rectify them in future (formative assessment).

In the preceding review, we bring to light the nature of the relationship between language assessment and testing where the later is considered to be one feature of the former. We provided a general view on language assessment including its definition, types, role in learning/teaching process, its purposes, and the role of feedback. Additionally, we focus on language testing where we moved among its various aspects. We start by giving a definition, its historical developmental stages, and its common forms. Fundamentally, we offered a detailed description of useful test, the well known frame of speaking tests, and scoring rubrics. Finally, we conclude with the common problems associated with tests. Have not developed well, so they take steps to overcome. Having introduced assessment as a field of study very crucial for teaching, we now narrow down to the speaking skill as the focus of the current study

1.14. Language Oral Proficiency and Speaking

Effective communication in English is crucial in today's globalized world, where proficiency in English is often required for academic, professional, and personal reasons. We outlines a literature review regarding two key variables relevant to the study of improving the speaking skill of adult intermediate-level English as a Foreign Language (EFL) learners through the Communicative Language Teaching (CLT) approach. We now provide a theoretical presentation of the English language speaking skill. We include an overview of the major components of speaking, its conceptualizations, and its importance in the foreign language learning domain. Additionally, we highlight some of the difficulties that EFL learners encounter while attempting to master this skill. A detailed theoretical background of the CLT approach, including its principles, objectives, and techniques, as well as the various is presented.

1.14. Language Speaking Skill

The changes that have occurred in the twenty-first century has resulted in a gradual shift in various domains, including technology, education, and linguistics. As a consequence of these changes, English has emerged as the most widely used language in the world. It has been recognized as the international language or 'Lingua Franca,' and is deemed necessary to engage in global communication.

For the majority of students, learning English as a second language (L2) or foreign language (FL), with a focus on developing all four skills, particularly speaking, has become their main objective. However, the acquisition of speaking skills is not an effort less task, and it has led to the emergence of various theories and approaches. Among them, Communicative Language Teaching (CLT) has been widely acknowledged as the most accepted method in English language classrooms, with speaking proficiency being prioritized as its main aim.

We aim here to shed light on the crucial key terms that will be examined in depth later on, which include the definition, components, challenges, and importance of the speaking skill, as well as the definition, goal, and characteristics of CLT.

1.14.1. Background and Conceptualization of Language Speaking Skill

In the process of learning English as a second language (L2) or foreign language (FL), learners have to develop four fundamental skills, namely, writing, reading, listening, and speaking. Among these skills, speaking is considered the most crucial as it enables individuals to transfer information and maintain social relationships effectively. However, achieving proficiency in this skill is challenging due to the complexities of expressing spoken and non-verbal cues, requiring extensive experience and practice. Despite the difficulties, the mastery of English speaking skills has become a primary goal for both teachers and students and

deserves greater attention in the teaching and learning process

fundamental skill among the four skills that individuals must acquire when learning English as a foreign language (EFL). Numerous English language experts have examined it in various ways, but it is not feasible to discuss all of them.

According to Longman's 'Dictionary of Contemporary English,' Speaking is derived from the word 'speak,' which refers to the act of uttering words (Procter, 2004). Brudden (1995) delved further into this definition by stating that speaking is an activity performed by individuals to communicate with others, express ideas, feelings, and opinions, as well as accomplish specific goals. In other words, speaking is a complex verbal process of human communication that occurs anywhere and at any time, with the aim of conveying different objectives to the listener.

In the realm of EFL and According to Harmer (2002) teaching and learning, speaking is considered a productive skill that necessitates students' active production of language. Additionally, students generally have specific objectives in mind when seeking to enhance their speaking proficiency. These objectives can be categorized as either transactional purposes, such as conveying information and facilitating the exchange of goods and services, or interpersonal purposes, such as maintaining and establishing positive relationships between individuals (Harmer, 2002).

Traditionally, teachers have believed that grammar and vocabulary are the most critical components of language teaching, and thus, they spend most of the class time teaching them. Despite the fact that it has been acknowledged for many years that communication is the primary objective of language teaching and learning, and that allowing students to speak in

class is the most effective approach to achieving that objective, students are often not provided with sufficient opportunities to practice and enhance their speaking skills in the classroom. To address this issue, it is crucial to give students ample opportunities

To speak in class as it is the most effective way to achieve the goal of promoting communication skills (Harmer, 2002).

In modern times, teachers acknowledge that producing language output that can be understood necessitates an aware speaker who can correctly use grammar, pronunciation, and vocabulary (Ur, 2012). As a result, they place a greater emphasis on enhancing students' speaking abilities by utilizing various forms of speaking. According to Blaz (1999), these forms include conversation/discussion, circumlocution (description), memorized speech, oral reports, or interviews.

In conclusion, speaking is an indispensable language skill, and it is a natural means of human communication which is, used to convey ideas, emotions, and viewpoints to achieve a specific objective.

1.14.2. The Components of Language Speaking Proficiency

Richards (2008) states that “the mastery of speaking skills in English is a priority for many second language or foreign language learners” (p.19). Therefore, to attain a high level of speaking proficiency, both teachers and learners must consider language-building sub-skills, such as the components of speaking. Harris (1974) contends that there are five components of speaking that must be taken into account to evaluate oral performance, namely: pronunciation, grammar, vocabulary, fluency, and comprehension.

1.14.2.1. Grammar

Savage et al. (2010) argue that when speaking, individuals often integrate language norms into their discourse, although these norms are more strictly adhered to in writing. Beginning with basic sounds, language learners develop their oral skills by constructing words, phrases, and sentences. To produce grammatically correct sentences, speakers must be aware of various linguistic components, such as grammar, and possess knowledge of language mechanics. Grammar is a set of principles that govern the arrangement and structure of linguistic items, facilitating the relationship between form and meaning. There are two types of grammar; descriptive grammar, which describes real language usage, and prescriptive grammar, which outlines rules for proper usage. Although grammar may seem to involve merely memorizing patterns and verb forms, it should be taught as an enabling, motivating, and self-sufficient skill (Savage et al., 2010). Indeed, mastering grammar is crucial to develop speaking proficiency and enhancing communication between speakers and listeners. Misuse of grammar rules can result in an incomprehensible message, but the ability to correct such errors underscores the self-sufficiency of grammar as a skill.

1.14.2.2. Vocabulary

In the process of acquiring a second or foreign language, learners must also acquire its vocabulary, which involves learning words and their meanings. Vocabulary is considered a crucial goal in language teaching and learning as it forms the foundation of speech. Through practice, speakers can enhance their vocabulary, which in turn contributes to their ability to communicate effectively with others. As Alqahtani (2015) notes, “vocabulary is the total number of words that are needed to communicate ideas and express the speakers’ meaning” (p. 25). Essentially, it refers to the terminology that is intrinsic to any language and enables

speakers to express their thoughts and ideas.

However, one challenge faced by speakers is selecting the appropriate words to use in communication. In some cases, synonyms do not carry the same meaning in all contexts, and using them incorrectly can distort the intended message. Therefore, speakers need to use words and expressions accurately. As Harmer (2002) points out, knowledge of word classes allows speakers to form well-structured utterances. In summary, acquiring an extensive vocabulary and utilizing it appropriately is crucial for developing proficiency in the language speaking skill.

1.14.2.3. Pronunciation

Due to its pronunciation, the English language was regarded as a challenging language to acquire for many years. According to Mene (2019), pronunciation is a crucial factor to consider when talking about English proficiency. Learners must practice pronunciation overall, in order to enhance their English speaking skills.

Thornbury (2005) described pronunciation as the ability of learners to produce understandable utterances that meet the requirements of the task. In other words, learners must be aware of the various sounds and features of English, such as word stress, rising and falling intonation, and phonetic aspects. Harmer (2002) identified several sound components, including pitch, intonation, minimal pairings, spelling rhythm, and stress that are deeply connected to pronunciation. The speaker must understand the function of these components to transmit meaning effectively, resulting in less stressful and more engaging discourse.

Therefore, pronunciation entails producing words in a manner that is commonly accepted or understood based on various aspects of sounds.

1.14.2.4. Fluency

For many non-native learners, becoming a native-like English speaker is a desirable goal, but it requires significant time and practice to speak English fluently. According to Zhang (2009), fluency is characterized by speaking at a regular rate without hesitation and conveying a message in a clear and understandable manner. In other words, fluency is the ability to speak confidently and effortlessly.

Fluency has become an important sub-skill that teachers must focus on. The teacher should encourage learners to speak freely using their own language, even if they don't have enough language knowledge. Furthermore, the focus should be on the meaning and context, and less emphasis should be placed on grammatical mistakes. The teacher's support should be relevant to the classroom setting so that learners can become fluent speakers in real-world situations (Vorholt, 2018).

Furthermore, Fillmore (1979) identified five categories for assessing the fluency of student speech. These include the number of words used, the amount of time spent in silence and reflection, the number of instances of word, phrase, or clause reformulation, and the average length of speech.

In conclusion, one important skill that foreign learners must master is the ability to speak fluently in every day or formal situations.

