

**People's Democratic Republic of Algeria**  
**Ministry of Higher Education and Scientific Research**  
**Mohamed Boudiaf University, M'sila**

**Faculty of Letters and Languages.**

**Department of Letters and English Language.**

**N°:.....**



**Domain: Foreign Languages**  
**Stream: English Language**  
**Option: Didactics and Applied Languages.**  
**Online programme**

**Investigating the Validity and Reliability of Oral Proficiency Assessment for EFL Learners: Case of Secondary Education Students at Mouloud Quassem Secondary School, Setif, Algeria.**

Dissertation Submitted to the Department of Letters and English Language in  
Partial Fulfilment of the Requirements for the Master's Degree in  
**Didactics and Applied Languages – Online Programme**

**Candidates:**

Aissa BEKRAR

Lyazid BOUKERROUCHA

**Board of Examiners**

<b>Dr. Nouredine REFICE</b>	<b>University of M'sila</b>	<b>Chairperson</b>
<b>Dr. Mohammed Djemoui SABER</b>	<b>University of M'sila</b>	<b>Supervisor</b>
<b>Mr. Abdelmadjid TAYOUB</b>	<b>University of M'sila</b>	<b>Examiner</b>

**June 2025**

## Abstract

This study investigates the validity and reliability of oral proficiency assessments in the context of English as a Foreign Language (EFL) education at Mouloud Quassem Secondary School in Algeria. Employing a mixed-methods approach, the research combined quantitative and qualitative data collection techniques, including Likert-scale questionnaires administered to 170 students and 10 teachers. The findings revealed that while students exhibited moderate self-perceptions of their oral proficiency, they reported higher anxiety during formal assessments and preferred informal methods such as storytelling and role-plays. Teachers acknowledged the challenges of large class sizes, student anxiety, and the lack of standardized rubrics, yet they perceived students' oral proficiency levels as generally good. The study highlighted the importance of ensuring validity and reliability in oral assessments through diversified methods, technological integration, and transparent grading criteria. The results underscore the need for institutional support, teacher training, and student-centered strategies to enhance oral proficiency evaluation. In conclusion, the research advocates for balanced assessment practices that prioritize communicative competence while addressing contextual challenges in EFL settings.

**Keywords:** Oral proficiency assessment, Student perceptions, Teacher challenges, Language learning strategies, Communicative competence.



## **Dedication**

I dedicate this work to my family, whose sacrifices, encouragement, and unconditional support made it possible.

Aissa Bekrar





## **Dedication**

This thesis is dedicated to all those who contributed, directly or indirectly, to my academic journey.

To every educator who taught me not only lessons from books, but also values of perseverance, critical thinking, and self-belief.

To every student who wakes up each day with determination to learn and improve, even in the face of challenges this work is for you.

And most of all, to the inner voice within me that refused to give up, even in the darkest moments I thank you for keeping me going.

Lyazid Boukerroucha



## **Acknowledgments**

First and foremost, I would like to thank God Almighty for granting me the strength, patience, and determination to complete this academic journey.

I would like to express my sincere thanks and deep appreciation to my supervisor, Dr. Mohammed Djemoui SABER, for his invaluable guidance, constructive feedback, and constant support during every stage of this research. His insight and dedication played a crucial role in shaping this work, and I remain truly thankful.

## List of Abbreviations

ACTFL	American council on the Teaching Foreign languages
ADEPT	A Developmental English Language Proficiency Test
ASR	Automatic Speech Recognition
CAOT	Computer-Assisted Oral Tests
CEFR	The common European framework of reference for languages
COPI	Computerized Oral Proficiency Interview
EFL	English as a Foreign Language
EIT	Elicited Imitation Task
ELLs	English Language Learners
FL	Foreign language
H <sub>0</sub>	Null Hypothesis
H <sub>1</sub>	Alternative Hypothesis
OPI	Oral Proficiency Interview
SLA	Second language acquisition
TBLT	Task-based language teaching

## List of Tables

Title	Page
Table 01: Reliability Statistics based on Cronbach's Alpha.	45
Table 02: Validity statistics based on Pearson Correlation	46
Table 03: Gender of the teachers	46
Table 04: Years of Teaching Experience	47
Table 05: The average class size in your oral lessons	48
Table 06: Language(s) of Teaching	48
Table 07: Teachers' perceptions of current assessment practices.	49
Table 08: Descriptive statistics of teachers' responses on current assessment practices.	50
Table 09: Descriptive results of teachers' challenges in assessing oral proficiency.	52
Table 10: Descriptive results on how ensuring fairness and objectivity in oral assessment.	55
Table 11: Descriptive results on strategies and tools used by teachers to improve students' oral skills.	55
Table 12: Frequency and percentage of teachers' beliefs on the effectiveness of self assessment in improving students' oral proficiency.	56
Table 13: Descriptive results of teachers' responses on additional support and resources for conducting oral assessments	58
Table 14: Reliability Statistics based on Cronbach's Alpha	60
Table 15: Validity statistics based on Pearson Correlation.	60
Table 16: Students' perceptions of the oral proficiency components.	61
Table 17: Descriptive results of students' responses of the oral proficiency components.	61
Table 18: Students' perceptions of teaching and learning methods	63
Table 19: Descriptive results of students' responses on teaching and learning methods	63

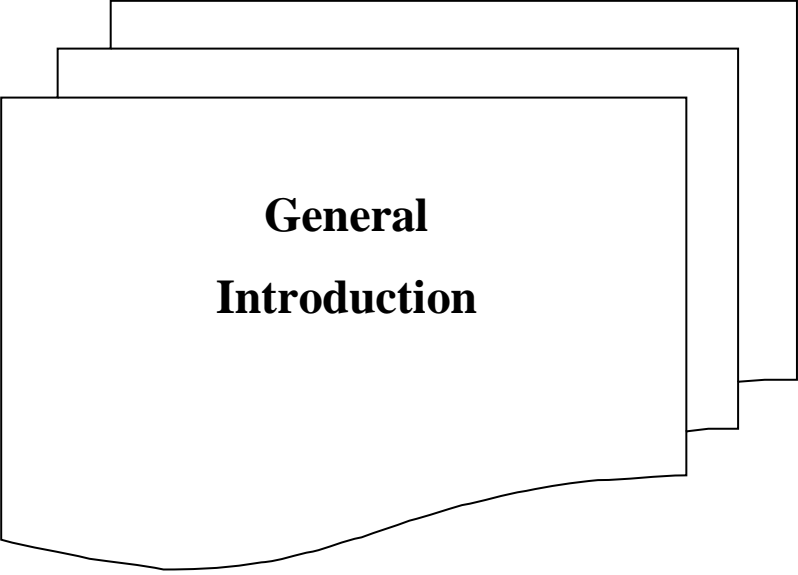
Table 20: Students' perceptions of assessment methods.	65
Table 21: Descriptive results of students' responses on assessments methods.	65
Table 22: Students' perceptions of challenges and suggestions.	67
Table 23: Descriptive results of students' responses on challenges encountered and suggestions for improvement.	68
Table 24: T-Test Analysis of Students' self-perceived oral proficiency levels.	70
Table 25: T-Test Results for students' Acceptance of Teaching and Learning Methods.	71
Table 26: T-Test Results for Students' challenges in Improving oral proficiency.	73
Table 27: T-Test Results for Teachers' perceptions of students' oral proficiency level.	74

## Table of Contents

Table of Contents.....	IX
List of Tables.....	XI
List of Abbreviations .....	V
General Introduction:.....	1
Statement of the Problem.....	1
Aims of the Study .....	2
Significance of the Study .....	2
Students’ Research Questions: .....	3
Students’ Hypotheses:.....	3
Teachers’ Research Questions:.....	4
Teachers’ Hypotheses: .....	4
Chapter one: The literature review .....	6
Introduction .....	6
1.1 Definition of oral proficiency: .....	6
1.2 Key components of oral proficiency:.....	7
1.2.1 Fluency and accuracy: .....	8
1.2.2 Vocabulary and comprehension:.....	8
1.2.3 Communication skills:.....	8
1.2.4 Cultural and linguistic integration:.....	9
1.3 Key factors enhancing oral proficiency:.....	9
1.3.1 Linguistic knowledge: .....	9
1.3.2 Practice opportunities: .....	10
1.3.3 Motivation: .....	10
1.3.4 Corrective feedback: .....	11
1.3.5 Exposure to authentic input: .....	11
1.4 The importance of assessing oral proficiency:.....	12
1.4.1 Measures real world communication skills: .....	12
1.4.2 Enhances language learning and teaching: .....	13
1.4.3 Facilitates professional success:.....	13
1.4.4 Promotes social and cultural integration: .....	13
1.4.5 Develops intercultural competency:.....	13
1.5 Models of oral proficiency assessments: .....	14

1.5.1 Presentation: .....	15
1.5.2 Dialogues and class participation: .....	15
1.5.3 Oral discussions: .....	16
1.5.4 Role plays: .....	16
1.5.5 Interviews: .....	16
1.5.6 Story telling: .....	17
1.5.7 Simulations: .....	17
1.5.8 Picture description:.....	17
1.6 Framework of oral proficiency assessments: .....	18
1.6.1 Validity definition: .....	19
1.6.2 Reliability .....	28
1.6.3 Challenges in ensuring reliability in oral proficiency assessments: .....	31
1.6.4 Examples of studies emphasizing validity and reliability in oral proficiency assessments: .....	32
1.7 Assessing the validity and reliability of oral proficiency assessments: .....	34
1.7.1 Technological Integration .....	35
1.7.2 Multiple Assessment Methods .....	35
1.7.3 Statistical and Methodological Approaches .....	36
1.7.4 Socio-Cultural and Contextual Considerations .....	36
1.8 Challenges in achieving validity and reliability in oral proficiency assessments: .....	37
1.8.1 Complexity of Language and Assessment Design: .....	37
1.8.2 Cultural and Contextual Factors: .....	38
1.8.3 Resource Limitations and Standardization: .....	38
1.8.4 Reliability and Consistency: .....	38
Conclusion.....	39
Chapter two: Methodology, results and discussion. ....	42
Introduction .....	42
2.1 Research Methodology and Design.....	42
2.1.1 The Research Method.....	42
2.1.2 The population and study sample:.....	43
2.1.3 The study setting: .....	43
2.1.4 Data Collection Instruments .....	43
2.1.5 Procedure.....	44
2.1.6 Data Analysis tools .....	45
2.2 Data analysis and discussion .....	45
2.2.1 Analysis of the teachers' questionnaire data.....	45
2.2.2 Analysis of the students' questionnaire data.....	60

2.3 Hypothesis Testing:.....	70
2.3.1 Students' Hypotheses: .....	70
2.3.2 Teachers' Hypotheses: .....	74
Conclusion.....	78
General conclusion .....	81
Limitations of the study .....	81
Sample characteristics.....	81
Methodological Constraints .....	80
Contextual and theoretical boundaries .....	81
Pedagogical Implications .....	81
Curriculum Development.....	81
Assessment Practices .....	83
Teacher Training and Support .....	83
Student-Centered Approaches .....	84
Recommendations for Future Research .....	84
Methodological Advancements .....	84
Theoretical Expansions .....	84
Technological Innovations .....	85
Policy and Practice .....	85
List of references .....	89
Appendices.....	94
Abstract in Arabic.....	106
Abstract in French.....	107



**General  
Introduction**

## General Introduction:

### 1. Statement of the Problem

In Algerian secondary educational institutions, learners of English as a Foreign Language (EFL) frequently encounter difficulties in achieving oral proficiency, primarily due to the insufficient exposure to authentic communicative contexts, a predominance of teacher-centered pedagogical approaches, and curricula that prioritize examination outcomes.

Within the context of Mouloud Quassem Secondary School, students confront supplementary obstacles, which include: low self-confidence in verbal communication stemming from apprehension regarding potential errors and an absence of practical engagement. Additionally, the excessive focus on grammatical accuracy and written expression within the national curriculum, which allows minimal opportunity for interactive speaking exercises. Also the socio-cultural and linguistic barriers, as students seldom utilize English beyond the confines of the classroom.

Educators also experience challenges, such as managing extensive class sizes, possessing limited professional development in oral assessment methodologies, and relying on conventional pedagogical strategies such as the evaluation methods centered on rote memorization.

Despite these complications, there exists little research examining the perceptions of Mouloud Quassem students regarding their oral competencies or the modalities through which educators evaluate these skills. This study investigates the disconnect between students' self-perceptions, their acceptance of current teaching methods, and the challenges teachers face in fostering oral proficiency.

### 2. Aims of the Study

The primary goal of this study is to evaluate Mouloud Quassem Secondary School students' self-perceptions of their oral proficiency, focusing on fluency, grammar, pronunciation, vocabulary, and confidence. Additionally, it aims to measure students' acceptance of existing teaching methods and oral assessment formats, determining which approaches cause anxiety and which they find most beneficial. From the teachers' perspective, the study examines the challenges they encounter in assessing oral skills, including time constraints, large class sizes, and the absence of reliable evaluation tools. Finally, the research proposes context-specific strategies to enhance oral proficiency, such as integrating technology-assisted learning, implementing low-pressure speaking activities, and providing teacher training on communicative language techniques. These objectives collectively aim to bridge the gap between students' needs and current instructional practices.

### 3. Significance of the Study

This research holds significant implications for EFL education at Mouloud Quassem Secondary School and similar Algerian institutions. For students, the findings could lead to more engaging and less stressful speaking activities, thereby boosting confidence, participation and encouraging autonomous practice through technology and peer collaboration.. Teachers may benefit from evidence-based recommendations on alternative assessment methods and classroom strategies that accommodate large groups while fostering oral skills. At an institutional level, the study highlights the need for curriculum adjustments to balance written and oral proficiency development, as well as potential investments in teacher training and digital tools. Beyond the immediate school context, the research contributes to broader discussions on EFL pedagogy in Algeria, advocating for reforms that prioritize communicative competence alongside grammatical accuracy. Ultimately, the study

## General Introduction

---

aspires to create a more supportive and effective environment for developing students' English speaking abilities.

### **4. Students' Research Questions:**

- What is the level of students' self-perceptions of their oral proficiency in terms of fluency, grammar, pronunciation, vocabulary, and confidence?
- To what extent do students accept the current teaching and learning methods?
- How do students perceive the different oral assessment methods?
- Are there any challenges that students face in improving their oral proficiency?

### **5. Students' Hypotheses:**

- Students' self-perceptions of their oral proficiency are generally good, but they tend to rate fluency and confidence lower than other components (grammar, pronunciation, vocabulary) due to anxiety during spontaneous speaking.
- There is a moderate level of acceptance of teaching and learning methods among students.
- There is a moderate level of acceptance of oral assessment methods among students, where they report higher anxiety during formal assessments (e.g., presentations, interviews), and prefer informal assessments (e.g., storytelling, picture description) due to lower stress levels.
- Students face challenges in improving their oral proficiency, mainly due to a lack of practice opportunities outside the classroom and cultural barriers. They suggest implementing more interactive activities and technology-based tools (e.g., apps, recordings)

## General Introduction

---

### 6. Teachers' Research Questions:

- How do teachers evaluate students' oral proficiency?
- What challenges do teachers face in assessing oral proficiency?
- What additional support do teachers need to improve oral proficiency assessment?

### 7. Teachers' Hypotheses:

- Teachers perceive that students generally have a good level of oral proficiency.
- Teachers face numerous challenges in assessing oral proficiency, particularly large class sizes and student anxiety. Meanwhile, the use of standardized rubrics and peer assessments are among the most commonly used strategies to ensure fairness.
- Teachers need additional support to improve oral proficiency assessment, including training workshops, speech analysis software, and smaller class sizes to enhance assessment quality.