1.14.2.5. Comprehension

In the realm of language learning, four fundamental skills exist, each with unique characteristics yet inter dependent upon one another. Listening, as one of these skills, carries a crucial role in attaining proficiency in the target language for non-native speakers. During

conversations, listeners should be attentive to the speaker and strive to understand the message to communicate it effectively (Omari, 2016). Essentially, if the listener grasps a coherent, precise, and comprehensive message, comprehension will ensue. Thus, the ability to comprehend spoken language represents an art that contributes to a learner's ability to speak the language fluently.

1.14.3. The Significance of Enhancing L2 Speaking Skill

Speaking is the most crucial aspect of language teaching and learning, according to the perspective of many language experts (Richards, 2008). This view is motivated by several reasons. First, speaking ability is often used as a yardstick to evaluate the English proficiency of learners, with their competence, skill level, and shortcomings all being assessed through their spoken output (Celce-Murcia, 2001). Second, effective communication in both classroom and real-world settings can be facilitated through speaking competency (Celce-Murcia, 2001). Lastly, proficiency in a language is frequently equated with speaking proficiency, as highlighted by Nunan's (1991) assertion that “To most important factors of learning a second or foreign language and success is measured in terms of the ability to carry out a conversation in the language” (p.39).

1.15. Communicative Language Teaching CLT

The development of communicative competence through the four language skills, especially speaking, is a primary objective in teaching and learning a foreign language. Communicative Language Teaching (CLT) is an approach that emerged to support this goal. It involves various advancements in syllabus design and has proven effective in helping students improve their ability to communicate effectively.

1.15.1. Background Overview of the CLT Approach

In recent years, the Communicative Language Teaching (CLT) approach has gained significant attention as a method for teaching foreign languages, particularly English. It originated in Europe and the USA in the late 1960s as a response to dissatisfaction with traditional language teaching methods such as the Audio-lingual and Grammar Translation method. Scholars such as Hymes, Chomsky, Wilkins, Van Ek, Alexander, and the Council of Europe contributed to the development of CLT, arguing that previous methods failed to provide learners with the ability to communicate effectively in real-life situations due to an emphasis on structure rather than function and notion of language (Spada, 2007).

Howatt (1984) classified the CLT approach into weak and strong versions. The weak version highlights the importance of providing learners with opportunities to use English for communication purposes (i.e. learning to use English), while the strong version asserts that language is acquired through communication, emphasizing the use of English to learn communication, and it is not merely a matter of activating preexisting language knowledge, but rather of stimulating the development of the language system itself (i.e. the use of English as a means to acquire proficiency in it).

1.15.2. The Main Objective of the CLT Approach

Various approaches and methods have emerged over the years. However, most of them failed to recognize the importance of communication in the teaching process. The Communicative Language Teaching (CLT) approach stands out as a response to this issue by considering language as a means of communication. CLT has gained widespread acceptance and use because it aims to develop communicative competence, with a focus on the development of the four language skills, particularly speaking.

Hymes, (1972, as cited in Richards & Rodgers, 2001), asserted that the ultimate aim of language teaching is to cultivate what is referred to as 'communicative competence.' According to his perspective, individuals are not competent merely because they can express their ideas, but because they possess both the knowledge and skill to use language effectively. Canale (1980) further elaborated on Hymes' theory by proposing that communicative competencies made up of four components, namely grammatical competence, sociolinguistic competence, strategic competence, and discourse competence. Together, these components provide a comprehensive definition of what a learner requires to be communicatively competent in a given speech community.

1.15.3. Principles Underlying the Communicative Language Teaching Approach (CLT)

The emergence of the CLT approach was a response to the inadequacies of traditional language teaching methods. It emphasized the importance of language teaching by highlighting four characteristics of the language (Shaw & McDonough, 1993);

- Language serves as a system for expressing meaning.
- Interaction and communication are the primary functions of language.
- The structure of language reflects its functional and communicative uses.
- The primary unit of language is not only its grammatical and structural features but also categories of functional and communicative meaning as exemplified in discourse.

According to Littlewood (1981, as cited in Richards & Rodgers, 1986), "One of the most characteristic features of communicative language teaching is it pays systematic attention to functional as well as structural aspects of language" (p.66).

The principles of Communicative Language Teaching (CLT) are centered on language teaching goals, classroom activities that promote learning, and the respective roles of teachers

and learners in the classroom. Aromdee (2012) notes that CLT is characterized by the integration of different language skills, enabling students to communicate and interact meaningfully using authentic materials. CLT emphasizes the use of communicative activities that are completed through pair and group work, such as role-plays, games, information-gap activities, and interviews, which provide opportunities for students to engage in authentic communication.

CLT also emphasizes the roles of teachers and learners in the classroom. Teachers are encouraged to facilitate communication between students by providing appropriate pair or group activities and tasks. They are also expected to act as independent participants within the learning-teaching group by sharing their knowledge and tolerating errors until activities are completed. In turn, students are expected to interact with their classmates in pair or group work.

1.15.4. CLT as a Basis to Develop EFL Speaking Skills

As previously mentioned, speaking is a fundamental means of communication between individuals. However, traditional language teaching methods failed to prioritize this critical skill, leading to the development of CLT. CLT emphasizes the importance of communication in language learning and aims to improve students' ability to communicate effectively in real-life situations outside of the classroom.

In the field of teaching EFL, CLT prioritizes the development of students' speaking skills, believing that communication is the primary means of acquiring a foreign language. Saputra (2015) confirms that CLT is the most effective method for teaching English speaking. CLT encourages active communication among students in the classroom, with both teachers and learners practicing and improving their speaking skills.

Harmer (2002) notes that in CLT, activities focus on real or realistic communication, as it simulates real-life communication situations. He adds that the primary goal in CLT is for learners to successfully accomplish communicative tasks. This means that while accuracy in language use is still important, it is not the sole or overriding concern. Learners are encouraged to use the language they know to convey their message and achieve their communicative objectives, even if they make some errors along the way. Therefore, students should observe their classmates' actions to build confidence in their ability to communicate. Teachers play a crucial role in facilitating successful oral communication by motivating, assessing, facilitating, and correcting students while providing appropriate classroom activities to encourage real-life language use.

1.16. Types of CLT Speaking Activities

In the realm of language learning and teaching, it is essential to focus on communication as the ultimate goal. By doing so, students will be better equipped to use the language effectively in real-life situations that are relevant to their own interests, needs, and opportunities. To achieve this, the implementation of CLT in teaching speaking skills requires teachers to provide classroom activities that resemble real-life scenarios, making it easy for students to apply what they have learned outside of the classroom.

In order to create a context that fosters effective communication, teachers must organize a wide range of classroom activities that keep students motivated and engaged in speaking activities. In the third edition of 'Teaching English as a Second or Foreign Language', Celce-Murcia (2001) classifies speaking activities into four types:

- **Structured activities:** These activities are employed to achieve the goal of communication between students in and out of the classroom. According to Celce-

Murcia, EFL classroom activities are classified based on a continuum that ranges from manipulative activities, which often disregard the meaning of the task, to communicative activities that depend on authentic situations. Manipulative activities, such as drills, provide prepackaged material that enables students to create genuine, real, and captivating communication.

- **Performance activities:** These activities occur after the students have prepared themselves. The students are aware of the topic that will be discussed, and their thoughts and ideas are already equipped in their minds during the speech. This preparation makes them feel self-confident and contributes to the success of the process of delivering the message to their classmates in a comprehensible way. For instance, explaining a process, telling a story from daily life experiences, or engaging in role-plays, dramas, and debates all serve as performance activities that create a communicative atmosphere in the class and allow for the participation of all students. As a result, the learners' level of speaking improves, as does their knowledge.
- **Participation activities:** These activities are also commonly used in EFL classrooms. They take place in a natural setting in which the teacher guides students to discuss a specific topic or problem in small groups. The teacher introduces the topic through a brief reading, and the students then suggest solutions, resolutions, or complications.
- **Observation activities:** These activities provide students with a sense of comfort, as they are not required to participate or interact. In these tasks, students simply observe or record verbal and non-verbal interactions between native or fluent speakers, without the fear of making errors in performance.

1.17. Teaching Techniques and Activities in CLT

CLT offers a variety of techniques to improve EFL learners' speaking skills. The teacher must possess knowledge of these techniques to select the appropriate ones that facilitate the expression of learners' opinions, emotions, and information gathering, thereby preparing them for communication beyond the classroom.

According to Larsen-Freeman (2000), in his book, several effective techniques exist in CLT to enhance students' speaking skills. These include role-play activities and the utilization of authentic materials. Oradee (2013) suggested that communicative techniques like problem-solving and information gap activities can also aid in developing learners' speaking abilities. Hedge (2000) supported the use of free discussion and role-playing to enhance students' oral production. The following subsections will elaborate on each of these techniques, which will be employed in the current study.