**Chapter one:**  
**The literature review**

### 1 Chapter one: The Literature review

#### Introduction

The assessment of oral proficiency is a critical component of English as a Foreign Language (EFL) education, as it directly reflects learners' communicative competence and real-world language use. However, ensuring the validity and reliability of oral proficiency assessments remains a persistent challenge for educators and researchers, particularly in secondary education contexts where standardized testing may not fully capture students' speaking abilities. This chapter explores the theoretical foundations of oral proficiency assessment, focusing on the importance of assessing oral proficiency, key concepts such as validity, reliability, examples of studies emphasizing validity and reliability in oral proficiency assessments, and the specific challenges associated with evaluating EFL learners' speaking skills. By examining established frameworks and models of communicative competence, this section aims to provide a comprehensive understanding of how oral proficiency assessments should be designed, administered, and evaluated to ensure its validity and reliability. The discussion will also address factors influencing assessment outcomes, such as rater bias, complexity of language, assessment design, and contextual constraints.

#### 1.1 Definition of oral proficiency:

Oral proficiency refers to the ability of effectively communicate with others verbally in a certain language, including different linguistic and communicative competencies. It is such an important skill for learners who are learning foreign language (FL), since it does not only require the mechanics of speech production but also the capability of transmitting utterances effectively and appropriately in various situations.

## Chapter one: The Literature review

---

Assessing oral proficiency on the other side, is considered a crucial matter for most EFL learning and teaching syllabi, as it poses a challenge for teachers to make a valid, reliable and effective oral assessment test to evaluate students' oral performance in classrooms.

Oral proficiency involves the ability to articulate thoughts and ideas clearly and effectively in a spoken language. It encompasses various linguistic and communicative competencies like: pronunciation, grammar, vocabulary, fluency and naturalness, as well as the pragmatic competencies which are the more practical and concrete language skills necessary to engage successfully in daily interactions such as making requests and responding to questions (Bachman & Palmer, 2010). In this regard, (Fulcher, 2003) stated the oral proficiency encompasses fluency, pronunciation, vocabulary, grammatical accuracy and the ability to engage in meaningful discourse.

Oral proficiency is an essential and indispensable aspect of the process of learning new languages, as it enables individuals to interact and actively participate in real life settings and situations, including both social and professional engagements (Celce-Murcia, D. Brinton, & Snow, 2014).

### **1.2 Key components of oral proficiency:**

Oral proficiency is a complex skill that incorporates various components vital and essential for effective and proficient communication. These components encompass fluency, accuracy, vocabulary, comprehension and communication skills, each one of them plays a significant role in the overall capability of speaking a language with a high level of proficiency.

### **1.2.1 Fluency and accuracy:**

Fluency refers to the ability to speak smoothly and effortlessly. Fillmore described fluency as “the ability to talk at length with few pauses, the ability to fill time talk.”(Fillmore, 1979). Respectively, Lennon defined fluency as the “rapid, smooth, accurate, lucid and efficient translation of thoughts or communicative intention under the temporal constraints of on-line processing.”(Lennon, 2000). While accuracy involves the correct usage of grammatical structures and appropriate lexical items. Research studies suggest that the effect of listening to various TV shows and programs significantly enhances different components of speaking proficiency. In fact, the result showed that the enhancement of vocabulary usage, as a crucial element of speaking proficiency, is more pronounced. Conversely, the enhancement in accuracy is less significant (Krashen S. , 1985).

### **1.2.2 Vocabulary and comprehension:**

Having an extensive vocabulary and good understanding abilities are crucial and vital for expressing ideas clearly and understanding others. Vocabulary enhancement is frequently emphasized as a significant area of improvement in language learning (Krashen S. , 1985). On the flipside, comprehension is enhanced through the integration of cultural and linguistic components in language learning. Integrating these aspects is crucial for developing oral proficiency (Byram, 1997).

### **1.2.3 Communication skills:**

Communicative skills are essential competencies that enable proficient interaction and the exchange of information within diverse professional and educational settings. These skills include a wide range of abilities inclusive of verbal and non-verbal communication, active listening and the capability to effectively manage various communicative scenarios. In the realm of language teaching and learning, communicative skills are developed through

interactive methodologies such as “role-play, discussion and situation modeling”. These methods improve students’ motivation and efficacy in acquiring language proficiency, which is vital for their professional success in non-linguistic fields (Canale & Swain, 1980).

### **1.2.4 Cultural and linguistic integration:**

Cultural and linguistic integration represents a complex process that incorporates cultural elements into language learning and various social contexts to foster communication and comprehension. This incorporation of cultural elements into language education not only enhances linguistic proficiency but also enriches the overall communicative abilities. Understanding cultural contexts facilitates the development of more authentic and meaningful competencies (Byram, 1997). So, this integration is indispensable for learners to effectively navigate real world interactions.

### **1.3 Key factors enhancing oral proficiency:**

Oral proficiency in the realm of second language acquisition (SLA) constitutes a complex construct that is influenced by an array of factors, which encompass linguistic knowledge, opportunities for practice, motivation, corrective feedback, and exposure to authentic linguistic input. These components interact in an intricate manner to define learners’ abilities to communicate proficiently in second language. Researchers in SLA have examined these factors intensively, providing valuable insights into their contributions to the advancement of oral proficiency.

#### **1.3.1 Linguistic knowledge:**

Linguistic knowledge is the fundamental basis of oral proficiency, incorporating vocabulary, grammar and pronunciation. It is imperative for learners to develop a comprehensive understanding of these elements to engage effectively in communication (Juffs & DeKeyser, 2006). The development of linguistic knowledge is shaped by both

internal cognitive processes and external pedagogical strategies, which can be customized to enhance the comprehension and use of the language (Nunan, 2001).

### **1.3.2 Practice opportunities:**

Opportunities to practice, especially within interactive environments are of a paramount importance for the enhancement of oral proficiency. Group work and pair activities in classroom provide learners with the opportunity to negotiate meaning and refine interaction strategies, which are essential for the enhancement of oral skills (Krashen S. D., 1982). Longitudinal researches have demonstrated that augmented exposure and rigorous practice in foreign language classrooms significantly improve fluency and lexico-grammar, thereby emphasizing the significance of regular and consistent practice opportunities (Larson-Hall, 2008). Numerous studies have indicated that task-based language teaching (TBLT) represents an ideal opportunity for the practice of communication and serves as a distinguished approach to enhancing oral proficiency among language learners by prioritizing meaningful communication and the real world language use. This methodology underscores the importance of engaging learners in tasks that replicate authentic language scenarios. TBLT fosters a dynamic learning environment that promotes active participation and interaction among students. For example, the integration of drama into TBLT has been demonstrated to evoke enthusiasm and self confidence in students, thereby significantly enhancing oral English proficiency by creating a positive and engaging atmosphere for language learning (Stinson & Winston, 2011).

### **1.3.3 Motivation:**

Motivation serves as a fundamental facilitator for the success of language acquisition endeavors. It significantly affects the level of engagement of learners with the language, as well as their tendency to actively pursue practice opportunities and authentic linguistic input.

Learners who exhibit a high degree of motivation are more inclined to persevere in their language studies and utilize the resources at their disposal, leading to improved oral proficiency outcomes (Fathman, 1976).

### **1.3.4 Corrective feedback:**

Corrective feedback serves as crucial role in enhancing the oral competencies of learners. It assists learners in identifying and rectifying language inaccuracies, indicates instances that are deemed unacceptable, and proposes alternative solutions, thus promoting the acquisition of language skills (Lyster & Ranta, 1997). To ensure the effectiveness of feedback, it must be adapted to the specific context while also accounting for the preferences of learners and the interactive dynamics between the feedback giver and taker. This approach guarantees that the feedback is constructive and fosters the learner's development (Lyster & Ranta, 1997).

### **1.3.5 Exposure to authentic input:**

Exposure to authentic input is critical for the advancement of oral proficiency. Authentic materials equip learners with real world language use examples, thereby augmenting their ability to comprehend and generate the language across different contexts (Krashen S. D., 1982). The significance of input is emphasized in numerous theories of second language acquisition (SLA), which underscore its critical role is supplying learners with the essential linguistic data for effective language acquisition (Krashen S. , 1985).

While these factors are critical for developing oral proficiency, it is imperative to acknowledge that they do not work independently. The interaction among these components may fluctuate based on the distinctive characteristics of individual learners and environmental contexts. As an illustration, the success of corrective feedback may hinge on the learners' intrinsic motivation and the quality of input they are provided. Furthermore, although

linguistic knowledge and opportunities for practice are fundamental, these elements must be supported by conducive learning environment that fosters the utilization of the target language. Understanding these dynamics can assist educators to better devise more effective language learning programs that address the diverse needs of learners.

### **1.4 The importance of assessing oral proficiency:**

Assessing oral proficiency is critical in the realm of language education and professional settings, as it significantly influences the effectiveness of communication and the resultant of learning outcomes. Oral proficiency assessment enables educators and employers to assess individual's capacity for oral communication in a foreign language, which is indispensable for achieving both academic excellence and professional effectiveness. The assessment of oral proficiency encompasses not only the measurement of linguistic capabilities but also the promotion of learner autonomy and intercultural competence. This comprehensive approach guarantees that learners are adequately equipped to confront the complexities of real world communication. The following sections engage in a comprehensive examination of the significance of assessing oral proficiency, supported by perspectives derived from diverse research context.

#### **1.4.1 Measures real world communication skills:**

Oral proficiency assessments systematically evaluate the capability of individuals to employ language effectively across diverse contexts, including social interactions, academic discussions and professional presentations (Fulcher, 2003). In contrast to written assessments, oral assessments target fluency, pronunciation and interactive competence, all of which are vital for successful communication.

### **1.4.2 Enhances language learning and teaching:**

Oral proficiency assessments provide valuable insights for both learners and educators. They enable the identification of strengths and weaknesses in verbal communication, thus allowing for targeted pedagogical strategies (Fulcher, 2003). By assessing oral competencies, educators can adapt their instructional approaches to improve students' speaking skills, thereby fostering more comprehensive process of language learning.

### **1.4.3 Facilitates professional success:**

In a professional, proficient oral communication skills are essentially correlated with career growth, as employers emphasize the importance of effective verbal communication in organizational contexts. Therefore, oral communication education is critical for students' future personal and professional accomplishments. It emphasizes the vital importance of communication in improving organizational processes and the comprehensive dynamics of the organization. Developing as a whole person and realizing achievement as an individual in one's career and within the sphere of business (Morreale, Osborn, & Pearson, 2000).

### **1.4.4 Promotes social and cultural integration:**

For individuals who have migrated and those acquiring a new language, oral proficiency assessment is of paramount importance for successful social integration. Evaluative framework such as the common European framework of reference for languages (CEFR) provides a means for individuals to exhibit their competency in participating in everyday conversations, thereby enhancing their involvement within the community (Europe, 2020).

### **1.4.5 Develops intercultural competency:**

Projects such as CEFcult underscore the significance of intercultural competence in conjunction with oral proficiency, equipping learners for effective professional

## Chapter one: The Literature review

---

communication within varied cultural frameworks. The European-funded initiative CEFcult (2009-2011) was established with the objective of producing a digital platform dedicated to the collaborative evaluation of oral competencies and intercultural proficiency in the context of foreign language acquisition. CEFcult offers a collection of authentic communication scenarios, each comprising various tasks through which learners can evaluate their competencies for intercultural professional engagement in multiple European languages. While the tool is amenable to autonomous use, in a classroom environment, educators can assign appropriate scenarios, facilitate the execution of speaking tasks, and document the oral communication activities. Upon uploading the recorded samples to the online platform, they may be evaluated by the instructor, by other specialists, or by fellow learners. Furthermore, the tool has been constructed to assist in self-assessment. By fostering self- and peer-evaluation, CEFcult effectively engages learners in a reflexive learning process that ultimately promotes enhanced autonomy. Nevertheless, this framework also empowers the educator to play an integral role in this same developmental process (Beaven & Neuhoff, 2011).

### **1.5 Models of oral proficiency assessments:**

The employment of presentation, dialogue, class participation, oral discussion, role play, and interview as assessment instruments substantially augments the enhancement of oral proficiency among learners of English as a Foreign Language (EFL). These methodologies furnish authentic, interactive, and stimulating platforms for learners to engage in the practice and refinement of their speaking abilities. They assist in reducing anxiety, augmenting motivation, and enhancing fluency and accuracy, which are imperative for the acquisition of proficiency in a foreign language. Each of these instruments presents distinct advantages and challenges, which collectively foster a comprehensive learning environment conducive to the advancement of oral proficiency.

### 1.5.1 Presentation:

Oral presentations represent a prevalent approach for the assessment of speaking competencies among EFL learners. Such assignments necessitate that students carefully organize and structure their thoughts, thereby facilitating the improvement of their fluency and accuracy over an extended duration. Empirical research has indicated a notable advancement in learners' performance when assessments were conducted utilizing a structured rating scale that encompassed elements of preparation, organization, and presentation (Mazdayasna, 2012). Nevertheless, obstacles such as performance anxiety and insufficient presentation skills may impede the efficacy of this assessment method. Educators assume a pivotal role in equipping students with the requisite skills for proficient presentations by transmitting fundamental competencies and providing direction throughout the preparatory process (King, 2002).

### 1.5.2 Dialogues and class participation:

Engagement in dialogues and participation in classroom activities are fundamental to the enhancement of oral proficiency, as they promote the spontaneous application of language and the development of interactional competence. Such activities are instrumental in mitigating the affective filter while simultaneously enhancing learners' self-confidence (Swain & Lapkin, 1998).

The implementation of dialogues within a controlled environment, exemplified by the utilization of a multimodal dialogue system, facilitates the provision of reliable evaluations of interactional competencies by adaptively modifying the complexity of topics to appropriately challenge the limits of learners' proficiency (Johnson & Valente, 2009).

### 1.5.3 Oral discussions:

Discussions serve as a crucial mechanism for the enhancement of oral proficiency, as they foster critical thinking skills and enable learners to articulate their thoughts in a coherent framework. Analytical results confirm that the adoption of the discussion task model substantially elevates oral communication skills and the grasp of critical thinking among EFL learners (Iman & Angraini, 2019). The integration of topics that resonate with students' interests can substantially elevate their motivation and engagement, thereby contributing to improved performance in oral proficiency evaluations (Swain & Lapkin, 1998).

### 1.5.4 Role plays:

Role plays present a dynamic and interactive modality for evaluating speaking competencies, emphasizing learners' capacity to employ language within diverse contexts. They serve as particularly useful tools for evaluating interactional competence and the capability of language use in various situational contexts (Johnson & Valente, 2009).

The integration of role plays alongside alternative methodologies, such as interviews, can yield a holistic evaluation of a learner's oral proficiency by encompassing an extensive array of language functions (Johnson & Valente, 2009).

### 1.5.5 Interviews:

Interviews constitute a conventional technique for evaluating oral proficiency, facilitating an in-depth evaluation of a learner's linguistic capabilities. They afford learners the opportunity to exhibit their competence in responding to inquiries and participating in prolonged dialogues (Johnson & Valente, 2009). The efficacy of interviews can be augmented through the implementation of a systematic methodology that incorporates designated linguistic functions and modifies topic complexity in accordance with real-time evaluations (Johnson & Valente, 2009).

### **1.5.6 Story telling:**

The practice of storytelling signifies a dynamic and participatory educational approach that effectively involves learners and stimulates them to take an active role in the process of language learning. This method fosters a supportive atmosphere wherein students can practice their speaking skills without the concern of committing errors (Kim & Kellogg, 2007). The practice of storytelling significantly contributes to the enhancement of fluency and vocabulary by encouraging learners to utilize language in a creative manner and in various contexts. Furthermore, it facilitates the development of narrative competencies, which are essential for proficient communication (Isbell, Sobol, Lindauer, & Lowrance, 2004). Storytelling frequently incorporates cultural dimensions, which can aid learners in establishing an emotional connection with both the language and the associated content, thereby augmenting their capacity for self-expression (Kim & Kellogg, 2007).

### **1.5.7 Simulations:**

Simulations afford learners the opportunity to engage in linguistic practice within authentic contexts, thereby enhancing their capacity for spontaneous and contextually appropriate language use (Sato, 1985). Through simulations, learners are prompted to engage in problem-solving and critical thinking, both of which are imperative for the advancement of oral proficiency. Such activities necessitate that learners express their ideas with clarity and persuasiveness. Simulations frequently incorporate components of feedback and reflection, enabling learners to evaluate their performance and identify areas for enhancement (Sato, 1985).

### **1.5.8 Picture description:**

The act of describing pictures engages visual stimuli to facilitate language generation, thereby assisting learners in structuring their cognitive processes and articulating their

## Chapter one: The Literature review

---

thoughts with greater clarity (Ellis & He, 1999). The task of articulating descriptions of images necessitates that learners employ precise vocabulary, which has the potential to augment their lexical repertoire and enhance their capacity to describe objects, actions, and scenes (Newton, 2013). Tasks centered on picture description can serve as assessment instruments for various dimensions of oral proficiency, encompassing fluency, accuracy, and pronunciation. Such tasks offer a systematic approach to evaluating the speaking competencies of learners (Ellis & He, 1999).

While these methodologies demonstrate efficacy in the evaluation of oral proficiency, obstacles such as speech-related anxiety, insufficient encouragement, and an emphasis on accuracy rather than fluency may impede the advancement of speaking competencies. It is paramount for educators to balance these dimensions and foster a supportive environment that promotes risk-taking and engagement in speaking activities. Furthermore, the incorporation of technology and innovative assessment methodologies can significantly enhance the reliability and validity of oral proficiency evaluations within EFL contexts.

### **1.6 Framework of oral proficiency assessments:**

The framework related to oral proficiency assessments constitutes a sophisticated model that calls for detailed analysis of both validity and reliability to support an exact evaluation of linguistic proficiency. Validity involves the measure of how faithfully an assessment mirrors the idea it aims to quantify, while reliability deals with the stability and consistency of results extracted from the assessment throughout various timeframes and evaluations. Numerous empirical research efforts have investigated these dimensions across a range of contexts, offering important perspectives on the efficiency and barriers linked to oral proficiency assessments.

### 1.6.1 Validity definition:

The concept of validity within the context of oral proficiency assessment is complex, encompassing multiple dimensions that ensure a test accurately measures what it is intended to measure. Validity does not signify an inherent property of a test; instead, it is shaped by the interpretations and applications associated with the resultant test scores. In the domain of oral proficiency, the significance of validity is paramount in ensuring that assessments are both reliable and meaningful, thereby offering accurate representations of an individual's capabilities. The principal categories of validity encompass construct validity, content validity, criterion-related validity, face validity, and consequential validity, with each type contributing distinctively to the comprehensive validity framework.

#### 1.6.1.1 Construct validity:

Construct validity is about how successfully an assessment tool measures the theoretical concept that it aims to measure. In the domain of oral proficiency assessments, this requires the confirmation that the evaluation genuinely reflects the diverse components of speaking competence, encompassing fluency, coherence, and interactional proficiency.

The construct validity is crucial for determining that oral proficiency assessments accurately measure the specified linguistic competencies. For example, the investigation concerning the ACTFL (American Council on the Teaching of Foreign Languages) Proficiency Guidelines provided substantial evidence for the construct validity of their oral interview methodology, signifying that the guidelines serve as a reliable framework for proficiency evaluations (Dandonoli & Henning, 1990).

In the context of the TOEFL iBT speaking tasks, construct validity was evaluated through the application of natural language processing tools, which supported the employment of a variety of speaking tasks to assess speaking proficiency, thereby emphasizing the significance

of task diversity in capturing distinct linguistic features (Kyle, A. Crossley, & McNamara, 2016).

Despite its importance, the achievement of construct validity poses significant challenges. The investigation pertaining to oral proficiency scales uncovered a halo effect, which can diminish the reliable variance within scales evaluated during a single hearing, thereby highlighting potential concerns regarding specificity and agreement among raters (Yorozuya & Oller, 1980).

In order to enhance construct validity, it is advisable to employ authentic tasks that accurately represent real-world language application. The investigation into communicative competence within an EFL curriculum underscored the necessity for assessments that facilitate authentic interaction and the processing of unpredictable data in real-time contexts (Tseng, 2009).

Collaborative alliances with fields such as cognitive and computational neuroscience, in conjunction with artificial intelligence, are recommended to tackle the demands of language assessment in the 21st century, thereby potentially enhancing the construct validity of assessments measuring oral proficiency (Aryadoust, 2023).

### **1.6.1.2 Content validity:**

Content validity evaluates the extent to which an assessment instrument accurately reflects the content domain it claims to address. This procedure requires the engagement of subject matter experts to ensure that all relevant aspects of oral proficiency are integrated within the assessment (Messick, 1995).

Within the field of language evaluations, content validity is typically assessed through expert analysis and its compliance to established language proficiency criteria (Alderson, Clapham, & Wall, 1995).

## Chapter one: The Literature review

---

Content validity guarantees that the items within the assessment are representative of the specific language competencies under evaluation, which is imperative for the accuracy of oral proficiency evaluations. This process necessitates the involvement of subject matter experts to validate the representativeness and relevance of the test material, as illustrated in the study of Spanish as a Foreign Language placement assessment. This study clarifies the concept of content validity and describes the methodology utilized to validate the content of a speaking placement assessment in Spanish as a Foreign Language through expert analysis employing the focus group technique. The outcomes underscore the critical role of the proposed methodology in fostering agreement among experts and in confirming the representativeness and relevance of the test material (Alonso, 2017).

The lack of content validity adversely influences the effectiveness of teaching methodologies within the field of English language teaching. Assessment practices in education create substantial pressure for both teachers and pupils; however, these practices receive relatively lower prioritization from educators compared to their other instructional obligations. This challenge arises from the uncertainties surrounding the discourse on this subject. Assessments ought to concentrate on fundamental elements of the curriculum to evaluate the extent of learner accomplishment. In the absence of content validity in examinations, the predominant focus of instructional efforts tends to shift towards examination strategies instead of engaging in meaningful and real teaching. Furthermore, it was observed that learners did not utilize their textbooks as vital sources of knowledge, instead focusing their efforts predominantly on acquiring grades. The consequences of these educational practices culminated in the failure to meet the pedagogical objectives of language education (Siddiek, 2010).

The definition of the testing domain and the choice of relevant linguistic tasks involve sophisticated methodologies that demand detailed analysis and the engagement of subject matter experts (Alonso, 2017).

Assessments characterized by elevated content validity foster pedagogical practices that are congruent with the fundamental curriculum, thereby facilitating authentic language acquisition as opposed to mere rote memorization for evaluative purposes (Siddiek, 2010).

### **1.6.1.3 Criterion- related validity:**

Criterion validity in the assessment of oral proficiency represents a fundamental aspect that guarantees the accuracy and reliability of linguistic evaluations. This process entails the comparison of test scores with an external criterion, which may include an alternative validated assessment or a measure of performance in real-world scenarios, to evaluate the extent to which the test accurately predicts or correlates with the criterion. Such a form of validity is imperative for confirming that the assessment effectively measure what is intended to measure, especially within the realm of oral language proficiency. The impacts of criterion validity on linguistic evaluation are profound, as it affects the legitimacy and applicability of the test outcomes within both educational and professional domains. Numerous critical elements and implications associated with criterion validity exist in relation to oral proficiency assessment.

The assessment of criterion validity often requires an analysis of the relationships between test scores and those obtained from established evaluations. For instance, the Duolingo English Test demonstrated criterion validity through significant correlations with the TOEFL and IELTS scores, thereby confirming that the DET functions as a reliable instrument for assessing oral proficiency among non-native speakers.

## Chapter one: The Literature review

---

The DET is carefully designed to evaluate integrated English competencies in reading, writing, listening, and speaking, in alignment with the Common European Framework of Reference (CEFR). The CEFR serves as an internationally recognized standard for the assessment of individual language proficiency, employing a cumulative score to represent the test taker's overall linguistic capabilities. The framework encompasses six levels, ranging from A1 for novices to C2 for proficient learners, thereby providing test takers with clarity regarding the score they must attain. Overall, given that the DET can explain test takers' language proficiency effectively, and can be compared with other reputable language assessments, it may be regarded as a suitable language evaluation tool for the public, adhering to professional assessment standards (Chen, 2024).

An alternative approach to determining criterion validity involves the use of real world performance measures. This involves the comparison of assessment outcomes with authentic language usage in professional or social contexts. For instance, the employment of interviews in evaluating oral proficiency can yield valuable insights into an individual's capacity for effective communication in real life scenarios, thereby reinforcing the criterion validity of the assessment (Rosser, 1979).

Criterion validity may also be evaluated utilizing multi-component analysis, wherein distinct dimensions of linguistic proficiency, including fluency, pronunciation, and grammatical accuracy, are examined both in isolation and in conjunction. This methodology facilitates an understanding of the extent to which each component enhances the comprehensive validity of the assessment (Henning, 1983).

Criterion validity enhances the accuracy and reliability of linguistic evaluations by verifying that the assessment appropriately measures the intended construct. This is

## Chapter one: The Literature review

---

particularly significant in scenarios involving high-stakes assessments, where decisions about language proficiency may have considerable consequences (Thu, 2019).

Comprehending criterion validity can significantly contribute to the evolution and enhancement of linguistic assessments. Through the identification of specific elements within a test that exhibit strong correlations with external standards, test developers are enabled to concentrate on improving those areas to enhance the overall validity of the evaluation (González-Such, Meliá, & Bakieva, 2013).

Criterion validity possesses significant implications within educational and professional contexts, where accurate assessments of oral proficiency are essential for the purposes of placement, certification, and employment decisions. Evaluations demonstrating strong criterion validity produce reliable data that can guide these essential decisions (Rubin & Lim, 2006).

### **1.6.1.4 Face validity:**

Face validity in oral proficiency assessment pertains to the degree to which a test appears to be effective concerning its declared objectives, especially from the viewpoint of the test-takers. It constitutes an essential component of test validity, as it significantly impacts the acceptance and perceived credibility of the assessment by individuals undertaking it. When evaluating language, face validity is particularly important as it affects how motivated and engaged test-takers feel, which can later change how reliable the assessment results are.

Face validity is often defined as the degree to which an assessment seemingly measures the constructs it claims to evaluate, as perceived by the examinees and other pertinent stakeholders. It includes the clarity, relevance, and appropriateness of assessment items as perceived by the intended audience (Mosier, 1947).

## Chapter one: The Literature review

---

Assessment items should demonstrate clarity and be easily understandable to the test-takers. This underscores the importance of using language and formats that are familiar to the intended audience (Mosier, 1947).

The content contained within the test must maintain a considerable relevance with the lived experiences of the test-takers, as well as the contextual framework under which the test takes place (Bachman & Palmer, 1996).

The test must demonstrate a correspondence to language utilization in the real world and be regarded as authentic by the test-takers undertaking the test (Downing, 2006).

Test items should be designed in a manner that is both non-judgmental and non-intrusive, as any negative perceptions may significantly hinder the performance of the test-takers (Mosier, 1947).

### **1.6.1.4.1 Influence on reliability of language evaluation:**

An increase in face validity has the potential to significantly improve the motivation and engagement levels of individuals taking assessments, thereby supporting the realization of more accurate and reliable assessment findings. When individuals undergoing evaluation consider a test as relevant and fair, they are likely to operate at the highest level of their competence (Mosier, 1947).

Assessments marked by significant face validity are presumed to produce data that reflects greater accuracy, given that those taking these tests are better equipped to comprehend and relate to the questions asked. Such accuracy is crucial for reliable language evaluation, as it secures that the assessment findings truly reflect the real linguistic capabilities of the candidates (Mosier, 1947).

## Chapter one: The Literature review

---

The concept of face validity plays a significant role in fostering the acceptance of the assessment by both test-takers and relevant stakeholders. A test that is deemed valid is more likely to receive acceptance and trust, which are essential for its broad implementation and credibility (Bachman & Palmer, 1996).

Critics argue that face validity constitutes the least rigorous form of validity due to its dependence on subjective assessments rather than on empirical evidence. Nevertheless, within the realm of oral proficiency evaluations, where the engagement and motivation of test-takers are paramount, face validity assumes a vital function in guaranteeing the efficacy and reliability of the assessment process (Downing, 2006).

### **1.6.1.5 Impact (consequential) validity:**

The concept of impact, or consequential validity, within the sphere of oral proficiency assessment represents a crucial component in evaluating the effectiveness and implications of language assessments. The idea of consequential validity involves examining how the results of an assessment correspond to its intended objectives and the larger consequences for different parties, including learners, instructors, and educational systems. This principle is vital in confirming that assessments not only measure what they intend to but also encourage positive impacts on educational strategies and outcomes.

Consequential validity is characterized as the extent to which the results derived from an assessment correspond with its designated objectives and the wider effects it exerts on various stakeholders (Royal & Puffer, 2014).

It is considered a fundamental aspect of validity, as it encompasses the implications of test application and interpretation for both individuals and larger systems.

## Chapter one: The Literature review

---

The evaluation of consequences is essential to the test validation process, ensuring that both expected positive outcomes and possible unforeseen negative impacts are appropriately recognized (Mehrens, 2005).

Consequential validity examines the impacts of assessments on various stakeholders, including students, educators, and educational institutions. For students, assessments can significantly affect learning outcomes, motivational levels, and self-perception. For example, diagnostic feedback derived from oral proficiency evaluations can assist students in recognizing their strengths and weaknesses, consequently directing their learning process (Zhao, 2013).

Educators and educational institutions are similarly influenced, given that evaluative measures can significantly shape pedagogical methodologies and curriculum design. The implementation of interactive task formats within oral proficiency evaluations, for instance, incorporates a social aspect that can influence both the design of assessments and the criteria for scoring (McNamara & Roever, 2006).

One of the significant challenges in evaluating consequential validity is the possibility of unintended consequences, including heightened levels of anxiety or stress experienced by individuals undertaking the assessment (Kiani, Alibakhshi, & Akbari, 2009).

The socio-cultural aspect of oral proficiency evaluations, in which performance is shaped by engagements with peers, presents significant concerns regarding the validity and reliability of the assessments (McNamara & Roever, 2006).

Evaluating the impact of assessments requires the employment of both quantitative and qualitative research methods to capture the entire range of outcomes (Cook & Lineberry, 2016).

## Chapter one: The Literature review

---

Frameworks designed for the evaluation of consequential validity necessitate a thorough analysis of both the intended and unintended repercussions of assessments on diverse stakeholders. The collection and interpretation of evidence concerning consequential validity can be systematically arranged using frameworks that reflect the type, quantity, and methodological stringency of the requisite evidence, depending on the particular assessment and its intended use (Cook & Lineberry, 2016).

Qualitative studies, such as those investigating the impact of English for Specific Purposes (ESP) examinations in Iran, yield valuable perspectives on the psychological, socio-cultural, and financial consequences of assessments on both students and educators (Kiani, Alibakhshi, & Akbari, 2009).

While consequential validity constitutes an essential dimension of oral proficiency evaluation, it frequently remains underrepresented and restricted in scope. This underscores the necessity for more extensive assessments that take into account the wider repercussions of evaluations on various stakeholders and educational frameworks. Through the incorporation of consequential validity into the assessment paradigm, educators and policymakers can determine that examinations not only effectively measure proficiency but also yield beneficial contributions to educational outcomes and methodologies.

### 1.6.2 Reliability

Reliability in the assessment of oral proficiency refers to the consistency and accuracy with which oral linguistic competencies are measured. It constitutes a fundamental part in demonstrating that the outcomes of an assessment are sound and can be relied upon with certainty for discerning decision-making. The concept of reliability doesn't reflect an essential attribute of the test; instead, it relates to the results obtained from it. This notion possesses considerable significance within the domain of language evaluation, where the variability

## Chapter one: The Literature review

---

intrinsic to human performance and subjective assessment may result in inconsistencies. The measurement of reliability involves a variety of statistical techniques and metrics that are vital for clarifying and improving the consistency of oral proficiency assessments.

Reliability is characterized as the consistency of scores across various replications, which may encompass divergent items, forms, raters, or temporal contexts. This consistency is essential for guaranteeing that the assessment outcomes remain stable and are not subject to the influence of random errors (Bachman L. F., 2004).