1.17.1. Role-Playing Activities

Role-playing activities enable learners to practice communication in diverse social contexts and roles, making it the most enjoyable for many students, especially those who like to imitate others. Role-playing is also an opportunity for students who are afraid of public speaking to participate without pressure. Harmer (2022) describes it as an authentic teaching and learning strategy, in which student-student interaction is vital for developing motivation and engagement in the activity. When using role-play activities, the teacher must choose a suitable topic that meets the learners' needs and interests to ensure their participation without pressure. Students can work in pairs, small groups, or large groups, and each student must play a specific role. The success of the activity is determined by the participation of all students, even the weaker ones, as everyone can select a role that matches their linguistic abilities. For

instance, asking students to play various characters from a narrative story is an easy, fun way to memorize language phrases. In other words, while performing the role-play of a story, students participate in language learning and develop their speaking skills (Harmer, 2022).

1.17.2. Problem-Solving Activities

Problem-solving activities can be an engaging technique to enhance EFL learners' speaking skills. According to Richards and Rodgers (2001), problem-solving activities involve presenting a problem or an issue to students along with some relevant information and then asking them to devise an appropriate solution. The language used during these activities depends on the nature of the problem and the topic under discussion. In addition, Klippel (1984) suggests that students can be encouraged to make suggestions, provide reasons, and even challenge or accept the suggestions and reasons provided by their peers.

These activities provide students with the opportunity to exercise their critical thinking skills and apply their language knowledge to real-life situations.

1.17.3. Interviewing Activities

In the communicative approach, interviews can be an effective way to improve students' speaking skills by enhancing their fluency. This is achieved when students choose topics that are of interest to them and then interview various individuals. Conducting interviews in this way provides students with the opportunity to practice their speaking skills and makes them more sociable, as noted by Kayi (2008). For the interview to be successful, both the interviewer and interviewee must be good listeners, enabling the question-and answer sequence to develop into a meaningful conversation. As suggested by Klippel (1984).The teacher may provide students with a rubric to guide them on the types of questions to ask and the sequence to follow in the interview. However, it is essential for the students to prepare the

interview questions on their own, which helps them to take ownership of their learning process (Kayi, 2008).

1.17.4. Using Authentic Materials

Developing language speaking skills among non-native EFL learners is a challenging task, as they may not have ample opportunities to practice the language in authentic situations (Oura, 2001). To address this issue, the Communicative Language Teaching (CLT) approach emphasizes the use of authentic materials that expose students to language as it is actually used in real-life contexts. The incorporation of authentic materials that are not specifically designed for language teaching purposes has been found to be effective in enhancing various language skills, mainly the speaking skill (Nunan, 1989). The selection of appropriate materials is crucial and must take into account the learners' background, the purpose of instruction, and the best way to use the materials (Ellis & Johnson, 1994).

1.18. Assessment of Oral Proficiency

The complex nature of the construct of oral proficiency implies that teachers must have a clear understanding of what speaking is in order to guarantee reliable assessment protocols (Luoma, 2004). Most importantly, they must consider the special nature of spoken grammar and spoken vocabulary. Speech is less sophisticated than written language in terms of vocabulary and grammar, and it is organized in idea units or phrasal chains instead of complete sentences (Luoma, 2004; Milton, 2013). Moreover, teachers must take into account that speech is based on interaction, which implies that speakers react to each other and take turns. The implications of these ideas in terms of assessment include that teachers should not require students to use long phrases and advanced vocabulary in speaking tests, and that teachers should reward the use of

interaction strategies (Luoma, 2004).

The act of speaking happens in real time and it involves many different features such as pronunciation, pausing, rhythm and monitoring interaction, as well as resources such as vocabulary and grammar. Therefore, speaking is often regarded as the most difficult skill to assess reliably among the traditional four skills (reading, writing, listening and speaking) (Lazaraton, 2014; Luoma, 2004; Rydell, 2014). The use of analytic assessment protocols (i.e., based on different aspects or categories) instead of holistic (i.e., based on a global judgment) is recommended by the CEFR (Council of Europe 2001), as analytic assessment protocols encourage a closer observation and minimize the risk of basing assessment on one or two language features (Council of Europe, 2001; Rydell, 2014). In order to make reliable evaluations, it is important to describe the assessment categories as clearly as possible (Brown, 2004). For example, Ellis provides a list of six different criteria that may be used in analytic assessments: pronunciation, fluency, vocabulary, grammar, discourse features such as cohesion and appropriateness, and accomplishing the objective of the task. In any case, EFL teachers must create numerous and varied opportunities for output, especially in the form of interaction, to provide a solid base for assessment (Ellis, 2014; Rydell, 2014).

Conclusion

Throughout this chapter, it has been outlined and revealed that the Communicative Language Teaching (CLT) approach has been conceived as a successful teaching methodology that has a high potential to enhance EFL learners speaking skills. The effectiveness of this approach has been established through an exploration of its underlying principles, techniques, and classroom activities. All in all, this chapter has sought to administer a prevailing theoretical understanding of the CLT approach and Language Speaking. The subsequent chapter will detail the methodology employed in the present study, followed by a comprehensive analysis of the main research findings, in order to test the hypotheses proposed in this work.

Chapter Two:

Methodology and Data Analysis

2.1. Research Methodology and Data Collection

2.1. 1. Research Design

This study adopts a correlational research design to investigate the relationship between subjective assessments (teachers' ratings) and objective measurements of English oral proficiency among third-year students at Msila University's English Department. The approach allows for the analysis of the degree to which the subjective ratings of oral proficiency align with objective measures.

2.1. 2. Participants

The participants in this study were third-year students enrolled in the English Department at Msila University. A total of 6 students were randomly selected from among 60 others. Additionally, five experienced university teachers participated in rating the students' oral proficiency.

2.1.3. Data Collection

- **Recording of Oral Performances**

Each student was asked to perform a standardized oral task about the same topic of daily routine, which was recorded for subsequent analysis. The task involved a spontaneous speaking activity to narrate someone daily activity.

- **Subjective Assessment**

The recorded oral performances were rated by the five university teachers. Each teacher independently rated the students on four dimensions: Oral Proficiency Accuracy, Fluency, and Pronunciation. A rating scale of 1 to 20 has been designed and administered to 8 teachers at the University of Msila along with less than a 5 minutes recording for each student in which students narrate their daily routine activities. Teachers listened to and rated each student for **overall oral**

proficiency first then separately for **fluency**, **pronunciation**, and **accuracy** criteria. Five teachers out of eight responded and completed the rating task.

- **Objective Measurement**

The recordings were analyzed using specialized software (Praat) to obtain objective measurements of the students' oral performances. The analysis focused on: speech rate (words per minute), error frequency (number of grammatical errors per 100 words), pause duration and frequency (mean length and frequency of pauses), phonological accuracy (percentage of correctly pronounced phonemes)

2.1.4. Data Analysis

- **Correlational Analysis**

To investigate the relationship between the subjective and objective assessments, the following steps were undertaken:

- **Subjective Ratings Correlation:** The teachers' ratings for oral proficiency were correlated with their ratings for accuracy, fluency, and pronunciation using Pearson's correlation coefficient. This analysis aimed to determine the consistency and relationship among different dimensions of oral performance as perceived by the teachers.
- **Objective vs. Subjective Correlation:** The objective measurements of speech rate, error frequency, pause duration, and phonological accuracy were correlated with the teachers' subjective ratings of oral proficiency. Pearson's correlation coefficient was used to assess the strength and direction of these relationships.

The correlational analyses were used to validate whether the subjective ratings and objective measures were indeed assessing the same underlying constructs of oral proficiency.

2.1.5. Instruments

The instruments used in this study comprised PRAAT speech analysis software and teachers' evaluation. When test-takers took the test PRAAT and Transcriber are computer-based software used for speech analysis, which are useful for distinguishing silent pauses from phonations. By providing oscillography pictures, they separated silent pauses from phonations. The oscillography pictures were used to measure the lengths of pauses.

2.2. Data Analysis, Findings, Discussion and Recommendations

The following section presents teachers' evaluations of each student. S1 in the table refers to student one, S2 to student 2 and the way is so until S6.

Analysis of Teachers' Evaluations of Students' Voice Recordings

Analysis of Teacher1 Evaluation

Table 2: *Teacher 1 Evaluation of Students' Voice Recordings*

Criteria	Students					
	S1	S2	S3	S4	S5	S6
Oral Proficiency	13	15	13	13.5	13	12
Pronunciation	10	14.5	10	11	10.5	10
Accuracy	10	14	10	11	10.5	10
Fluency	11	13	11	12	11	10

Table 1 shows the evaluations by Teacher 1 of six students (S1 to S6) on four criteria: overall oral proficiency, pronunciation, accuracy, and fluency. Each criterion is scored on a 20 points scale. Overall oral proficiency scores range from 12 to 15, indicating a moderate level of variability among the students.