Inter-rater Reliability constitutes a fundamental element in oral assessments, where multiple raters evaluate the same performance. Increased inter-rater reliability indicates that various evaluators provide consistent scores, which is essential for the credibility of the assessment. For example, The ACTFL Oral Proficiency Interview (OPI) demonstrated remarkably high inter-rater reliability across 19 languages, underscoring the importance of a well- designed interview framework and thorough rater training (Surface & C.Dierdorff, 2003).

Test-retest Reliability evaluates the consistency of scores across different time intervals. In the context of oral proficiency assessments, it is essential that scores demonstrate consistency across various testing instances, presuming there are no substantial changes in the test-takers' capabilities. This reliability is frequently analyzed by utilizing parallel forms of the assessment over a short time period to lessen practice effects and to establish that any observed score changes are due to measurement error rather than real variations in proficiency (Daniel, 2010).

Internal Consistency pertains to the degree to which various components of the assessment evaluate the same underlying construct. A research study was conducted to

## Chapter one: The Literature review

---

examine the reliability and validity of 30 oral presentations that occurred over a duration of four months, incorporating factors such as the students' gender, proficiency, and the subjects of the presentations. It is frequently quantified utilizing Cronbach's Alpha, with coefficients ranging from 0 to 1. Elevated internal consistency suggests that the assessment items are effectively measuring the same fundamental skill. In an investigation of oral presentations, reliability coefficients varied from 0.70 to 0.95, signifying a high level of internal consistency (Brown & Hudson, 2002).

Test Validity serves as a significant indicator that guarantees the assessment accurately quantifies the intended constructs it seeks to evaluate. Oral proficiency evaluations must account for a broad range of language skills, which should involve fluency, pronunciation quality, and grammatical accuracy. The use of various assessment methods, like interviews, imitation, and completion tasks, has the capability to elevate the credibility of the validity by enabling a more comprehensive assessment of oral abilities (Henning, 1983). Furthermore; the incorporation of authentic, interactive activities has the potential to enhance the construct validity of the assessments. The evaluation of oral proficiency has experienced an increasing inclination towards interactive task formats as the framework within which to assess a learner's spoken capabilities. Tasks that require learners to engage in authentic and meaningful interactions with their peers or their teacher (McNamara & Roever, 2006).

The involvement of an interlocutor encompasses a social dimension that has the potential to influence assessment outcomes. This interaction may result in variability in performance, as the dynamics present between the test-taker and the interlocutor could significantly affect the results. It is imperative to construct tasks that reduce the influence of

these social factors and concentrate on the linguistic competence of the test-taker (McNamara & Roever, 2006).

### **1.6.3 Challenges in ensuring reliability in oral proficiency assessments:**

Factors including the specific occasion, the individual conducting the interview and the subject matter can substantially influence test scores. Empirical studies suggest that variations in speech styles and topics may result in variability in scores, thereby underscoring the necessity for standardized testing conditions to guarantee reliability (Shohamy, 1983).

Evaluative measures of oral proficiency obligate assessors to conduct subjective evaluations of language use, thereby potentially resulting in inconsistencies. The creation of specific and clear scoring rubrics, alongside the implementation of technological instruments to record and assess performances, may help to reduce these difficulties (Fulcher, 2003).

The assurance of fairness in oral assessments is critical, especially when assessing heterogeneous populations. Bias may manifest from variations in culture or the fundamentally subjective aspects of the assessment process. The application of comprehensive training and calibration for raters can significantly minimize bias and enhance the principle of fairness (Memon, Joughin, & Memon, 2010).

While the challenges associated with ensuring the reliability of oral proficiency assessments are considerable, innovations in technology and assessment methodologies present viable solutions. The implementation of digital recording and analytical tools has the potential to augment the objectivity and consistency of evaluative processes. Furthermore, continuous research and development within this domain are imperative to address the intricacies of oral assessments and to enhance their reliability and validity.

### **1.6.4 Examples of studies emphasizing validity and reliability in oral proficiency assessments:**

The importance of validity and reliability in the assessment of oral proficiency is highlighted by several of distinguished studies, each illuminating various aspects of these critical evaluation attributes. Validity guarantees that the assessment accurately evaluates the intended construct, whereas reliability assures the consistency of outcomes across diverse administrations and raters. Collectively, these studies underscore the necessity for rigorous assessment methodologies to effectively measure oral proficiency, which is vital in both educational and professional contexts.

#### **1.6.4.1 Pair Interaction and Task Design:**

The investigation conducted by Lynette May concerning English for Academic Purposes (EAP) programs highlights the significant impact of task design on the validity and reliability of assessments related to oral proficiency. This study contrasts oral proficiency interviews with pair interactions, illustrating that the characteristics of the task can significantly influence the dynamics of interaction and the resulting assessment outcomes, consequently impacting the validity and reliability of the findings (May, 2000).

#### **1.6.4.2 Self, Peer, and Teacher Assessments:**

The research conducted by Brown and Hudson examines the reliability and validity of self, peer, and teacher assessments in the context of oral presentations. The findings of the study indicated high reliability indices (Cronbach's Alpha values ranging from 0.70 to 0.95) and observed that the aspect of validity experienced improvement over time. This implies that the implementation of consistent assessment methodologies has the potential to enhance the reliability and validity of evaluations concerning oral proficiency (Brown & Hudson, 2002).

### ***1.6.4.3 Elicited Imitation Task:***

The investigation conducted by Harriet Wood Bowden regarding the Spanish Elicited Imitation Task (EIT) illustrates its effectiveness in reliably evaluating L2 oral proficiency. The EIT showed considerable external validity when examined in conjunction with simulated oral proficiency interviews, thereby reinforcing its application as a standardized instrument in the realm of second language acquisition research (Bowden, 2016).

### ***1.6.4.4 Oral Proficiency Interview (OPI):***

The examination conducted by Margaret E. Malone regarding the ACTFL Oral Proficiency Interview (OPI) highlights the persistent discourse surrounding its validity and reliability. In spite of the criticisms leveled against it, the OPI continues to be a prevalent instrument, with scholarly investigations concentrating on enhancing its inter-rater reliability and addressing the concerns related to its validity (Malone, 2003).

### ***1.6.4.5 Self-Assessment Accuracy:***

The thesis authored by Angels Piñana investigates the accuracy of self-assessment in oral proficiency, emphasizing the significance of proficiency level and implicit training. The research indicated that learners with higher proficiency levels are inclined to produce more accurate self-assessments; however, biases and subjective perceptions may compromise the reliability of these evaluations (Piñana, 2009).

### ***1.6.4.6 Computerized Oral Proficiency Interview:***

The research of Dai Zhao-hui compares the standard Oral Proficiency Interview (OPI) with the Computerized Oral Proficiency Interview (COPI), indicating considerable reliability and validity across both assessment approaches. Nevertheless, the COPI is regarded as exhibiting a higher degree of objectivity yet is perceived as less interactive, a factor that may influence the performance and preferences of test-takers (Zhao-hui, 2011).

### ***1.6.4.7 Classroom-Based Feedback:***

The investigation conducted by Bijani, Hashempour, Khaled Ahmed Abdel Al Ibrahim, Orabah, & Heydar nejad regarding classroom-based feedback and the training of raters underscores the significance of systematic training in attaining inter-rater reliability. The findings of the study reveal that although training enhances consistency among raters, it does not eradicate inherent individual differences among them, thereby signifying the necessity for continuous training to maintain reliability (Bijani, Hashempour, Khaled Ahmed Abdel Al Ibrahim, Orabah, & Heydarnejad, 2022).

While these investigations confirm the significance of validity and reliability in oral proficiency assessments, they also reveal challenges and areas for enhancement. For example, the inherently subjective nature of oral assessments, along with the impact of task design and rater attributes, may complicate endeavors to attain reliable and valid outcomes. These complexities emphasize the imperative for ongoing research and the refinement of assessment instruments to guarantee their alignment with the high standards required for educational and professional evaluations.

### **1.7 Assessing the validity and reliability of oral proficiency assessments:**

Assessing the validity and reliability of oral proficiency assessments in the context of language acquisition is imperative for ensuring precise measurement of learners' linguistic competencies. A variety of methodologies have been investigated to enhance these assessments, each possessing distinct advantages and drawbacks. The incorporation of technological advancements, the implementation of diverse assessment techniques, and the consideration of socio-cultural influences represent fundamental strategies in enhancing the validity and reliability of oral proficiency assessments.

### **1.7.1 Technological Integration**

#### **1.7.1.1 Automatic Speech Recognition (ASR):**

The application of ASR technology within the realm of oral proficiency assessments has demonstrated considerable potential in augmenting both reliability and validity. ASR possesses the capability to deliver consistent scoring alongside immediate feedback, which serves as advantageous for learners. The integration of ASR with bilingualism-oriented methodologies has been observed to exert a favorable influence on learners' attitudes and the acquisition of oral proficiency (Karatay & Karatay, 2021).

#### **1.7.1.2 Computer-Assisted Oral Tests (CAOT):**

With Computer-Assisted Oral Tests (CAOT), there's a structured strategy for assessing oral abilities, creating a regulated atmosphere that minimizes human judgment. These assessments have been favorably compared to traditional in-person oral evaluations, revealing distinct advantages in terms of reliability and validity (Chapelle & Chung, 2010).

### **1.7.2 Multiple Assessment Methods**

#### **1.7.2.1 Diverse Testing Approaches:**

The utilization of an array of testing methodologies, including interviews, imitation tasks, and completion exercises, has the potential to significantly enhance the validity of assessments pertaining to oral proficiency. Among these methodologies, the imitation technique has been recognized as exhibiting the greatest overall validity, followed in rank by both the interview and completion methods (Henning, 1983).

#### **1.7.2.2 Self, Peer, and Teacher Assessments:**

The integration of self-assessment, peer evaluation, and teacher assessments can provide a holistic perspective on a learner's oral proficiency. These evaluative methods have

exhibited high reliability indices and possess the capability to accurately predict achievement scores (Brown & Hudson, 2002).

### **1.7.3 Statistical and Methodological Approaches**

#### **1.7.3.1 Statistical Analysis:**

Instruments such as Cronbach's alpha and Pearson r analysis are employed to assess the internal consistency and concurrent validity of oral proficiency evaluations. For example, the A Developmental English Language Proficiency Test (ADEPT) assessment exhibited moderate to high internal consistency along with diverse levels of concurrent validity observed across various academic years (Chavez, 2013).

#### **1.7.3.2 Frameworks and Standards:**

Compliance with international standards such as the Common European Framework of Reference (CEFR) and the American Council on the Teaching of Foreign Languages (ACTFL) guidelines guarantees that assessments possess both validity and reliability. These frameworks underscore the significance of authenticity, practicality, and socio-cultural factors in the formulation of assessment methodologies (Europe., 2020).

### **1.7.4 Socio-Cultural and Contextual Considerations**

#### **1.7.4.1 Cultural Sensitivity:**

A comprehensive assessment of oral proficiency necessitates the incorporation of cultural sensitivity as well as an awareness of the socio-cultural context of the learners. This encompasses the assessment of abilities in listening comprehension and oral expression, both of which are fundamental to the attainment of proficiency (González-Such, Meliá, & Bakieva, 2013).

## Chapter one: The Literature review

---

While these methodologies offer significant frameworks for evaluating oral proficiency, certain challenges persist. The adoption of technological resources, despite their benefits, requires careful attention to reduce possible technical biases. Additionally, we must take the socio-cultural influences on learners into consideration to secure fairness and equity in assessment strategies. Continuous professional development for evaluators remains essential to stay informed about emerging trends and challenges in the domain of language assessment.

### **1.8 Challenges in achieving validity and reliability in oral proficiency assessments:**

Achieving the validity and reliability of oral proficiency assessments in the context of language acquisition poses numerous challenges, as highlighted by various research studies. These challenges arise from the intrinsic complexity of language as an abstract construct, the varied contexts in which language is employed, and the subjective nature of oral assessments.

#### **1.8.1 Complexity of Language and Assessment Design:**

The inherent abstract nature of language renders its assessment a complex endeavor. This complexity is further intensified by the imperative for assessments to uphold both validity and reliability, which requires a distinct differentiation between the competencies being assessed and the methodologies utilized to elicit those competencies (Bachman & Palmer, 1996).

The design of oral proficiency assessments often merges the competencies under evaluation with the methodologies employed for elicitation, thus posing challenges in establishing validity. For instance, the ACTFL Oral Proficiency Interview has faced critique for its insufficient differentiation between these components, which subsequently undermines the validity of its evaluative ratings (Bachman, 1988).

### **1.8.2 Cultural and Contextual Factors:**

Cultural biases and the imperative for cross-linguistic validity present considerable challenges, particularly in assessments pertaining to bilingual children from minority-language communities. These assessments must be both culturally and cognitively appropriate to ensure equity and fairness (Sánchez, 2006).

The socio-cultural and contextual variables require thorough investigation to maintain the integrity of language assessments. This encompasses an appreciation of the varied backgrounds of test-takers and the circumstances in which they use the language (Europe., 2020).

### **1.8.3 Resource Limitations and Standardization:**

The insufficiency of resources constitutes a considerable barrier to the development and validation of effective language proficiency assessments. This phenomenon is particularly pronounced in the realm of English Language Learners (ELLs), where states encounter challenges in fulfilling legislative requirements for valid assessments due to limitations in available resources (Wolf, Farnsworth, & Herman, 2008).

Standardized assessments for high-incidence languages, including Spanish and French, are available; however, their applicability tends to diminish over time as linguistic attrition occurs within immersion settings. This situation necessitates the use of informal assessment measures, which may exhibit reduced reliability (McKellar & Unruh, 2017).

### **1.8.4 Reliability and Consistency:**

The assurance of reliability within oral proficiency assessments necessitates the involvement of multiple evaluators and the execution of repeated evaluations, which can be considerably demanding in terms of resources. The exploration of automatic speech

## Chapter one: The Literature review

---

recognition systems as a potential means of delivering objective scoring has been undertaken; however, their practicality and reliability are still subjects of ongoing research (Walt, Wet, & Niesler, 2008).

Assessments conducted through self-evaluation, peer evaluation, and teacher evaluation of oral presentations have demonstrated high indices of reliability, yet the validity of these assessments may fluctuate over time and across varying levels of proficiency. The self-assessment method, notably, is vulnerable to how learners perceive their abilities and levels of proficiency, which can significantly influence its overall correctness (Brown & Hudson, 2002).

While these challenges are significant, they also present opportunities for innovation and enhancement in language assessment practices. The integration of technology, continuous professional development for evaluators, and the development of culturally and contextually appropriate assessment tools are potential solutions for addressing these challenges. However, it is crucial to balance these innovations with the need for rigorous validation processes to ensure that assessments remain fair, reliable, and valid across diverse contexts and populations.