Correlating Overall Oral Proficiency with Pronunciation, Accuracy, and Fluency

The data show s that S2, who scored highest in overall oral proficiency, also has the highest scores in pronunciation, accuracy, and fluency, indicating a strong positive correlation. This suggests that Teacher 1 views these components as integral to overall proficiency. The table also reveals that lower overall proficiency scores correlate with lower scores in pronunciation, accuracy, and fluency, reinforcing the idea that these aspects are critical to their overall speaking ability. It is also noticed that **pronunciation and accuracy** have identical scores for each student, indicating that Teacher 1 evaluates pronunciation and accuracy very similarly or sees them as closely linked. As for fluency scores, while not identical to pronunciation and accuracy, show similar trends. It means that students with higher fluency scores tend to have higher overall proficiency scores. Finally, the results also indicate that teacher 1's evaluations are highly consistent. Overall, the results indicate that teacher one intuitive overall oral proficiency scores are influenced by pronunciation, accuracy, and fluency. This is statistically evident by calculating correlation coefficients for teacher 1 evaluations, correlation between Oral Proficiency and Pronunciation: ($r = 0.913$), Correlation between Oral Proficiency and Accuracy: ($r = 0.898$) Correlation between Oral Proficiency and Fluency: ($r = 0.969$). Fluency: ($r = 0.969$) has the strongest correlation with oral proficiency, followed by Pronunciation: ($r = 0.913$), and then Accuracy: ($r = 0.898$). This implies that, for Teacher 2, a student's fluency in language use is slightly more indicative of their overall oral proficiency compared to pronunciation and accuracy.

These correlation coefficients suggest a strong positive linear relationship between oral proficiency and the other criteria in teacher 1 evaluations. This implies that teachers tend

intuitively judge oral proficiency depending on the three constructs of accuracy, fluency, and pronunciation.

Analysis of Teacher 2 Evaluation

Table 3: *Teacher 2 Evaluation of Students' Voice Recordings*

Criteria	Students					
	S1	S2	S3	S4	S5	S6
Oral Proficiency	12	13	11	12	11	9
Pronunciation	12	12	10	11	9	7
Accuracy	11	13	12	11	11	10
Fluency	11	12	10	11	10	9

Table 2 presents the evaluations by Teacher 2 of the same six students (S1 to S6) on four criteria: overall oral proficiency, pronunciation, accuracy, and fluency. **Overall oral proficiency** scores range from 9 to 13, indicating a moderate level of variability among the students. S2 scored the highest (13) and S6 scored the lowest (9). **Pronunciation** scores range from 7 to 12, with S1 and S2 having the highest score (12) and S6 has the lowest score (7). **Accuracy** scores range from 10 to 13, showing that S2 is perceived to have the highest grammatical accuracy (13) and S1 and S6 scored the lowest (10), which may imply frequent grammatical errors. As for **fluency** scores, they range from 9 to 12, with S2 scoring highest (12), indicating better fluidity and ease of speaking and S6 has the lowest fluency score (9), suggesting possible issues with speaking smoothly or at a consistent pace.

Correlating Overall Oral Proficiency with Pronunciation, Accuracy, and Fluency

The data show that S2, who scored highest in overall oral proficiency, also has high

scores in pronunciation, accuracy, and fluency, indicating a strong positive correlation. Table 2 also shows that lower overall proficiency scores for other students consistently correlate with scores in pronunciation, accuracy, and fluency, reinforcing the idea that these aspects are critical to their overall speaking ability overall intuitive judgments.

Our calculation of correlation coefficients for teacher 2 evaluations revealed the following results; correlation between Oral Proficiency and Pronunciation: ($r = 0.492$), Correlation between Oral Proficiency and Accuracy: ($r = 0.502$), Correlation between Oral Proficiency and Fluency: ($r = 0.463$). These correlation coefficients suggest moderate correlations, suggesting a positive but not strong linear relationship between oral proficiency and the other criteria (pronunciation, accuracy, and fluency). Accuracy has the strongest correlation with oral proficiency, followed closely by Pronunciation ($r = 0.492$), and then Fluency ($r = 0.463$). This implies that, for Teacher 2, a student's accuracy in language use is slightly more indicative of their overall oral proficiency compared to pronunciation and fluency.

Comparative Analysis of Teacher 2 with Teacher 1's Evaluation

It appears that Teacher 2's scores are slightly lower on average compared to Teacher 1's scores. This might indicate a stricter evaluation approach or a different interpretation of the criteria. S2 is found to consistently score highest in both evaluations, indicating a consensus on this student's superior performance across the criteria. However, it appears that the range of scores for each criterion is broader in Teacher 2's evaluation, particularly in pronunciation, suggesting a more differentiated assessment of student abilities. This moderate correlation coefficients suggest also that Teacher 2 might have a more varied or nuanced approach to evaluating oral proficiency. The relationships between the different aspects of oral proficiency (pronunciation, accuracy, and fluency) are not as tightly coupled as observed in teacher 1

evaluations. This could indicate that Teacher 2 values or weights these criteria differently or that there might be other factors influencing their assessment of oral proficiency. The differences in evaluation by the two teachers, teacher 1 and teacher 2 suggest the need for standardized evaluation criteria and inter-rater reliability training to ensure consistency in assessments.

Analysis of Teacher 3 Evaluation

Table 4: *Teacher 3 Evaluation of Students' Voice Recordings*

Criteria	Students					
	S1	S2	S3	S4	S5	S6
Oral Proficiency	13	16	16	15	16	9
Pronunciation	11	16	14	15	13	9
Accuracy	11	16	13	15	13	9
Fluency	13	16	14	15	14	9

Table 3 above displays the evaluations by Teacher 3 of the six students (S1 to S6) on the four criteria: overall oral proficiency, pronunciation, accuracy, and fluency. The scores range from 9 to 16, showing a significant variation among the students. S2, S3, and S5 scored the highest (16), indicating excellent overall speaking abilities according to Teacher 3 while S6 scored the lowest (9), suggesting a need for improvement in overall oral proficiency. **Pronunciation** scores range from 9 to 16, with S2 having the highest score (16) and S6 scoring the lowest (9), indicating challenges in clear and correct pronunciation. Accuracy scores range from 9 to 16, with S2 scoring highest (16), suggesting high grammatical accuracy and S6 scored the lowest (9), which implies frequent grammatical errors. **Fluency** scores range from 9 to 16, with S2 scoring highest (16), indicating smooth speech and S6 scoring (9), suggesting issues with speaking smoothness or consistent pace.

Correlating Overall Oral Proficiency with Pronunciation, Accuracy, and Fluency

Reading the data from table three shows that S2's consistently high scores across all criteria. This suggests a strong positive correlation among the different aspects of speaking skills (pronunciation, accuracy, and fluency). Lower scores in overall proficiency are found to correspond to lower scores in pronunciation, accuracy, and fluency, indicating that these aspects are critical for overall speaking ability in teachers' intuitive judgments.

The statistical calculations of correlation coefficients for teacher 3 evaluations indicated the following; correlation between Oral Proficiency and Pronunciation: ($r = 0.913$), Correlation between Oral Proficiency and Accuracy: ($r = 0.936$), and Correlation between Oral Proficiency and Fluency: ($r = 0.960$). This implies that, for Teacher 3, a student's fluency in language use is slightly more indicative of their overall oral proficiency followed by accuracy then pronunciation.

These correlation coefficients suggest a strong positive linear relationship between oral proficiency and the other criteria in teacher 1 evaluations. This implies that teachers tend intuitively judge oral proficiency depending on the three constructs of accuracy, fluency, and pronunciation.

Comparative Analysis of Teacher 3 with Teacher 1 and Teacher 2's Evaluations

Our comparison showed that S2 consistently scores high across all teachers' evaluations, indicating agreement on this student's superior performance. S6 consistently scores lowest across evaluations, suggesting a clear area for improvement. Regarding the **variability in teachers' scores**, the data shows that Teacher 3's scores are slightly higher on average compared to Teachers 1 and 2's scores, indicating a possibly more lenient evaluation approach. The high

variability in Teacher 3’s scores, especially for overall oral proficiency, shows a more exact assessment of student abilities.

Analysis of Teacher 4 Evaluation

Table 5: Teacher 4 Evaluation of Students’ Voice Recordings

Criteria	Students					
	S1	S2	S3	S4	S5	S6
Oral Proficiency	9	15	16	14	11	11
Pronunciation	5	14	15	14	10	10
Accuracy	10	16	14	13	12	12
Fluency	8	16	15	14	11	11

Table 4 presents Teacher 4's evaluations of the six students' voice recordings across various criteria: Oral Proficiency, Pronunciation, Accuracy, and Fluency. Teacher 4 rated student 3 (S3) as the highly proficient in oral proficiency scoring 16 out of 20, Student 1 (S1) has the lowest score of 9 in oral proficiency. In pronunciation, student 3 (S3) also has the highest pronunciation score of 15 and S6 has the lowest score of 5. In accuracy, S2 and S3 have the highest accuracy score of 16 and S5 has the lowest score of 10. In fluency, S2 has the highest fluency score of 16 and S1 and S6 again have the lowest fluency scores of 8.

Correlating Overall Oral Proficiency with Pronunciation, Accuracy, and Fluency

There appears to be a positive correlation between Oral Proficiency and the other criteria (Pronunciation, Accuracy, and Fluency), as higher scores in Oral Proficiency generally coincide with higher scores in these areas. The more detailed statistical analysis to come will better capture the strength of correlation.

To better capture the correlations of oral proficiency with the other criteria (fluency, accuracy, and pronunciation) statistically, we calculated Pearson correlation coefficients and found that for teacher 4 evaluations, correlation between Oral Proficiency and Pronunciation: ($r = 0.810$), Correlation between Oral Proficiency and Accuracy: ($r = 0.976$) Correlation between Oral Proficiency and Fluency: ($r = 0.869$). These correlation coefficients suggest a strong positive linear relationship between oral proficiency and the other criteria in teacher 4 evaluations. This implies that students who perform well in oral proficiency tend to also excel in accuracy, fluency, and pronunciation.

Analysis of Teacher 5 Evaluation

Table 6: *Teacher 5 Evaluation of Students' Voice Recordings*

Criteria	Students					
	S1	S2	S3	S4	S5	S6
Oral Proficiency	13	15	13	13.5	13	12
Pronunciation	10	14.5	10	11	10.5	10
Accuracy	10	14	10	11	10.5	10
Fluency	11	13	11	12	11	10

The scores of Teacher 5's evaluation of oral proficiency range from 12 to 15; with S2 having the highest score of 15 and S6 is having the lowest score of 12. In pronunciation, scores range from 10 to 14.5, with S2 having the highest 14.5 score and S1 having the lowest score of 10. In accuracy, scores range from 10 to 14, with S2 again having the highest accuracy score of 14 and S1 is having the lowest score of 10. In fluency, scores range from 10 to 13, with S4 having the highest fluency scores of 13 and S6 the lowest score of 10.

Careful reading of the data showed us a positive correlation between Oral Proficiency and the other criteria (Pronunciation, Accuracy, and Fluency) according to the evaluations of Teacher 5. Here are correlation coefficients calculated: correlation between Oral Proficiency and Pronunciation: ($r = 0.992$), Correlation between Oral Proficiency and Accuracy: ($r = 0.991$), and Correlation between Oral Proficiency and Fluency: ($r = 0.986$). Once again, these correlation coefficients suggest a strong positive linear relationship between oral proficiency and the other criteria in teacher 5 evaluations.

Comparative Analysis of Teacher 4 with Teacher 5 Evaluations

Teacher 4's evaluations show higher variability across all criteria compared to Teacher 5's evaluations, indicating potentially different grading standards or perceptions of student performance between the two teachers. Additionally, Teacher 5's evaluations show slightly higher correlation coefficients compared to Teacher 4's evaluations, indicating a potentially more consistent assessment approach regarding the relationship between oral proficiency and the other evaluated criteria.

Strength of Correlation with Oral Proficiency

Based on the correlation coefficients provided by each teacher's evaluations, we can determine the overall strength of the relationship between oral proficiency and the other criteria (fluency, accuracy, and pronunciation). These criteria will be first ranked according to their average correlation strength with oral proficiency across all teachers. The table below provides a concise summary:

Table 5 A : *The correlation coefficients of the 3 criteria across the 5 teachers*

Teacher	Oral Proficiency & Pronunciation	Oral Proficiency & Accuracy	Oral Proficiency & Fluency
T 1	0.913	0.898	0.969
T2	0.492	0.502	0.463
T3	0.970	0.936	0.960
T4	0.810	0.976	0.869
T5	0.992	0.991	0.986
Average	0.8354	0.8606	0.8494

Based on the average correlation coefficients, the criteria can be ordered according to their strength of correlation with oral proficiency as follows:

- **Accuracy** has been found to have the strongest correlation with oral proficiency, suggesting that students' accuracy in language use is most closely associated with their overall oral proficiency, which influences teachers' overall intuitive judgments of students' oral proficiency. This latter is influenced by the accuracy construct in the first place.
- **Fluency** follows closely, indicating that the smoothness and flow of speech are also strongly related to teachers' oral proficiency intuitive judgments.
- **Pronunciation** has been shown to have a slightly weaker correlation compared to accuracy and fluency but still shows a significant relationship with teachers' oral proficiency intuitive judgments. This ordering highlights that, while all three criteria are important for evaluating oral proficiency and influential in teachers' oral proficiency intuitive judgments, accuracy tends to be the most strongly associated factor, followed by fluency, and then pronunciation.

Student fluency control temporal variable

Table 7 : *Evaluation of Student 1 Voice Recording*

Criteria	Students					
	T1	T2	T3	T4	T5	Av
Oral Proficiency	13	12	13	9	13	12
Pronunciation	10	12	11	5	10	9.6
Accuracy	10	11	11	10	10	10.4
Fluency	11	11	13	8	11	10.8

Table 8: *Evaluation of Student 2 Voice Recording*

Criteria	Students					
	T1	T2	T3	T4	T5	Av
Oral Proficiency	15	13	16	15	15	14.8
Pronunciation	14.5	12	16	14	14.5	14.2
Accuracy	14	13	16	16	14	14.6
Fluency	13	12	16	16	13	14

Table 9: *Evaluation of Student 3 Voice Recording*

Criteria	Students					
	T1	T2	T3	T4	T5	Av
Oral Proficiency	13	11	16	16	13	13.8
Pronunciation	14.5	12	16	14	14.5	14.2
Accuracy	10	12	13	14	10	11.8
Fluency	11	10	14	15	11	12.2

Table 10: *Evaluation of Student 4 Voice Recording*

Criteria	Students					
	T1	T2	T3	T4	T5	Av
Oral Proficiency	13.5	12	15	14	13.5	13.5
Pronunciation	11	11	15	14	11	12.4
Accuracy	11	11	15	13	11	12.2
Fluency	12	11	15	14	12	12.8

Table 11: *Evaluation of Student 5 Voice Recording*

Criteria	Students					
	T1	T2	T3	T4	T5	Av
Oral Proficiency	13	11	16	11	13	12.8
Pronunciation	10.5	9	13	10	10.5	10.6
Accuracy	10.5	11	13	12	10.5	11.4
Fluency	11	10	14	11	11	11.4

Table 12: *Evaluation of Student 6 Voice Recording*

Criteria	Students					
	T1	T2	T3	T4	T5	Av
Oral Proficiency	12	9	9	11	12	10.6
Pronunciation	10	7	9	10	10	9.2
Accuracy	10	10	9	12	16	11.4
Fluency	10	9	9	11	10	9.8

Temporal Data from Praat Software:

Table 13: *Combination of Teachers Evaluation with Students Objective Results*

	Criteria	Students					
1		S1	S2	S3	S4	S5	S6
2	Oral Proficiency	12	14.8	13.8	13.5	12.8	10.6
3	Pronunciation	9.6	14.2	14.2	12.4	10.6	9.2
4	Accuracy	10.4	14.6	11.8	12.2	11.4	11.4
5	Fluency	10.8	14	12.2	12.8	11.4	9.8
6	T+P	20.52	194.92	194.72s	194.92	171.56	114.87
7	PT	4.98	55.64	30s	69.4	37	41
8	T-P	15.54	125.52	164.72	125.52	134.56	73.85
9	NS	55	199	41	55	449	249
10	NP	6	84	84	84	59	43
11	NR	7	85	85	85	60	44
12	MLR	7.85	2.34	0.48	7.85	7.31	5.6
12	SR	2.68	25.8	25.8	0.28	0.34	2.16
14	AR	3.53	1.58	14.4	0.79	3.33	3.37
15	LP	0.83	0.62	1.96s	1.49	2.28	0.16
16	PAM	0.14	0.035	0.14	0.10	0.011	0.008
17	PPM	0.30	0.04	0.12	0.09	0.013	0.008

To facilitate reading table 12, the following table of abbreviations is provided

1	T+P	Total time of the sample including pause time
2	PT	Total pause time in the sample
3	T-P	Total time of the sample excluding pause time
4	NS	Total number of syllables in the sample
5	NP	Total number of pauses in the sample
6	NR	Total number of fluent runs in the sample
7	MLR	Mean length of run
8	SR	Speech rate
9	AR	Articulation rate

10	LP	Mean length of pause
11	PAM	Proportion of accuracy mistakes
12	PPM	Proportion of pronunciation mistakes

Correlating Teachers’ Judgments of Oral Proficiency with their Judgments of Pronunciation
(row 2 with row 3)

The correlation coefficient between the teachers' evaluations of the students' oral proficiency and pronunciation is approximately 0.922. This indicates a **very strong positive relationship** between the two variables. This means that students who are rated highly in oral proficiency also tend to be rated highly in pronunciation, and vice versa. This high correlation suggests that the teachers' evaluations are consistent across these two criteria. It implies that the aspects the teachers consider important for oral proficiency might overlap significantly with those for pronunciation. This consistency can be useful for both students and educators. For students, it means that improving in one area is likely to be associated with improvements in overall oral proficiency. For educators, it may indicate that their evaluation criteria are aligned, which can simplify the assessment process.