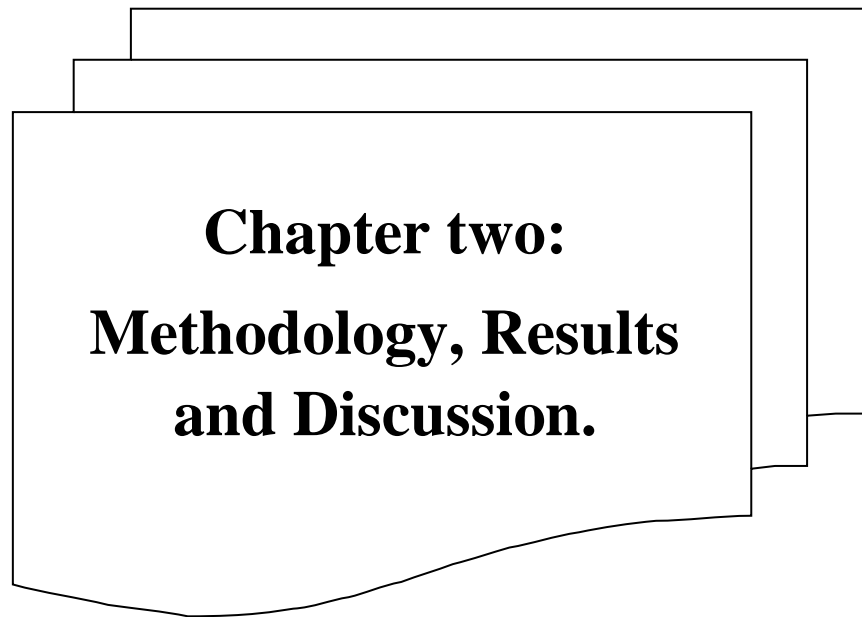
### **Conclusion**

The theoretical exploration in this chapter underscores the complexity of assessing oral proficiency in EFL learners, emphasizing the need for rigorous validation and reliability measures to ensure fair and accurate evaluations. Validity and reliability are not merely technical requirements but fundamental principles that determine the credibility and usefulness of assessment outcomes. The discussion has highlighted the importance of aligning assessment tasks with real-world communication demands while mitigating potential biases

## **Chapter one: The Literature review**

---

and inconsistencies. The insights drawn from this theoretical framework will inform the subsequent empirical investigation, guiding the analysis of oral proficiency assessment practices among secondary education students. Ultimately, this chapter sets the stage for a deeper examination of how current assessment methods can be refined to better capture and foster EFL learners' communicative abilities.



**Chapter two:  
Methodology, Results  
and Discussion.**

## 2 Chapter two: the practical chapter.

### Introduction

Chapter one addressed the theoretical background and relevant studies on the oral proficiency assessment techniques, laying the foundation for the next chapter. Therefore, this chapter focuses on the fieldwork. It presents the practical aspect of the research. This chapter is devoted to offer a detailed explanation of the research methodology and design, including the research methods, setting, population, and instruments, along with the data collection and analysis tools. It also centers on the analysis and interpretation of the data, as well as a discussion of the results.

### 2.1 Research Methodology and Design.

This section explains the research methodology and outlines the steps taken to collect, analyze, discuss, and interpret data. It also details the methods and approaches used to achieve the research aims and answer the research questions.

#### 2.1.1 The Research Method

This study employed a mixed-methods approach, combining quantitative and qualitative data collection techniques to comprehensively investigate oral proficiency assessment practices from both student and teacher perspectives. This method was chosen because it allows for data triangulation, enhancing the validity of results by cross-verifying findings from different sources. It provides a broader and deeper understanding of the research problem than a single-method approach.

### 2.1.2 The population and study sample:

The study population comprised secondary school students and teachers of English and French at Mouloud Quassem Secondary School in Mezloug, Setif, during the 2024/2025 academic year. This encompassed 758 students organized into twenty-two groups and a census of ten language teachers. For student sampling, a convenience sampling technique was employed. The researcher selected five student groups based on their affiliation as the instructor of those groups. This choice facilitates access. From these five groups, a total of 170 students, representing 22.42% of the total student population, constituted the primary study sample. Regarding teacher sampling, a census approach was utilized. All ten English and French teachers at the school were included in the study sample.

### 2.1.3 The study setting:

The research was situated at Mouloud Quassem Secondary School in Mezloug, Setif, Algeria, during the 2024/2025 academic year. This single-site setting provided access to the target population of secondary school students and language teachers within the Algerian public education system. The researcher's position as an instructor within the school facilitated operational access to participants.

### 2.1.4 Data Collection Instruments

#### 2.1.4.1 Student Questionnaire:

A structured Likert-scale questionnaire was administered to students, comprising four sections:

**The first section: Oral Proficiency Components:** Focused on self-assessment of fluency, grammar, pronunciation, vocabulary, and confidence.

**The second section: Teaching and Learning Methods:** Explored the effectiveness of activities like role-playing, media use, and group discussions.

**The third section: Assessment Methods:** Evaluated perceptions of oral presentations, exams, and informal assessments.

**The fourth section: Challenges and Suggestions:** Identified barriers to oral proficiency and student recommendations for improvement.

### 2.1.4.2 Teacher Questionnaire:

The questionnaire was designed for teachers, divided into four sections:

**The first section: Demographic Information:** Collected background data on gender, experience, class size, and languages taught.

**The second section: Current Assessment Practices:** Used a Likert scale to gauge teachers' agreement on assessment frequency, methods, and institutional support.

**The third section: Challenges and Strategies:** Included multiple-choice and open-ended questions to identify obstacles and effective teaching strategies.

**The fourth section: Support and Recommendations:** Sought feedback on additional resources or training needed for better assessment practices.

### 2.1.5 Procedure

Both questionnaires were distributed electronically and in paper format to ensure accessibility. Students and teachers were given sufficient time to complete the surveys. Participants were assured of anonymity and confidentiality, with responses used solely for research purposes.

### 2.1.6 Data Analysis tools

**Quantitative Analysis:** Likert-scale responses from both questionnaires were analyzed using statistical tools to calculate frequencies, means, and standard deviations. This helped identify trends and patterns in perceptions and practices.

**Qualitative Analysis:** Open-ended responses and additional comments were thematically coded to extract recurring themes, challenges, and suggestions. This provided deeper insights into subjective experiences and recommendations.

## 2.2 Data analysis and discussion

### 2.2.1 Analysis of the teachers' questionnaire data

#### 2.2.1.1 Analysis and discussion of questionnaire reliability and validity

**Table 01: Reliability Statistics based on Cronbach's Alpha.**

Cronbach's Alpha	N of Items
0,744	13

The reliability statistics indicate a Cronbach's Alpha value of 0.744 for the 13 items in the questionnaire. This value suggests a good level of internal consistency among the items, meaning that the questions are measuring a coherent concept related to oral proficiency assessment. Generally, a Cronbach's Alpha above 0.7 is considered acceptable in social sciences, indicating that the questionnaire is reliable enough for research purposes. Therefore, we can trust that the responses collected from these 13 items provide consistent and dependable data.

**Table 02: Validity statistics based on Pearson Correlation**

<b>Pearson Correlation</b>	<b>N</b>	<b>Sig. (2-tailed)</b>
0.857	10	0.002

To ensure the validity of the scale measuring teachers' perceptions of students' oral proficiency. The results show a **strong positive correlation** with a Pearson coefficient of  $r = 0.857$ , based on  $N = 10$  cases. The **significance level is  $p = 0.002$** , which is well below the commonly accepted threshold of 0.05. This high correlation indicates that the items are **strongly related** to the overall construct and that the scale has **good construct validity**. In other words, the items are measuring what they are intended to measure "teachers' perceptions of oral proficiency" consistently and reliably. Hence, the scale used in this section can be considered **valid**, and the results derived from it are credible and can be interpreted with confidence.

### 2.2.1.2 Analysis of the first questionnaire section: Demographic Information

**Table 03: Gender of the teachers**

<b>Result</b>	<b>Frequency</b>	<b>Percentage</b>
<b>Male</b>	1	10
<b>Female</b>	9	90
<b>Total</b>	10	100

## Chapter two: Methodology, Results and Discussion

---

The gender distribution of the respondents shows a significant predominance of female teachers, representing 90% of the sample, while male teachers make up only 10%. This imbalance suggests that the majority of the participants in this questionnaire are female.

**Table 04: Years of Teaching Experience**

Result	Frequency	Percentage
Less than 1 year	0	0
1-5 years	5	50
6-10 years	3	30
More than 10 years	2	20
Total	10	100

The distribution of teaching experience among the respondents reveals that half of the teachers (**50%**) have between 1 to 5 years of experience. Additionally, **30%** have between 6 to 10 years, while **20%** possess more than 10 years of teaching experience. No teachers reported having less than one year of experience. This indicates a relatively balanced spread, with the majority being relatively early to mid-career educators, and a smaller portion being very experienced. Such a mix can provide diverse insights, combining fresh perspectives with seasoned expertise regarding oral proficiency assessment.

## Chapter two: Methodology, Results and Discussion

**Table 05: The average class size in your oral lessons**

<b>Result</b>	<b>Frequency</b>	<b>Percentage</b>
<b>Fewer than 20 students.</b>	0	0
<b>20 - 30 students.</b>	0	0
<b>30 - 40 students.</b>	5	50
<b>More than 40 students.</b>	5	50
<b>Total</b>	10	100

The data on average class size for oral lessons indicates that half of the teachers (50%) conduct classes with 30 to 40 students, while the other half manage even larger classes of more than 40 students. Notably, there are no classes with fewer than 30 students. This suggests that oral lessons are generally held in relatively large groups, which could present challenges in providing individual attention and thorough assessment of each student's oral proficiency. The large class sizes may impact the effectiveness of oral proficiency evaluations and require strategies tailored to managing many students at once.

**Table 06: Language(s) of Teaching**

<b>Result</b>	<b>Frequency</b>	<b>Percentage</b>
English	5	50
French	5	50

## Chapter two: Methodology, Results and Discussion

<b>Total</b>	10	100
--------------	----	-----

The responses show an equal split between teachers who teach English and those who teach French, each group representing **50%** of the sample. This balance indicates that the questionnaire reflects perspectives from educators in both language areas equally, which can provide a broader understanding of oral proficiency assessment across different languages. It also suggests that findings from the study may be relevant to teaching both English and French oral skills.

### 2.2.1.3 Analysis of the second questionnaire section: Current Assessment Practices (Likert Scale)

Please rate the following statements on a scale of 1 (**Strongly Disagree**) to 5 (**Strongly Agree**).

**Table 07: Teachers' perceptions of current assessment practices.**

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
5. I regularly assess students' oral proficiency in my classes.	0	3	0	7	0
6. My institution provides clear guidelines for oral proficiency assessment.	4	3	2	1	0
7. I use a variety of assessment methods (e.g., interviews, role-plays, presentations).	0	1	1	6	2
8. I believe oral proficiency assessments accurately measure students' speaking skills.	0	1	1	7	1
9. I receive adequate training on how to conduct oral assessments.	3	2	2	3	0
10. I provide feedback to students after oral assessments.	0	2	0	6	2
11. Students feel anxious during oral assessments.	0	1	1	1	7

## Chapter two: Methodology, Results and Discussion

12. Technology (e.g., speech recognition software) improves oral assessment reliability.	0	1	0	7	2
13. Cultural factors influence students' performance in oral assessments.	0	0	1	6	3
14. Peer assessment is an effective method for evaluating oral proficiency.	1	2	1	5	1

**Table 08: Descriptive statistics of teachers' responses on current assessment practices.**

Statement	Mean	Standard Deviation	Result
5. I regularly assess students' oral proficiency in my classes.	3,4000	0,96609	Agree
6. My institution provides clear guidelines for oral proficiency assessment.	2,0000	1,05409	Disagree
7. I use a variety of assessment methods (e.g., interviews, role-plays, presentations).	3,9000	0,87560	Agree
8. I believe oral proficiency assessments accurately measure students' speaking skills.	3,8000	0,78881	Agree
9. I receive adequate training on how to conduct oral assessments.	2,5000	1,26930	Disagree
10. I provide feedback to students after oral assessments.	3,8000	1,03280	Agree
11. Students feel anxious during oral assessments.	4,4000	1,07497	Strongly Agree
12. Technology (e.g., speech recognition software) improves oral assessment reliability.	4,0000	0,81650	Agree
13. Cultural factors influence students' performance in oral assessments.	4,2000	0,63246	Strongly Agree
14. Peer assessment is an effective method for evaluating oral proficiency.	3,3000	1,25167	Neutral

## Chapter two: Methodology, Results and Discussion

---

The results reveal that teachers generally have a positive attitude toward assessing oral proficiency. For instance, item 5 shows a mean of **3.4**, indicating that most teachers **agree** they regularly assess oral proficiency in their classes. This suggests that speaking skills are considered an integral part of their teaching practice.

However, item 6 has a notably low mean of **2.0**, with teachers **disagreeing** that their institutions provide clear guidelines for oral proficiency assessment. This may reflect a lack of institutional support or standardized frameworks, which could impact the consistency of assessment across classrooms.

In item 7, the high mean of **3.9** shows that teachers **agree** they use a variety of assessment methods, such as interviews and presentations. This variety is essential in evaluating different aspects of oral communication and indicates that teachers are using diversified tools.

Similarly, item 8 (mean **3.8**) and item 10 (mean **3.8**) show agreement that oral assessments accurately measure speaking skills and that teachers provide feedback afterwards—two essential components for effective assessment and student improvement.

Item 9, however, stands out with a mean of **2.5**, which falls under **disagree**. This highlights that teachers feel they **do not receive sufficient training** on how to conduct oral assessments, which may hinder their ability to apply best practices.

When it comes to student experience, item 11 shows a very high mean of **4.4**, indicating that teachers **strongly agree** students often feel **anxious** during oral assessments. This suggests the emotional impact of these assessments is significant and should be taken into account during planning and implementation.

## Chapter two: Methodology, Results and Discussion

Technology is seen positively in item 12, with a mean of **4.0**, showing that teachers **agree** on its value in improving reliability. Likewise, item 13 (mean **4.2**) reveals **strong agreement** that **cultural factors influence student performance**, which is a critical observation in multilingual or multicultural classrooms.

Finally, item 14 presents a mean of **3.3**, which reflects a **neutral** stance toward the effectiveness of **peer assessment** in evaluating oral proficiency. This may indicate uncertainty or mixed experiences with peer-based methods.

### 2.2.1.4 Analysis of the third questionnaire section: Challenges and Strategies.

#### 15. What challenges do you face when assessing students' oral proficiency?

*(Select all that apply or add your own)*

**Table 09: Descriptive results of teachers' challenges in assessing oral proficiency.**

Result	Frequency	Percentage	Percent of case
Lack of time for individual assessments	8	16.8	80
Student anxiety or nervousness	9	20.9	90
Difficulty in maintaining objectivity	2	4.7	20
Limited resources (e.g., recording devices, software)	6	14	60
Large class sizes	7	16.3	70
Varying proficiency levels in the same class	7	16.3	70
Lack of standardized rubrics	3	7	30
Other	1	2.3	10

## Chapter two: Methodology, Results and Discussion

---

<b>Total</b>	43	100	-
--------------	----	-----	---

The data illustrates a wide range of challenges that teachers face when assessing students' oral proficiency, with the most prominent being **student anxiety or nervousness**, reported by **90%** of participants. This highlights a psychological barrier that can significantly affect learners' performance and suggests a need for assessment environments that are more supportive and less intimidating.

The **lack of time for individual assessments** comes second, affecting **80%** of respondents. This is likely due to heavy teaching loads and large class sizes, which limit opportunities for personalized assessment.

Indeed, **large class sizes** and **varying proficiency levels** were both selected by **70%** of teachers, reinforcing the idea that classroom diversity and overcrowding hinder effective oral evaluation. These structural issues make it harder to adapt assessments to individual student needs.

**Limited resources**, such as access to technology or recording tools, were also mentioned by **60%** of teachers, pointing to logistical and infrastructural constraints in schools.

Other challenges, though less frequently cited, are still important. For instance, **30%** indicated a **lack of standardized rubrics**, which could lead to inconsistency in grading. Meanwhile, **20%** mentioned **difficulty in maintaining objectivity**, suggesting some teachers struggle to remain unbiased in subjective evaluations.

One teacher added a personal observation: "**The insufficient academic level of the learners**" ,reported by **10%** of participants ,which reflects concern about learners' general preparedness and could explain difficulty in performing oral tasks.

### 16. How do you ensure fairness and objectivity in oral assessments?

**Table10: Descriptive results on how ensuring fairness and objectivity in oral assessment.**

<b>Result</b>	<b>Frequency</b>	<b>Percentage</b>	<b>Percent of case</b>
Using standardized rubrics/scoring criteria	4	26.7	40
Recording assessments for later review	1	6.7	10
Multiple assessors (peer/co-teacher evaluations)	4	26.7	40
Blind grading (assessing without knowing student identities)	0	0	0
Providing clear instructions and practice beforehand	6	40	60
Other	0	0	0
<b>Total</b>	15	100	-

The strategies used by teachers to ensure fairness and objectivity in oral assessments show a preference for clear communication and standardized procedures. The most common method, selected by **60%** of teachers, is **providing clear instructions and practice beforehand**. This approach likely helps students feel more prepared and reduces ambiguity during the evaluation process, which contributes to a sense of fairness.