Correlating Teachers’ Judgments of Oral Proficiency with their Judgments of Accuracy
(row 2 with row 4)

The correlation coefficient between the teachers' evaluations of the students' oral proficiency and accuracy is approximately 0.740. This indicates a strong positive relationship between oral proficiency and accuracy. This suggests that students who perform well in oral proficiency tend to also perform well in accuracy, and vice versa. The potential implication of these findings for students is that focusing on improving accuracy could also positively impact their overall oral proficiency. For teachers, it can be invested in designing integrated lessons that simultaneously target both accuracy and oral proficiency. Overall, the strong positive correlation of 0.740 indicates that there is a significant relationship between the evaluations of oral proficiency and accuracy, suggesting that improvements in one area are likely to correspond with improvements in the other.

Correlating Teachers' Judgments of Oral Proficiency with their Judgments of Fluency (row 2 with row 5)

The correlation coefficient between the teachers' evaluations of the students' oral proficiency and fluency is approximately 0.969. This indicates a **very strong positive relationship** between oral proficiency and fluency. This means that students who are rated highly in oral proficiency are also rated highly in fluency, and vice versa. This suggests that the teachers' evaluations for oral proficiency and fluency are highly consistent and the criteria used for assessing these two aspects are likely to overlap significantly. This result implies that working on improving fluency can have a significant positive impact on students overall oral proficiency. For teachers, it implies that teachers can address both fluency and oral proficiency together and that enhancements in one are likely to result in improvements in the other.

Correlating Teachers' Judgments of Oral Proficiency with Temporal Measurement of Accuracy through Proportion of Accuracy Mistakes (PAM) (row 2 with 16)

The correlation coefficient between teachers' intuitive judgments of students' oral proficiency in and the proportion of accuracy mistakes (PAM) made by the students is approximately 0.218. This positive correlation indicates that as the teachers' ratings of oral proficiency increase, the proportion of accuracy mistakes tends to decrease slightly. However, since the correlation is low, this relationship is weak because a correlation of 0.218 suggests a very weak linear relationship between the two variables. This means that the teachers' intuitive judgments and the accuracy mistakes made by students are not strongly associated. This weak correlation implies that other factors may be influencing teachers' judgments of oral proficiency beyond just the accuracy of the speech.

Correlating teachers' judgments of oral Proficiency with Temporal Measurement of Pronunciation through Proportion of Pronunciation Mistakes (PPM) (row 2 with row 17)

The correlation coefficient between teachers' intuitive judgments of students' oral proficiency and the proportion of pronunciation mistakes (PPM) made by the students is approximately -0.076. This negative correlation indicates that as the teachers' ratings of oral proficiency increase, the proportion of

pronunciation mistakes tends to decrease, though this relationship is very weak. This result suggests a very weak linear relationship between the two variables. This means that the teachers' intuitive judgments and the pronunciation mistakes made by students are not strongly associated, which implies that pronunciation mistakes have little to no effect on the teachers' judgments of oral proficiency. Other factors are likely playing a more significant role in these judgments. Overall, referring to our previous analysis of accuracy mistakes and oral proficiency correlation, the weak correlation with pronunciation mistakes suggests that teachers' ratings of oral proficiency are influenced by multiple factors, not just the proportion of pronunciation mistakes.

Correlating teachers' Judgments of Fluency with Temporal **Measurement** of Fluency (row 5 with rows 6 through 15)

Temporal evaluations of students' accuracy and pronunciation:

Table 13: *Number of students' mistakes in the speech recoding*

Criteria	Students					
	S1	S2	S3	S4	S5	S6
Oral Proficiency	12	9	9	11	12	10.6
Pronunciation	17	8	5	5	6	2
Accuracy	8	7	6	6	5	2
PAM	0.14	0.035	0.14	0.10	0.011	0.008
PPM	0.30	0.04	0.12	0.09	0.013	0.008

➤ **Proportion of Accuracy Mistakes(PAM)**

PAM is the proportion of accuracy mistakes per a hundred syllables. It is calculated as follows:

$$\text{Number of accuracy mistakes in the sample} \div \text{number of syllables in the sample}$$

For example: as shown in table 13 above, Student 1 (S1) committed **8** accuracy mistakes and produced **55** syllables. So, $8 \times 100 \div 55 = 14.54$ mistakes per 100 syllables

$$14.54 \div 100 = 0.14$$

➤ **Proportion of Pronunciation Mistakes (PPM)**

PPM is the proportion of **pronunciation** mistakes per a hundred syllables. It is calculated as follows:

$$\frac{\text{Number of } \mathbf{pronunciation} \text{ mistakes in the sample}}{\div \text{number of syllables } \mathbf{in the sample}}$$

Correlating Teachers' Judgments of Fluency with Sample Time including Pauses (T+P) (row 5 with rows 6)

The correlation coefficient (r) between the two variables is found to equal 0.678. The correlation coefficient (r) of approximately 0.678 indicates a moderately positive correlation between teachers' intuitive judgments of students' oral fluency and the time of speaking including pause time in students' speech samples. This means that as fluency increases, the time of speaking tends to increase as well, although not extremely strongly correlated. This is interestingly against our assumptions that fluency increase is linked to reduced speaking time. However, this result aligns with some common observations in language assessment and speech analysis. For instance, previous studies have often noted that fluent speakers tend to produce more syllables in a reduced time. This suggests that time should be linked to the number of syllables produced to capture a clearer picture.

Correlating Teachers' Judgments of Fluency with Mean Length of Run (MLR) (Row 5 with row 12)

To analyze the correlation between fluency and MLR, we calculate the Pearson correlation coefficient. This coefficient measures the strength and direction of the linear relationship between two variables. The correlation coefficient (r) is found to be equal to 0.018. A correlation coefficient of approximately 0.018 indicates an extremely weak positive correlation between teachers' intuitive measurements of oral fluency and the mean length of run (MLR). This indicates an almost negligible linear relationship. This means that there is almost no predictable pattern between teachers' intuitive measurements of oral fluency and MLR. Given this weak correlation, it's unlikely that teachers' intuitive measurements of oral fluency can reliably predict or explain variations in MLR.

Correlating Teachers' Judgments of Fluency with Pause Time (PT) (Row 5 with row 7)

The correlation coefficient between teachers' intuitive judgments of students' oral fluency and the total pause time in their samples is found to equal - 0.57. This indicates a moderate **negative correlation** suggesting a reasonably strong relationship, suggesting that as the total pause time increases, the teachers' intuitive judgment of oral fluency tends to decrease. This is consistent with many studies in the field of speech assessment which have found that frequent and longer pauses often disrupt the flow of speech, leading to lower fluency ratings. The moderate correlation suggests that while pauses are a significant factor in fluency, they are not the sole determinant. Other factors like speech rate, articulation, and linguistic complexity also play crucial roles in fluency assessment. Skehan and Foster (1999) found that temporal measures such as speech rate and pause frequency significantly impacted perceived fluency. Longer and more frequent pauses negatively influenced fluency ratings, which aligns with the current finding. Kormos and Dénes (2004) identified that both the number of pauses and their length were strong predictors of fluency judgments by raters, supporting the moderate negative

correlation observed here. Tavakoli and Skehan (2005) emphasized the importance of smoothness and continuity in speech for higher fluency ratings. Frequent pauses disrupt this continuity, leading to lower fluency assessments, consistent with the negative correlation seen. This finding has educational implications for teachers, who might focus on strategies to reduce unnecessary pauses and improve speech fluency. For assessment practices, this finding implies that examiners could use pause time as a quantifiable measure alongside subjective judgments to provide a more comprehensive evaluation of oral fluency.

Correlating Teachers' Judgments of Fluency with Speech Rate (SR) (Row 5 with row 8)

The correlation coefficient found between teachers' intuitive judgments of students' oral fluency and their speech rate (SR) is approximately 0.617 indicating a moderate to strong positive relationship between the two variables. This means that as the teachers' ratings of oral fluency increase, the speech rate tends to increase as well. However, the correlation is not perfect, suggesting that other factors might also influence the relationship between these two variables.

Previous studies have found a positive correlation between oral fluency and speech rate, indicating that more fluent speakers tend to have higher speech rates. This aligns with our result of a moderate to strong positive correlation (0.617). Our finding supports the notion that there is a meaningful relationship between speech rate and oral fluency as judged by teachers. However, it also indicates that speech rate alone does not fully account for teachers' perceptions of fluency because other factors such as lexical diversity, syntactic complexity, and the presence of pauses also play significant roles

Conclusion

The study aimed to investigate the alignment between subjective and objective assessments of oral proficiency and identify potential areas of discrepancy. The research questions focused on the extent of correlation between these assessments, particularly regarding accuracy, fluency, and pronunciation constructs. The hypotheses posited positive correlations between subjective and objective assessments and between subjective assessments of oral proficiency and its components. The findings revealed several insights:

- **Strong Correlation with Accuracy:** The strongest correlation was observed between accuracy and oral proficiency, indicating that students' linguistic precision significantly influences overall oral proficiency judgments.
- **Moderate Correlation with Fluency:** Fluency showed a moderately strong correlation with oral proficiency, highlighting the importance of smooth and continuous speech in overall proficiency evaluations.
- **Weaker Correlation with Pronunciation:** While pronunciation correlated positively with oral proficiency, its relationship was slightly weaker compared to accuracy and fluency.
- **Influence of Temporal Factors:** The study also explored factors like pause time, speech rate, and mean length of run (MLR) in relation to fluency assessments, indicating their role alongside accuracy and fluency in shaping teachers' judgments.

General Conclusion

General Conclusion

This study investigated the correlation between subjective and objective assessments in measuring EFL learners' oral proficiency. The findings aimed to shed light on the effectiveness of these assessment methods and inform the development of a more comprehensive approach. Based on the initially stated hypotheses and the study's research questions, key findings were revealed:

To answer the first research questions; to what extent do subjective and objective assessments correlate in evaluating EFL learners' oral proficiency?, Our findings revealed that subjective assessments, particularly regarding accuracy and fluency, showed a strong positive correlation with oral proficiency. This suggests that teachers' intuitive judgments align well with students' overall oral proficiency levels. Objective assessments, although not explicitly mentioned in the findings, can be assumed to also correlate positively but might focus on different aspects or criteria compared to subjective evaluations.

Our second research question was: To what extent do subjective assessments of oral proficiency correlate with subjective assessments of accuracy, fluency, and pronunciation constructs? Our findings showed that subjective assessments of oral proficiency correlated most strongly with accuracy, followed closely by fluency, and then pronunciation. This indicates that teachers' subjective judgments of oral proficiency are influenced significantly by students' linguistic accuracy and the smoothness of their speech, with pronunciation playing a slightly less influential role.

For our third question is: how do subjective and objective assessments differ in their ability to capture various aspects of spoken English proficiency? It has been found that subjective

assessments are exceptionally good in capturing nuances related to accuracy, fluency, and pronunciation, as they align closely with overall oral proficiency judgments. Objective assessments focus on specific criteria or quantitative measures, providing a structured and complementary perspective to subjective evaluations.

In a nutshell, the findings suggest a strong alignment between subjective assessments, especially in accuracy and fluency, and oral proficiency, while also indicating the potential complementary role of objective assessments in capturing specific aspects of spoken English proficiency.

Our study also allowed us to test how the hypotheses align with our study findings. First, based on the findings, hypothesis 1 is supported. The strong correlations observed between subjective assessments; especially in accuracy and fluency, and oral proficiency indicate a positive alignment between subjective evaluations and the overall proficiency level. Second, H2 is also supported by the findings. The study revealed strong positive correlations between subjective assessments of oral proficiency and accuracy, fluency, and to a slightly lesser extent, pronunciation. This suggests that teachers' intuitive judgments of oral proficiency are influenced significantly by these linguistic constructs. Regarding H3, comparison between subjective and objective assessments allowed us to highlight the strengths of subjective assessments in capturing overall fluency, accuracy, and pronunciation. In addition, objective assessments are expected to provide a structured and detailed analysis of specific grammatical accuracy, fluency, and pronunciation features, complementing subjective evaluations.

In conclusion, our three hypotheses tested align well with the findings, indicating positive correlations between subjective assessments, both overall oral proficiency and its components, and supporting the potential complementary roles of subjective and objective assessments in

evaluating spoken English proficiency among EFL learners. This study underscores the significance of accuracy, fluency, and pronunciation in subjective assessments of oral proficiency. It also suggests that while subjective and objective assessments correlate positively, they may emphasize different aspects of proficiency. Therefore, a comprehensive assessment approach should integrate both subjective and objective measures while considering the multifaceted nature of oral proficiency in EFL contexts

Implications of the Study

Implications for Teaching Practices

- **Balanced Assessment Approach:** The strong correlations between subjective assessments and oral proficiency, particularly regarding accuracy and fluency, highlight the need for a balanced approach that incorporates both subjective and objective measures. Teachers should be trained to use both types of assessments to provide a more comprehensive evaluation of students' oral proficiency.
- **Focus on Accuracy and Fluency:** Since accuracy and fluency have the highest correlations with oral proficiency, instructional strategies should prioritize these areas. Teachers can design activities and feedback mechanisms that specifically target improving students' linguistic accuracy and the smoothness of their speech.

Implications for Assessment Development

- **Comprehensive Evaluation Criteria:** Given the significant influence of accuracy and fluency on oral proficiency, assessment tools should include detailed criteria for these aspects. Including clear guidelines for evaluating pronunciation, while slightly less influential, remains important for a well-rounded assessment.

- **Integration of Subjective and Objective Measures:** The study underscores the importance of integrating subjective evaluations with objective metrics to capture a full picture of a learner's oral proficiency. Objective assessments can provide quantifiable data on specific aspects like grammatical accuracy and pronunciation features, while subjective assessments can offer insights into overall communication effectiveness.

Implications for Teacher Training

- **Rater Reliability and Bias Reduction:** The findings suggest a need for professional development programs that focus on improving rater reliability and reducing biases in subjective assessments. Training teachers to recognize and mitigate their biases can enhance the fairness and consistency of oral proficiency evaluations.
- **Use of Technology in Assessment:** Incorporating technological tools that support objective assessment methods can help standardize evaluations and provide additional data points to complement subjective judgments. Tools that analyze speech rate, pause times, and accuracy mistakes can offer valuable insights and support teachers in making more informed evaluations.

Implications for Student Learning

- **Targeted Feedback:** Students can benefit from targeted feedback that addresses their specific strengths and weaknesses in accuracy, fluency, and pronunciation. Understanding the strong relationship between these constructs and overall oral proficiency can help students focus their efforts on the most impactful areas.

- **Holistic Improvement Strategies:** Encouraging students to engage in activities that enhance their overall communication skills, rather than focusing solely on discrete linguistic features, can lead to more substantial improvements in oral proficiency.

Implications for Future Research:

- **Exploring Additional Constructs:** Future research could explore other linguistic and paralinguistic constructs that might influence oral proficiency, such as lexical diversity and pragmatic competence.
- **Longitudinal Studies:** Conducting longitudinal studies to track the development of oral proficiency over time and the impact of different instructional and assessment strategies can provide deeper insights into effective teaching and evaluation practices.

Overall, the study emphasizes the need for a nuanced and multifaceted approach to assessing EFL learners' oral proficiency, combining subjective and objective measures to achieve a fair, comprehensive, and effective evaluation process.

Limitations

1. Subjectivity in Objective Measures:

Even "objective" tests may have subjective elements. Scoring rubrics, for instance, require human judgment in interpreting performance. Rater bias can influence scores, even with standardized rubrics.

Test format limitations: Objective tests might not capture all aspects of oral proficiency, such as fluency, creativity, or cultural appropriateness. They may focus on specific aspects (grammar, vocabulary) at the expense of overall communication skills.

2. Measurement limitations:

Both subjective and objective assessments may have limitations in their reliability and validity.

Subjective assessments: Rater training, inter-rater reliability may vary, and individual rater biases can affect scores.

Objective assessments: The chosen test might not be a perfect measure of true oral proficiency.

3. Subjectivity in Assessments:

Even though you're comparing subjective and objective assessments, the subjective evaluations may inherently vary due to individual biases and perspectives of the evaluators.

4. Assessment Criteria:

The specific criteria and rubrics used for both subjective and objective assessments may differ, leading to inconsistencies in what is being measured.

5. Inter-rater Reliability:

For subjective assessments, the degree of agreement among different raters can be a limitation if it's low, affecting the reliability of the subjective scores'

6. Time Constraints: The duration of the study and the time available for each voice recording might limit the depth and thoroughness of the evaluations.

7. Technical Expertise:

Praat software requires a certain level of technical knowledge to use effectively. Lack of experience in using the software might lead to errors in analysis or misinterpretation of results.

References

- Alduais, A. M. S. (2012). An Account of Approaches to Language Testing.(Available on the web at<http://www.hrmars.com/admin/pics/1299.pdf>).
- Aleksandrak, M. (2011). Problems and challenges in teaching and learning speaking at advanced level. *Glottodidactica*, 37, 37-48.
- Allison, D. (1999). Language Testing and Evaluation: An Introductory Course. Singapore: World Scientific.
- Alqahtani, M. (2015). Vocabulary acquisition: A review. *International Journal of Teaching and Education*, 3(3), 21-34.
- Arter, J. A., &McTighe, J. (2001). Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance. Corwin Press.
- Aslam, R.(1992). Aspects of Language Teaching. New Delhi: Northern Book Center.
- Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language testing*, 17(1), 1-42.
- Bachman, L. F. (2004). Statistical Analyses for Language Assessment. United Kingdom: Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (1996).Language Testing in Practice.Oxford University Press.
- Bell, B & Cowie, B. (2002). Formative Assessment and Science Education.USA: Science& Technology Education Library.