**Standardized rubrics or scoring criteria** and **multiple assessors (such as peer or co-teacher evaluations)** were each chosen by **40%** of respondents. These strategies play a crucial role in minimizing subjectivity and enhancing reliability in scoring. The use of rubrics ensures all students are judged against the same benchmarks, while involving multiple assessors adds a layer of objectivity.

## Chapter two: Methodology, Results and Discussion

Interestingly, only **10%** of teachers reported **recording assessments for later review**, which can be a valuable tool for reassessment and ensuring consistency, especially in borderline cases or disputes. Yet, its low usage may be due to resource constraints or time limitations.

Notably, **no teacher** reported using **blind grading**, indicating that assessments are generally conducted with full awareness of student identities. While this is typical in oral exams, it may raise concerns about potential unconscious bias.

### 17. What strategies or tools have you found most effective in improving students' oral skills?

**Table 11: Descriptive results on strategies and tools used by teachers to improve students' oral skills.**

Result	Frequency	Percentage	Percent of case
Role-plays and simulations	7	29.2	70
Group discussions/debates	7	29.2	70
Storytelling activities	1	4.2	10
Pronunciation apps (e.g., ELSA, Speechling)	1	4.2	10
Video/audio recording for self-assessment	3	12.5	30
Regular one-on-one speaking practice	5	20.8	50
Other	0	0	0
<b>Total</b>	24	100	-

The data reveals that teachers find **interactive and communicative activities** the most effective strategies to improve students' oral skills. Both **role-plays and simulations** and **group discussions or debates** were chosen by **70%** of respondents, indicating a strong

## Chapter two: Methodology, Results and Discussion

---

preference for active student participation and real-life speaking practice. These methods promote spontaneous language use and boost confidence, which are essential for oral proficiency.

Additionally, **regular one-on-one speaking practice** was reported by **50%** of teachers, highlighting the importance of personalized attention in developing individual speaking skills.

Other tools like **video/audio recording for self-assessment** were less common but still used by **30%**, offering students the chance to reflect on their performance and monitor progress.

Less frequently used strategies include **storytelling activities** and **pronunciation apps**, both selected by only **10%** of participants. This might reflect limited access to technology or less emphasis on these techniques in the teachers' current practice.

Overall, the preference leans clearly towards interactive, communicative, and personalized approaches that engage students actively.

**18. Do you believe self-assessment helps students improve their oral proficiency? Why or why not?**

*(Select one and elaborate if needed)*

**Table 12: Frequency and percentage of teachers' beliefs on the effectiveness of self assessment in improving students' oral proficiency.**

<b>Result</b>	<b>Frequency</b>	<b>Percentage</b>
<b>Yes, because it encourages reflection and ownership of learning.</b>	2	20

## Chapter two: Methodology, Results and Discussion

---

No, because students may overestimate or underestimate their skills	2	20
<b>Sometimes</b> , depending on student maturity and training in self-assessment	6	60
Additional comments	0	0
<b>Total</b>	10	100

The responses reveal a cautious but generally positive attitude toward self-assessment as a tool for improving oral proficiency. The majority, **60%** of teachers, believe that self-assessment can be effective **sometimes**, depending largely on the student's maturity and the training they receive in how to evaluate themselves properly. This highlights the importance of guidance and skill-building to make self-assessment a reliable practice.

Only **20%** of teachers fully agree that self-assessment helps students improve because it encourages reflection and ownership of their learning process. This suggests that a smaller portion of teachers see self-assessment as inherently empowering and beneficial.

Conversely, another **20%** of teachers disagree, expressing concern that students may either **overestimate or underestimate** their speaking skills, which could undermine the accuracy and usefulness of self-assessment.

Overall, the results point to a nuanced view: while self-assessment has potential benefits, its effectiveness depends on proper implementation and student readiness.

### 2.2.1.5 Analysis of the fourth questionnaire section: Support and Recommendations.

19. What additional support or resources would help you conduct better oral assessments?

**Table 13: Descriptive results of teachers' responses on additional support and resources for conducting oral assessments.**

Result	Frequency	Percentage	Percent of case
Training workshops on oral assessment techniques	5	17.9	50
Access to speech analysis software (e.g., AI tools)	5	17.9	50
Pre-designed rubrics and assessment templates	4	14.3	40
Smaller class sizes or teaching assistants	7	25	70
More time allocated for assessments	7	25	70
Other	0	0	0
<b>Total</b>	28	100	-

The data shows that teachers emphasize the need for **more practical support and better working conditions** to improve oral assessments. The top two supports mentioned by **70%** of respondents are **smaller class sizes or the presence of teaching assistants** and **more time allocated for assessments**. This highlights the challenge of managing large groups and the time constraints that hinder thorough oral evaluation.

Equally important, **50%** of teachers expressed the need for **training workshops on oral assessment techniques** and **access to speech analysis software, including AI tools**. This indicates a desire for professional development and technological support to enhance the accuracy and effectiveness of oral assessments.

## Chapter two: Methodology, Results and Discussion

---

Furthermore, **40%** of teachers requested **pre-designed rubrics and assessment templates**, pointing to a need for standardized tools to ensure consistency and fairness in evaluation.

Overall, the findings suggest that improving oral proficiency assessment requires both **human resources and technical resources**, alongside proper training and time management.

### **20. Any other comments or suggestions regarding oral proficiency assessment?**

*(Optional free response)*

The open responses reflect several important concerns and suggestions from teachers regarding oral proficiency assessment. A recurring theme is the **lack of adequate vocabulary among students**, which negatively impacts their motivation and attitude toward oral assessments. One teacher pointed out that this limited vocabulary makes students dislike oral tests. Another significant suggestion is the **need for more training and professional development** for teachers. Multiple comments emphasize that teachers should receive enhanced training on assessment techniques to conduct evaluations more fairly and effectively. The use of **technology and standardized rubrics** is also highlighted as a crucial improvement area to ensure consistency and objectivity in oral assessments. Additionally, one response stresses the importance of **motivating students to speak without fear of making mistakes**, recommending varied oral activities such as presentations, dialogues, and debates to build students' confidence and fluency. Finally, the responses acknowledge the **general difficulty students face with oral expression**, reinforcing the need for supportive teaching strategies and assessment tools.

### 2.2.2 Analysis of the students' questionnaire data

#### 2.2.2.1 Analysis and discussion of questionnaire reliability and validity

**Table 14 :Reliability Statistics based on Cronbach's Alpha**

Cronbach's Alpha	N of Items
0,824	20

The Cronbach's Alpha value for the 20 items included in the student questionnaire is **0.824**, which indicates a **high level of internal consistency** among the items. Typically, a Cronbach's Alpha value above 0.8 is considered **reliable** and suggests that the items measure the same underlying concept—in this case, perceptions related to oral proficiency assessment.

This level of reliability also reflects that the questionnaire was well-designed, and the responses are dependable for further statistical analysis. It reduces concerns about random error or inconsistency in the answers.

**In conclusion**, we can say that the student questionnaire demonstrates a strong internal consistency, and the instrument is reliable for assessing students' views on oral proficiency evaluation.

**Table 15 : Validity statistics based on Pearson Correlation.**

Pearson Correlation	N	Sig. (2-tailed)
0.987	170	0.000

The **Pearson correlation coefficient** is **0.987**, which indicates a **very strong positive correlation** between the items and the overall score of the questionnaire. This means that the items are **highly consistent** with the general construct being measured (students' perceptions

## Chapter two: Methodology, Results and Discussion

of oral proficiency). Additionally, the **significance value** ( $p = 0.000$ ) is **less than 0.05**, which confirms that the correlation is **statistically significant**. Therefore, the questionnaire is considered to have **high construct validity**. The instrument used to measure students' oral proficiency perceptions demonstrates **excellent validity**, as indicated by a high Pearson correlation and statistically significant results.

### 2.2.2.2 Analysis of the first questionnaire section: Oral Proficiency Components.

**Table 16: Students' perceptions of the oral proficiency components.**

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I can speak English fluently without hesitation.	21	34	55	42	18
I use correct grammar when speaking English.	22	46	50	34	18
My pronunciation in English is clear and understandable.	16	36	47	44	27
I can express my ideas in English using a wide range of vocabulary.	24	37	46	41	22
I feel confident participating in English conversations.	22	21	48	54	25

**Table 17: Descriptive results of students' responses of the oral proficiency components.**

Statement	Mean	Std. Deviation	Result
I can speak English fluently without hesitation.	3,0118	1,17160	Neutral
I use correct grammar when speaking English.	2,8824	1,18583	Disagree
My pronunciation in English is clear and understandable.	3,1765	1,20822	Neutral
I can express my ideas in English using a wide range of vocabulary.	3,0000	1,24511	Neutral

## Chapter two: Methodology, Results and Discussion

I feel confident participating in English conversations.	3,2294	1,22608	Neutral
Section 1: Oral Proficiency Components	3,0600	0,93273	Neutral

The overall mean score for Section 1, which addresses oral proficiency components, is **3.06**, indicating a **neutral attitude** among students regarding their own oral English skills.

Looking more closely at the individual items:

The statement *"I can speak English fluently without hesitation"* has a mean of **3.01**, which reflects a **neutral** response. This suggests that students are unsure about their fluency, with some possibly confident and others still hesitant.

The item *"I use correct grammar when speaking English"* scored a mean of **2.88**, indicating **disagreement**. This shows that many students recognize weaknesses in their spoken grammar, which may impact their confidence in oral communication.

*"My pronunciation in English is clear and understandable"* received a mean of **3.17**, again a **neutral** stance. It seems students are not strongly confident in their pronunciation, though they may not see it as a major problem either.

The ability to express ideas using a wide range of vocabulary scored **3.00**, reflecting a **neutral** perception. This could suggest a limited vocabulary range or uncertainty in using diverse terms effectively during speech.

The item *"I feel confident participating in English conversations"* has the highest mean in this group, **3.23**, though it still remains within the **neutral** range. It indicates slightly more confidence in interactive speaking situations, yet not enough to suggest full assurance.

## Chapter two: Methodology, Results and Discussion

**In conclusion**, we can say that students generally hold a neutral view of their oral English proficiency. While they do not report major difficulties, they also do not express strong confidence, especially in grammar and vocabulary use. This highlights a need for targeted support in these areas to help learners develop greater fluency and accuracy in spoken English.

### 2.2.2.3 Analysis of the second questionnaire section: Teaching and Learning Methods.

**Table 18: Students' perceptions of teaching and learning methods**

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Role-playing activities help me improve my speaking skills.	12	19	29	56	54
Watching English TV shows/movies improves my oral proficiency.	5	10	21	66	68
Group discussions in class enhance my communication skills.	12	16	42	56	44
I receive useful feedback from my teacher on my speaking performance.	14	15	41	58	42
I feel motivated to practice speaking English in class.	13	11	39	58	49

**Table 19: Descriptive results of students' responses on teaching and learning methods**

Statement	Mean	Std. Deviation	Result
Role-playing activities help me improve my speaking skills.	3,7118	1,22324	Agree
Watching English TV shows/movies improves my oral proficiency.	4,0706	1,01221	Agree

## Chapter two: Methodology, Results and Discussion

Group discussions in class enhance my communication skills.	3,6118	1,17261	Agree
I receive useful feedback from my teacher on my speaking performance.	3,5824	1,19007	Agree
I feel motivated to practice speaking English in class.	3,7000	1,17594	Agree
Section 2: Teaching and Learning Methods	3,7353	0,80166	Agree

The average score for Section 2, focusing on teaching and learning methods, is **3.73**, which clearly indicates a general **agreement** among students that these methods contribute positively to their oral proficiency development.

Looking at individual items:

The statement *"Role-playing activities help me improve my speaking skills"* scored **3.71**, which shows that students **agree** that such interactive tasks are helpful in developing speaking abilities. It reflects an appreciation for dynamic, real-life simulations in the classroom.

*"Watching English TV shows/movies improves my oral proficiency"* received the **highest score** in this section at **4.07**, suggesting a **strong agreement** among students. They seem to recognize the value of authentic audiovisual input in improving pronunciation, vocabulary, and fluency.

The item *"Group discussions in class enhance my communication skills"* has a mean of **3.61**, indicating that students generally **agree** that peer interaction is beneficial. This suggests that collaborative learning environments contribute meaningfully to oral skill development.

*"I receive useful feedback from my teacher on my speaking performance"* scored **3.58**, showing **agreement** as well. Students appear to value the teacher's role in guiding and correcting their spoken English.

## Chapter two: Methodology, Results and Discussion

Finally, *"I feel motivated to practice speaking English in class"* scored **3.70**, showing that most students feel **encouraged and engaged** during speaking activities.

**In conclusion**, we can say that students perceive the teaching and learning methods applied in the classroom such as role-playing, group discussions, teacher feedback, and media exposure as effective and motivating tools for enhancing their oral proficiency.

### 2.2.2.4 Analysis of the third questionnaire section: Assessment Methods.

**Table 20: Students' perceptions of assessment methods.**

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Oral presentations are an effective way to assess my speaking skills.	15	20	55	54	26
I feel nervous during oral exams or speaking tests.	17	19	27	51	56
Pair/group activities (e.g., dialogues, debates) help me perform better in speaking assessments.	5	23	43	62	37
I prefer being assessed through informal methods (e.g., storytelling, picture description) over formal exams.	24	26	41	34	45
The grading criteria for oral tests are clear and fair.	20	26	50	55	19

**Table 21: Descriptive results of students' responses on assessments methods.**

Statement	Mean	Std. Deviation	Result
Oral presentations are an effective way to assess my speaking skills.	3,3294	1,13979	Neutral
I feel nervous during oral exams or speaking tests.	3,6471	1,31181	Agree
Pair/group activities (e.g., dialogues, debates) help me	3,6059	1,06197	Agree

## Chapter two: Methodology, Results and Discussion

perform better in speaking assessments.			
I prefer being assessed through informal methods (e.g., storytelling, picture description) over formal exams.	3,2941	1,37882	Neutral
The grading criteria for oral tests are clear and fair.	3,1588	1,17345	Neutral
Section 3: Assessment Methods	3,4071	0,65050	Agree

The average score for Section 3, which focuses on oral **assessment methods**, is **3.41**, reflecting a general **agreement** students that the current assessment approaches are fairly acceptable and supportive of their oral proficiency evaluation.

Breaking down the individual statements:

The statement *"Oral presentations are an effective way to assess my speaking skills"* received a **neutral** average of **3.33**. This suggests that students are somewhat uncertain about the effectiveness of oral presentations as a primary evaluation method, possibly due to performance anxiety or lack of practice.

*"I feel nervous during oral exams or speaking tests"* scored **3.65**, showing **agreement**. This reflects a common challenge among language learners, where test anxiety negatively affects performance, despite preparation.

The item *"Pair/group activities (e.g., dialogues, debates) help me perform better in speaking assessments"* received **3.61**, also showing **agreement**. Students seem to perform more confidently and effectively in collaborative and interactive settings compared to individual tests.

Regarding the statement *"I prefer being assessed through informal methods (e.g., storytelling, picture description) over formal exams"*, the average score was **3.29**, indicating a **neutral**

## Chapter two: Methodology, Results and Discussion

position. Students may not strongly lean toward informal methods, suggesting that preferences vary widely.

Lastly, "*The grading criteria for oral tests are clear and fair*" scored **3.15**, again falling in the **neutral** range. This shows that students might feel some ambiguity or inconsistency in the way oral assessments are graded.

**In conclusion**, we can say that while students generally appreciate collaborative and less stressful formats like group activities, they express concerns about test-related anxiety and lack of clarity in evaluation criteria. This suggests a need for more transparent and student-friendly assessment approaches in oral proficiency testing.