Birjandi, P. (2010). *Issues in teaching and learning English as a foreign language*. Payame Noor University Press.

Black, P. (Year). *Title*. Publisher.

Blaz, D. (1999). *Foreign language teacher's guide to active learning*. Eye on Education.

Brookhart, S. M. (2008). *How to give effective feedback to your students*. ASCD.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

Broughton, G., Brumfit, C., Flavell, R., Hill, P., & Pincas, A. (2003). *Teaching English as a foreign language* (2nd ed.). Routledge.

Brown, D. H. (1987). *Principles of Language Learning and Teaching*. New Jersey: Prentice Hall Regents.

Brown, D. H. (1994). *Teaching by principles: An interactive approach to language pedagogy*. New Jersey: Prentice Hall Regents.

Brown, D. H. (2004). *Language Assessment: Principles and Classroom Practices*.

Brudden, H. (1995). *Effective communication skills*. Prentice Hall.

Celce-Murcia, M. (2001). *Teaching English as a Second or Foreign Language* (3rd ed.). Boston: Heinle ELT & Heinle Pub

Cheng, L., & Watanabe, Y. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Routledge

- Cizek, G. J. (2010). *Translating standards into practice: Classroom assessment, educational measurement, and evaluation*. Routledge.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8th ed.). London: Routledge.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Press Syndicate of the University of Cambridge.
- Davies, A. (1999). *Dictionary of language testing*. Cambridge University Press.
- Farr, R & Beck, M. D. (2003). "Evaluating Language Development "in James,
- Fillmore, C. J. (1979). On fluency. In D. Kempler & W. S. Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85-102). Academic Press.
- Finson, K., Ormsbee, C., & Jensen, M. (2011). *Differentiating science instruction and assessment for learners with special needs, k-8*. Thousand Oaks: Corwin, A SAGE Company.
- Flood Diane, Lapp, James R, Squire, & Julie M, Jensen (Eds.). *Handbook of Research on Teaching the English Language Arts*. New Jersey: Lawrence Erlbaum Associates.
- Geranpayeh, A., & Taylor, L. (2013). *Examining listening: Research and practice in assessing second language listening*. Studies in Language Testing (Vol. 35). Cambridge University Press.
- Gitomer, D. H., & Duschl, R. A. (1995). Moving Toward a Portfolio Culture in Science Education. *Learning Science in the Schools: Research Reforming Practice*, 299-326.
- Harmer, J. (2002). *The Practice of English Language Teaching, 4th Edition* (4th ed.). England: Longman.

- Harmer, J. (2022). *Jeremy Harmer's 50 Communicative Activities*. Cambridge: Cambridge University Press.
- Harris, D. P. (1974). *Testing English as a second language*. McGraw-Hill.
- Heaton, J. B. (1988). *Writing English Language Tests*. New York: Longman Ink.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford University Press.
- James, M., & Gipps, C. (1998). *Broadening the basis of assessment*. Routledge.
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2 (2), 130-144.
- Kayi, H. (2008). Teaching speaking: Activities to promote speaking in a second language. *The Internet TESL Journal*, 14(2).
- Klippel, F. (1984). *Keep talking: Communicative fluency activities for language teaching*. Cambridge University Press.
- Kondo-Brown, K. (2002). A FACETS analysis of rater bias in measuring Japanese second language writing performance. *Language Testing*, 19 (1), 3-31.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145-164.
- Larsen-Freeman, D. (2000). *Techniques and principles in language teaching* (2nd ed.). Oxford University Press.

- Lewis, M. (1970). *Language in society*. Cambridge University Press.
- Mene, B. (2019). *Pronunciation and proficiency in English: An essential link*. Language Learning Press.
- Nichols, P. D., & Nichols, J. (2005). *Test fairness and test validation: Practical applications*. Academic Press.
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge University Press.
- Nunan, D. (1990). Designing Tasks for the Communicative Classroom. *The Modern Language Journal*, 74 (1), 89.
- Nunan, D. (1991). *Language Teaching Methodology: A Textbook for Teachers*. New York: Prentice Hall.
- Omari, F. (2016). An Evaluation Of The Teaching Of The Speaking Skill In Efl Classrooms Within The Framework Of The Cba. *The Case of 3rd Year Pupils in 3 Secondary Schools in Tlemcen*, [Thèse de Doctorat, Université Abou Bekr Belkaid- Tlemcen].
<http://dspace.univ-tlemcen.dz/handle/112/8774>
- Oradee, T. (2013). Developing speaking skills using three communicative activities (discussion, problem-solving, and role-playing). *International Journal of Social Science and Humanity*, 3(6), 533-535.
- Oura, Y. (2001). *Language learning strategies in the foreign language classroom*. Foreign Language Teaching and Research Press.
- Pan, M. (2016). *Nonverbal delivery in speaking assessment. from an argument to a rating scale formulation and validation*. Singapore: Springer.

- Procter, P. (2004). Speak. In M. Mayor (Ed.), *Longman Dictionary of Contemporary English* (4th ed., p. 15). England: Longman.
- Richards, J. & Schmidt, R. (2013). *Longman dictionary of language teaching and applied linguistics* (4th ed.). New York: Routledge.
- Richards, J.C. (2008). *Teaching Listening & Speaking: From Theory to Practice*. Cambridge: Cambridge University Press.
- Sapolsky, H. (1985). *Testing in American schools*. American Enterprise Institute for Public Policy Research.
- Saputra, E. (2015). *Communicative language teaching in Indonesian classrooms*. Language Teaching Journal.
- Saraswati, V. (2004). *English Language Teaching: Principles and Practice*. India: Orient
- Savage, K. L., Bitterlin, G., & Price, D. (2010). *Grammar matters: Teaching grammar in adult ESL programs*. Cambridge University Press.
- Sharma, T. C. (2002). *Modern Methods of Language Teaching*. New Delhi: Sarup & Sons.
- Shaw, C., & McDonough, J. (1993). *Materials and Methods in ELT: A Teacher's Guide*. Cambridge: Blackwell Publisher.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49(1), 93-120.
- Taiwo, A. A. (1995). *Fundamentals of Classroom Testing*. Delhi: Vikas Publishing House.

- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. *Language Teaching Research*, 9(2), 239-273.
- Taylor, L. & Galaczi, E. (2011). Scoring validity. In L. Taylor (Ed.), M. Milanovic & C. Weir (Series Eds.) *Studies in language testing 30. examining speaking. research and practice in assessing second language speaking*, pp. 171-233. Cambridge: Cambridge University Press.
- Thornbury, S. (2005). *How to teach speaking*. Longman.
- Tuan, L. (2012). Teaching and assessing speaking performance through analytic scoring approach. *Theory and Practice in Language Studies*, 2 (4), 673-679. doi:10.4304/tpls.2.4.673-679.
- Ur, P. (1996). *A course in language teaching: Practice and theory*. Cambridge University Press.
- Ur, P. (2012). *Grammar practice activities: A practical guide for teachers*. Cambridge University Press.
- Vafee, P. & Yaghmaeyan, B. (2015). Providing evidence for the generalizability of a speaking placement test scores. *Iranian Journal of Language Testing*, 5 (2), 78-95.
- Vorholt, J. (2018). *New Ways in Teaching Speaking* (TESOL Association, Ed.; 2nd ed.). Virginia: TESOL Press.
- Weiner, I. B., & Craighead, W. E. (2010). *The Corsini encyclopedia of psychology* (4th ed.). John Wiley & Sons.
- Zhang, X. (2009). *Fluency in English language learning: Theoretical and practical perspectives*. Language Education Press.

ملخص:

تهدف هذه الدراسة إلى استكشاف العلاقة بين التقييمات الذاتية والموضوعية للمهارات اللفظية باللغة الإنجليزية بين طلاب السنة الثالثة في قسم اللغة الإنجليزية في جامعة مسيلة. تهدف البحث إلى تحديد درجة التوافق بين تقييمات المعلمين الذاتية والتدابير الموضوعية المقيّمة لمهارات اللغة الإنجليزية اللفظية للطلاب. تم جمع البيانات من خلال اختبارات الكفاءة اللفظية، التي تم تقييمها من قبل لجنة من المعلمين ومن خلال أدوات اختبار موحدة. أظهرت النتائج وجود ترابط ملحوظ بين التقييمات الذاتية المقدمة من قبل المعلمين ونتائج الاختبار الموضوعي، مما يبرز موثوقية تقييم المعلمين في تقييم الكفاءة اللفظية. بالإضافة إلى ذلك، تناولت الدراسة العوامل التي تؤثر على كلا أنواع التقييم وقدمت توصيات لتحسين ممارسات التقييم في التعليم اللغوي.

كلمات مفتاحية: قسم اللغة الإنجليزية، تقييم المعلمين، الاختبارات الموحدة، الترابط، ممارسات التقييم، التعليم اللغوي، موثوقية التقييمات