### 2.2.2.5 Analysis of the first questionnaire section: Challenges and Suggestions

**Table 22: Students' perceptions of challenges and suggestions.**

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I lack opportunities to practice English outside the classroom.	14	28	33	45	50
Cultural differences make it harder for me to communicate in English.	13	34	56	38	29
I would like more technology (e.g., apps, recordings) to practice speaking.	5	11	28	53	73
My teacher's feedback helps me identify my weaknesses in speaking.	9	10	34	62	55
I need more interactive activities to improve my oral proficiency.	8	8	36	50	68

## Chapter two: Methodology, Results and Discussion

**Table 23: Descriptive results of students' responses on challenges encountered and suggestions for improvement.**

Statement	Mean	Std. Deviation	Result
I lack opportunities to practice English outside the classroom.	3,5235	1,29269	Agree
Cultural differences make it harder for me to communicate in English.	3,2118	1,17261	Neutral
I would like more technology (e.g., apps, recordings) to practice speaking.	4,0471	1,05926	Agree
My teacher's feedback helps me identify my weaknesses in speaking.	3,8471	1,10412	Agree
I need more interactive activities to improve my oral proficiency.	3,9529	1,10840	Agree
Section 4: Challenges and Suggestions	3,7165	0,71722	Agree

The average score for Section 4, which deals with the **challenges students face and their suggestions** for improving oral proficiency, is **3.72**, indicating a clear **agreement** with most of the statements. This reflects that students not only experience real obstacles but also have a strong awareness of what could enhance their speaking skills.

Let's look at the individual items:

The statement "*I lack opportunities to practice English outside the classroom*" scored **3.52**, showing **agreement**. This indicates that many students feel limited by their environment, lacking natural contexts to use English beyond school settings.

The item "*Cultural differences make it harder for me to communicate in English*" received a **neutral** score of **3.21**, which suggests that while some students may feel culturally disconnected when speaking English, it's not a universally shared struggle.

## Chapter two: Methodology, Results and Discussion

---

A high score of **4.05** was given to *"I would like more technology (e.g., apps, recordings) to practice speaking"*, clearly showing **agreement**. This highlights students' enthusiasm for integrating modern tools into learning, which could provide flexibility and more autonomous practice opportunities.

*"My teacher's feedback helps me identify my weaknesses in speaking"* scored **3.85**, indicating that students appreciate and benefit from the feedback they receive, which they find helpful for self-improvement.

Lastly, the statement *"I need more interactive activities to improve my oral proficiency"* scored **3.95**, again showing strong **agreement**. Students seem to favor dynamic, engaging activities that go beyond traditional lecture-based learning.

**In conclusion**, students recognize that challenges such as limited exposure and cultural barriers exist, but they also express a strong desire for more interactive, tech-based, and feedback-driven strategies to boost their speaking skills.

### 2.3 Hypothesis Testing:

#### 2.3.1 Students' Hypotheses:

1- Students' self-perceptions of their oral proficiency are generally good, but they tend to rate **fluency and confidence** lower than other components (grammar, pronunciation, vocabulary) due to anxiety during spontaneous speaking.

- **Null Hypothesis (H<sub>0</sub>):**

Students' self-perceptions of their oral proficiency are neutral, and there is no significant difference indicating a positive perception.

## Chapter two: Methodology, Results and Discussion

### • Alternative Hypothesis ( $H_1$ ):

Students' self-perceptions of their oral proficiency are significantly positive, even if some components (like fluency and grammar) are perceived lower due to anxiety.

**Table 24: T-Test Analysis of Students' self-perceived oral proficiency levels.**

N	Mean	Std. Deviation	T test	Df	Sig
170	3.0600	0.93273	42.775	169	0.000

From the table above, we notice that the overall mean score for students' self-assessment of oral proficiency is **3.06**, which is slightly above the neutral point (3.00). The T-test yielded a **statistically significant result** with  $p < 0.001$ , indicating that students' self-perceptions are **significantly more positive** than neutral.

Although the mean is not very high, it suggests that students have a **moderately positive view** of their oral proficiency. However, when we look closer at individual components from earlier analysis, we notice a pattern: students rated their **fluency (M = 3.01)** and **use of grammar (M = 2.88)** lower than their **confidence (M = 3.22)** or **pronunciation (M = 3.17)**. This reinforces the idea in the hypothesis that students experience **anxiety or hesitation when speaking spontaneously**, which likely affects their perceived fluency and grammar accuracy. Therefore, the hypothesis is **supported** by both the statistical result (significant positive perception overall) and the **detailed pattern** in subcomponents (lower fluency & grammar scores due to anxiety). Since the p-value is less than 0.05, we reject the null hypothesis ( $H_0$ ) and accept the alternative hypothesis ( $H_1$ ).

2- There is a **moderate level of acceptance** of teaching and learning methods among students.

## Chapter two: Methodology, Results and Discussion

---

- **Null Hypothesis (H<sub>0</sub>):**

Students show a neutral level of acceptance of teaching and learning methods, with no statistically significant difference.

- **Alternative Hypothesis (H<sub>1</sub>):**

Students show a statistically significant positive acceptance of teaching and learning methods.

**Table 25: T-Test Results for students' Acceptance of Teaching and Learning Methods.**

<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>T test</b>	<b>df</b>	<b>Sig</b>
170	3.7353	0.80166	60.752	169	0.000

From the table above, the mean value of **3.73** clearly exceeds the neutral midpoint of 3.00. The **T-value (60.752)** is very high, and the **p-value (0.000)** is less than 0.05, indicating that the result is statistically significant. Therefore, students demonstrate a strong and positive level of acceptance toward teaching and learning methods used to improve oral proficiency.

This aligns with previous descriptive results showing that students agreed with the usefulness of **role-playing (M = 3.71)**, **watching English media (M = 4.07)**, and **group discussions (M = 3.61)**. These activities seem to foster motivation and engagement in speaking tasks.

Since the p-value is less than 0.05, we reject the null hypothesis (H<sub>0</sub>) and accept the alternative hypothesis (H<sub>1</sub>).

There is a **moderate level of acceptance** of oral assessment methods among students, where they report **higher anxiety** during formal assessments (e.g., presentations, interviews), and prefer **informal assessments** (e.g., storytelling, picture description) due to lower stress levels.

## Chapter two: Methodology, Results and Discussion

---

- **Null Hypothesis ( $H_0$ ):** There is no significant level of acceptance of oral assessment methods among students.
- **Alternative Hypothesis ( $H_1$ ):** There is a significant level of acceptance of oral assessment methods among students.

N	Mean	Std. Deviation	T test	df	Sig
170	3.4071	0.65050	68.290	169	0.000

From the table above, we observe that the **mean score is 3.41**, which is **above the neutral value of 3.00**, indicating that students show a **significant acceptance** of oral assessment methods.

The T-value (68.290) is very high, and the **p-value (Sig.) is 0.000**, which is **less than the threshold of 0.05**, confirming the result is **statistically significant**.

This aligns with earlier descriptive analysis where students:

Reported **anxiety during formal assessments** (Mean = 3.64)

Showed **preference for informal assessments** like storytelling or picture description (Mean = 3.29)

Noted that **pair/group activities** help improve performance (Mean = 3.60).

These results suggest students accept oral assessment practices but feel **more comfortable and perform better** when assessed through **informal and less stressful methods**.

## Chapter two: Methodology, Results and Discussion

---

Since the p-value is less than 0.05, we reject the null hypothesis ( $H_0$ ) and accept the alternative hypothesis ( $H_1$ ).

Students face challenges in improving their oral proficiency, mainly due to a lack of practice opportunities outside the classroom and cultural barriers. They suggest implementing more interactive activities and technology-based tools (e.g., apps, recordings).

- **Null Hypothesis ( $H_0$ ):** Students do not significantly face challenges in improving their oral proficiency.
- **Alternative Hypothesis ( $H_1$ ):** Students significantly face challenges in improving their oral proficiency.

**Table 26: T-Test Results for Students' challenges in improving oral proficiency.**

N	Mean	Std. Deviation	T test	df	Sig
170	3.7165	0.71722	67.562	169	0.000

From the table above, the mean value is 3.72, which is clearly higher than the neutral point of 3.00. This suggests that students do indeed face significant challenges when trying to improve their oral English skills.

The T-value (67.562) is very high, and the p-value = 0.000, which is well below the significance level of 0.05, confirming the result is statistically significant.

This finding supports the earlier descriptive results:

## Chapter two: Methodology, Results and Discussion

Students reported **lack of opportunities outside the classroom** (Mean = 3.52)  
Mentioned **cultural barriers** (Mean = 3.21)  
Expressed strong desire for **more technology tools** (Mean = 4.05)  
Requested **more interactive activities** (Mean = 3.95) And acknowledged the benefit of **teacher feedback** (Mean = 3.84).

All these points indicate that students are **aware of the challenges** and have **clear suggestions** to address them.

Since **p-value < 0.05**, we **reject the null hypothesis (H<sub>0</sub>)** and **accept the alternative hypothesis (H<sub>1</sub>)**.

### 2.3.2 Teachers' Hypotheses:

Teachers perceive that students generally have a **good level of oral proficiency**.

• **Null Hypothesis (H<sub>0</sub>):**

*Teachers do not perceive that students have a good level of oral proficiency.*

• **Alternative Hypothesis (H<sub>1</sub>):**

*Teachers perceive that students have a good level of oral proficiency.*

**Table 27: T-Test Results for Teachers' perceptions of students' oral proficiency level.**

<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>T test</b>	<b>Df</b>	<b>Sig</b>
10	3,5300	0,61653	18.106	9	0.000

The descriptive statistics show a mean value of 3.53 and a standard deviation of 0.61. This suggests that, on average, teachers lean toward agreeing with the statement that students

## Chapter two: Methodology, Results and Discussion

---

have a good level of oral proficiency. The relatively moderate standard deviation implies that responses are somewhat consistent, with limited variability.

The T-value obtained is **18.106**, with **9 degrees of freedom (df = 9)**, and the **significance level (Sig.) is 0.000**, which is far below the threshold of **0.05**. This means the difference between the observed mean and the test value (**3.00**) is statistically significant. In other words, the result is not due to chance.

Given that the significance value is less than **0.05**, we reject the null hypothesis (which assumes no difference from the neutral perception). The result supports the idea that teachers perceive students' oral proficiency to be above average. Although the difference (**0.53**) is not dramatically large, it is statistically significant, and the consistency of responses (as reflected by the standard deviation) strengthens the finding.

However, it is worth noting that a mean of **3.53** is closer to the middle of the scale than to the maximum score of **5**. This suggests that while teachers generally perceive students to have an acceptable level of oral proficiency, their agreement is **moderate** rather than strong. It reflects **a cautiously positive evaluation** rather than full confidence.

Based on the statistical evidence, the hypothesis is **accepted**. Teachers indeed perceive that students possess a generally good level of oral proficiency. The mean score is significantly higher than the neutral value, confirming a positive perception. However, the moderate distance from the maximum score and the standard deviation suggest that this perception, while positive, is not overwhelmingly strong.

In conclusion, the hypothesis is confirmed: teachers show a statistically significant tendency to perceive students' oral proficiency as good. This result indicates a general satisfaction with students' oral performance, though with a degree of caution. This perception

## Chapter two: Methodology, Results and Discussion

---

can serve as a foundation for further pedagogical efforts aimed at refining and expanding oral language skills.

Teachers face **numerous challenges** in assessing oral proficiency, particularly **large class sizes** and **student anxiety**. Meanwhile, the **use of standardized rubrics** and **peer assessments** are among the most commonly used strategies to ensure fairness.

The hypothesis is **partially confirmed** by the data.

The results from **question 15** clearly support the first part of the hypothesis:

**Student anxiety or nervousness** was reported by **90%** of teachers, making it the **most cited challenge**.

**Large class sizes** were also a major concern, selected by **70%** of participants. These findings confirm that psychological and structural barriers are among the most significant obstacles in oral assessment.

However, the second part of the hypothesis concerning the **use of standardized rubrics and peer assessments** is **supported**.

According to **question 16**, only **40%** of teachers reported using **standardized rubrics**, and the same percentage used **peer or co-teacher evaluations**. While these are present strategies, they are **not dominant**. The **most common method** to ensure fairness was "**providing clear instructions and practice**", selected by **60%** of teachers, which is not mentioned in the hypothesis.

Moreover, **question 17** shows that **interactive strategies** like **role-plays** and **group discussions** are preferred to improve oral skills, rather than structural assessment tools like rubrics.

## Chapter two: Methodology, Results and Discussion

---

In **question 18**, teachers expressed **mixed feelings** about self-assessment—60% believe it is effective only under certain conditions, showing that teachers are cautious about methods that rely heavily on student responsibility.

**Conclusion:** While **student anxiety** and **large class sizes** are indeed key challenges, the **use of standardized rubrics and peer assessments**, although present, is not as widespread as the hypothesis suggests.

Teachers need **additional support** to improve oral proficiency assessment, including **training workshops, speech analysis software**, and **smaller class sizes** to enhance assessment quality.

The data strongly supports the need for training workshops. According to Question 19, half of the respondents (50%) selected “training workshops on oral assessment techniques” as a necessary form of support. Furthermore, in the open-ended responses of Question 20, many teachers explicitly stated the lack of sufficient preparation in oral assessment methods. They emphasized the need for specialized training that would enable them to assess students more fairly and confidently. This reveals a clear gap in professional development and aligns with the hypothesis that workshops are crucial for enhancing teachers’ assessment skills.

Another major component of the hypothesis is the importance of access to technology, particularly speech analysis software. The data shows that 50% of teachers requested such tools, confirming their relevance. In the open comments, several teachers discussed the potential role of AI and standardized rubrics to improve assessment objectivity and consistency. This implies that technological tools are seen not just as add-ons but as integral to improving the quality of oral testing. Therefore, the hypothesis correctly anticipates a growing demand for digital solutions in the assessment process.

## Chapter two: Methodology, Results and Discussion

---

The most frequently cited support option was “smaller class sizes or teaching assistants,” selected by **70%** of respondents. An equal percentage (**70%**) also indicated the need for more time to conduct oral assessments. These two points reflect structural issues within the classroom environment that make oral assessment difficult to manage. Large class sizes and limited time are barriers that affect the depth and fairness of assessment. This confirms that improving classroom conditions is a top priority for teachers and supports the hypothesis that institutional changes are necessary for better assessment quality.

Based on both the quantitative and qualitative data, the hypothesis is confirmed. Teachers express a consistent and significant need for professional development, technological assistance, and improved teaching conditions. The overlap in closed and open responses reinforces the argument that current support systems are insufficient. The findings demonstrate that multiple forms of support must be combined to truly enhance the oral proficiency assessment process.

In conclusion, Hypothesis 3 is strongly supported by the data. Teachers are clearly seeking multi-dimensional support to address the challenges they face in assessing oral skills. Their responses highlight the importance of training, technology, and institutional reform. Therefore, to improve the assessment of oral proficiency, education institutions must invest in teacher training programs, digital tools, and better classroom organization. These combined efforts can create a more effective and equitable assessment environment.

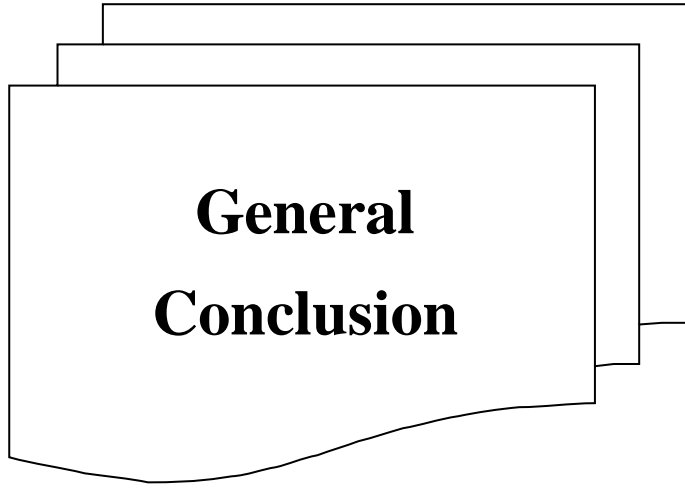
### Conclusion

The findings of this study suggest four challenges of oral proficiency assessment: students' fear; large class size; teacher lack of training; and the need to use interactive, clear, and technological approaches. Teachers and students agreed that oral proficiency assessment

## **Chapter two: Methodology, Results and Discussion**

---

needs to be done in a manner that encourages interactive and clear guidelines based on technological approaches. Overcoming these challenges can promote fairness and effectiveness of oral proficiency assessment and improve oral language learning. This chapter brings into focus that institutional support, pedagogical approaches, and students' interest go hand in hand with oral proficiency assessment.



**General  
Conclusion**

## General conclusion

---

### General conclusion:

#### Limitations of the study:

While this study has brought in valuable insights into assessing oral proficiency in EFL settings, it has a number of limitations as follows:

#### Sample characteristics

- Geographical and institutional constraints: The study was conducted in a specific regional setting (e.g. Algerian secondary schools) and as such the findings cannot be generalized to other educational systems or cultural settings. English speakers in other countries might have varying levels of exposure to the target language outside classroom time, which would affect oral proficiency differently.
- Sampling method: The use of convenience sampling (voluntary participation) means that there may be self-selection bias as students and teachers who chose to participate may have stronger opinions or been more positive toward oral assessments than others.
- Demographic representation: The sample does not necessarily represent all demographic groups (e.g. varying degrees of proficiency or socioeconomic background), which would skew the results.

#### Methodological Constraints

- Dependence on Self-Reports: The study relies heavily on questionnaires, vulnerable to response biases (e.g., social desirability bias, where respondents over-report positive behaviors); for instance, students may have rated their speaking ability higher than it actually is out of confidence issues.

## General conclusion

---

- **Lack of Objective Measures:** Without direct assessments (e.g., speaking tests recorded, external raters), the study only considers subjective perceptions of proficiency, which do not necessarily correspond to actual performance.
- **Cross-Sectional Design:** The study is cross-sectional only, capturing data at a single time point, so it cannot over time track any changes in oral proficiency or teaching methods. A longitudinal approach would provide deeper insights.

### **Contextual and theoretical boundaries**

- **Institutional variation:** Whilst differences in school resources, e.g. access to a language lab or teacher training programs, could have potentially impacted on the assessment strategies, this was not considered.
- **Cultural specificity:** The study did not investigate in depth how culture, e.g. cultural communication styles, may impact on oral assessment, e.g. learners' reluctance to speak up in class.
- **Narrow focus:** The study focused on specific components of oral proficiency, e.g. fluency, vocabulary, but not on the broader communicative competence, e.g. pragmatics, non-verbal communication.

### **Pedagogical Implications:**

Results provide practical suggestions for EFL teachers, curriculum designers, and policymakers to improve oral proficiency instruction and assessment.

### **Curriculum Development**

## General conclusion

---

- **Balanced Skill Integration:** Oral proficiency should be taught along with other language skills (e.g., listening, writing) to foster holistic language development. For example, task-based learning can integrate speaking with reading and critical thinking.
- **Task Variety:** Incorporate diverse speaking activities (e.g., debates, storytelling, simulations) to address different aspects of oral proficiency (fluency, accuracy, pronunciation).
- **Authentic Communication:** Use real-world tasks (e.g., role-playing job interviews, giving presentations) to prepare students for practical language use.

### Assessment Practices

- **Varied Assessment Methods:** Use a combination of formal assessments (e.g. oral exams) and informal assessments (e.g. peer feedback, self-recordings) to reduce anxiety and capture multiple performance measures.
- **Transparent Rubrics:** Develop clear, criterion-referenced scoring guides to ensure consistency and fairness in grading. For instance, rubrics could have different categories for fluency, grammar, and pronunciation.
- **Formative Feedback:** Provide meaningful, specific feedback after assessments to help students identify strengths and weaknesses.

### Teacher Training and Support

- **Professional Development:** Provide professional development on oral assessment techniques, such as how to conduct interviews and use technology for assessments.
- **Rater Calibration:** Train teachers on how to use rubrics consistently to reduce subjectivity in scoring.

## General conclusion

---

- Resource Allocation: Supply classrooms with resources such as recording devices and speech analysis software to facilitate self-assessment and peer review.

### Student-Centered Approaches

- Anxiety Reduction: Create low-stakes speaking opportunities such as pair work and small-group discussions to build confidence before high-stakes assessments.
- Self-Assessment: Encourage students to reflect upon their own progress through journals or recording tools to promote autonomy.
- Technology Integration: Use apps like ELSA Speak or Flipgrid for personalized pronunciation practice and feedback.

### Recommendations for Future Research:

To address the limitations of this study and build on its findings, future research could encompass the following directions:

#### Methodological Advancements

- Longitudinal Studies: Track students' oral proficiency development over time to identify effective teaching strategies.
- Mixed-Methods Approaches: Combine questionnaires with classroom observations, interviews, or recorded speaking tests for richer data.
- Experimental Designs: Test the impact of specific interventions (e.g., technology-enhanced assessments) on learning outcomes.

#### Theoretical Expansions

- Broader Proficiency Models: Investigate additional components of oral proficiency (e.g., pragmatic competence, discourse management).

## General conclusion

---

- Affective Factors: Examine how motivation, anxiety, and cultural background influence speaking performance.
- Intercultural Communication: Explore how students adapt their speaking skills in multicultural settings.

### Technological Innovations

- AI and Speech Recognition: Evaluate the effectiveness of tools like ChatGPT or speech-to-text software in providing real-time feedback.
- Virtual Reality (VR): Assess whether VR simulations (e.g., virtual classrooms, job interviews) improve speaking confidence.
- Automated Scoring Systems: Investigate the reliability of AI-based grading for large-scale oral assessments.

### Policy and Practice

- Teacher Training Programs: Research the effects of professional development on the quality of assessment.
- National Standards: Research how standardized oral proficiency guidelines (e.g. CEFR) are implemented between institutions.
- Equity in Assessment: Research how assessment methods impact diverse populations of students (e.g. non-native speakers, students with learning differences).

This study explored the complexity of speaking skill assessment in EFL contexts from the perspectives of students, teachers, and challenges in speaking skill assessment. The findings highlight several key insights that contribute to both theoretical understanding and practical application in language education.

## General conclusion

---

First, the research underscores the complexity of oral proficiency as a construct encompassing not only linguistic accuracy but also communicative fluency, cultural awareness, and psychological factors like confidence and anxiety. Students demonstrated varying levels of self-perceived competence across different components of speaking ability, with particular struggles in spontaneous fluency and pronunciation. These results align with existing literature on the multidimensional nature of oral communication while providing context-specific data from secondary education settings.

For pedagogical practice, the study reveals the effectiveness of interactive, student-centered approaches to oral skill development. Methods such as role-playing, group discussions, and technology-enhanced learning emerged as particularly beneficial, suggesting a need to move beyond traditional lecture-based instruction. The findings also emphasize the importance of aligning assessment methods with learning objectives, with evidence supporting the value of both formal evaluations and informal, continuous assessment strategies.

The investigation into assessment challenges uncovered significant barriers including student anxiety, large class sizes, and limited resources - issues that resonate with global concerns in language education. However, it also identified promising solutions through standardized rubrics, peer assessment, and technological integration, offering practical pathways for improvement.

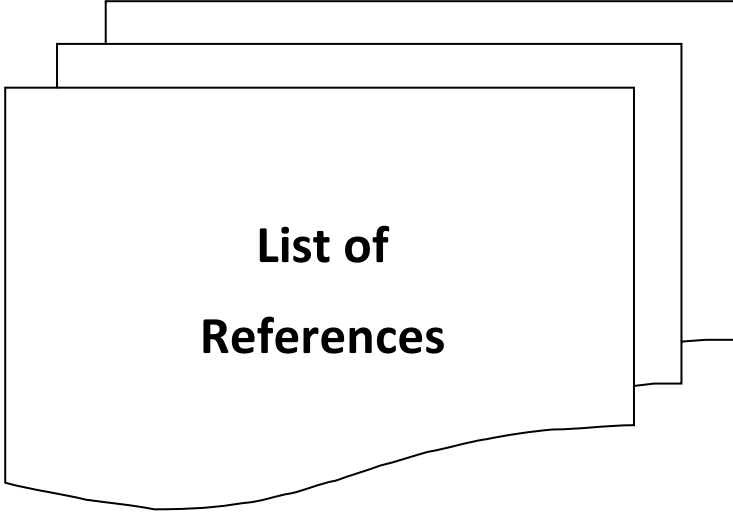
While the study provides valuable insights, its limitations regarding sample specificity and methodological constraints remind us that these findings should be interpreted within their particular educational context. The research nonetheless makes an important contribution to ongoing conversations about effective language assessment by:

## General conclusion

---

1. Validating the importance of communicative competence in EFL curricula
2. Highlighting the interplay between assessment methods and student motivation
3. Demonstrating the potential of alternative assessment formats to reduce anxiety
4. Identifying key areas for teacher professional development

As language education continues to evolve in response to globalization and technological advancement, this study reinforces the need for assessment practices that are not only valid and reliable but also equitable and conducive to genuine communicative development. The findings serve as a foundation for reimagining oral proficiency assessment in ways that better serve both learners and educators in diverse EFL contexts. Future research building on these insights could further refine assessment models and expand their applicability across different educational settings and learner populations.



**List of  
References**

## List of references

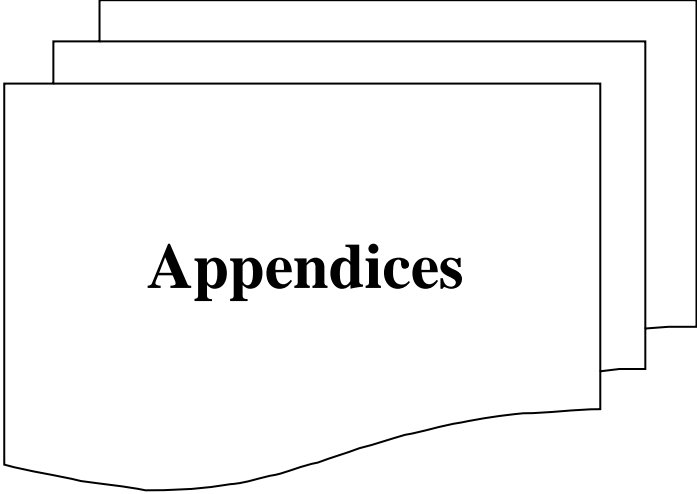
- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.
- Alonso, B. M. (2017). *Validación del contenido de una prueba oral de clasificación de E/LE* (Vol. 22). Universidad Antonio de Nebrija.
- Aryadoust, V. (2023). *The vexing problem of validity and the future of second language assessment* (Vol. 40). Language Testing.
- Bachman, L. F. (1988). *Problems in Examining the Validity of the ACTFL Oral Proficiency Interview* (Vol. 10). (C. U. Press, Éd.) Studies in Second Language Acquisition.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language testing in practice*. Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Beaven, A., & Neuhoff, A. (2011). *Assessing oral proficiency for intercultural professional communication: (Vol. 20)*. Nottingham, UK: Euro call: The Call Triangle: student, teacher and institution.
- Bijani, H., Hashempour, B., Khaled Ahmed Abdel Al Ibrahim, Orabah, S. S., & Heydarnejad, T. (2022). *Investigating the effect of classroom-based feedback on speaking assessment: a multifaceted Rasch analysis*. Language Testing in Asia.
- Bowden, H. W. (2016). *Assessing Second-Language Oral Proficiency for Research: The Spanish Elicited Imitation Task*. Studies in Second Language Acquisition.
- Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge University Press.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.
- Canale, M., & Swain, M. (1980). *Theoretical bases of communicative approaches to second language teaching and testing*. (Vol. 1). (<https://doi.org/10.1093/applin/1.1.1>, Éd.) Applied Linguistics.
- Celce-Murcia, M., D. Brinton, & Snow, M. A. (2014). *Teaching English as a second or foreign language* (4th ed.).
- Chapelle, C. A., & Chung, Y.-R. (2010). *The promise of NLP and speech processing technologies in language assessment*. (Vol. 27). Language Testing,.

- Chavez, G. (2013). *Examining the Reliability and Validity of ADEPT and CELDT: Comparing Two Assessments of Oral Language Proficiency for English Language Learners*. Digital Commons @ Loyola Marymount University.
- Chen, C. (2024). *Assessment of Test Validity in the Context of* (Vol. 14). (2. 1.-7. Open Journal of Modern Linguistics, Éd.) Scientific research publishing.
- Cook, D. A., & Lineberry, M. (2016). *Consequences validity evidence: Evaluating the impact of educational assessments*. Academic Medicine.
- Dandonoli, P., & Henning, G. (1990). *An Investigation of the Construct Validity of the ACTFL Proficiency Guidelines and Oral Interview Procedure* (Vol. 23). Foreign Language Annals .
- Daniel, M. H. (2010). *Reliability of AIMSweb Reading Curriculum-Based Measurement (R-CBM) (Oral Reading Fluency)*. <http://www.aimsweb.com/wp-content/uploads/Summary-of-Reliability-Information-for-AIMSweb-R-CBM.pdf>.
- Downing, S. M. (2006). *Face validity of assessments: faith-based interpretations or evidence-based science?* Medical Education .
- Ellis, R., & He, X. (1999). *The roles of modified input and output in the incidental acquisition of word meanings*. (Vol. 21). (<https://doi.org/10.1017/S0272263199002077>, Éd.) Studies in Second Language Acquisition,.
- Europe, C. o. (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.companion volume*. Council of Europe Publishing, Strasbourg,.
- Europe., C. o. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*.(<https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>, Éd.) Council of Europe Publishing.
- Fathman, A. (1976). *Variables Affecting the Successful Learning of English as a Second Language*. (Vol. 10). TESOL Quarterly .
- Fillmore, C. J. (1979). *Individual differences in language ability and language behavior* (éd. In C. J. Fillmore, D. Kempler, & W. S. Wang (Eds.)). (O. Fluency, Éd.) New York, New York , USA: Academic Press.
- Fulcher, G. (2003). *Testing second language speaking* (éd. 1st edition). London, UK: Routledge.
- Ginther, A. (2020). *Assessment of Speaking*. <https://doi.org/10.1002/9781405198431.wbeal0052.pub2>.
- González-Such, J., Meliá, J. M., & Bakieva, M. (2013). *Methodological Considerations Concerning the Assessment of Oral Competency in a Second Language (L2)* (Vol. 13). Revista Electrónica de Investigación Educativa.
- Henning, G. (1983). *Oral proficiency testing: comparative validities of interview, imitation, and completion methods*. (L. Learning, Éd.) Blackwell Publishing Ltd.

- Henning, G. (1983). *Oral proficiency testing: comparative validities of interview, imitation, and completion methods*. Blackwell Publishing Ltd.
- Iman, J. N., & Angraini, N. (2019). *Discussion Task Model in EFL Classroom: EFL Learners' Perception, Oral Proficiency, and Critical Thinking Achievements* (Vol. 133). Vytautas Magnus University.
- Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). *The Effects of Storytelling and Story Reading on the Oral Language Complexity and Story Comprehension of Young Children* (Vol. 32). (10.1023/B:ECEJ.0000048967.94189.a3, Éd.) Early Childhood Education Journal.
- Iskandarova, G. (2024). *Current issues in language assessment and language assessment research and its implication*. Baltic Journal of Legal and Social Sciences.
- Johnson, W. L., & Valente, A. (2009). *Tactical Language and Culture Training Systems: Using AI to Teach Foreign Languages and Cultures* (Vol. 30). (10.1609/aimag.v30i2.2240, Éd.) AI Magazine.
- Juffs, A., & DeKeyser, R. (2006). *Second Language Acquisition*. Pittsburgh: University of Pittsburgh.
- Karatay, Y., & Karatay, L. (2021). *Review of Automated speaking assessment: Using language technologies to score spontaneous speech* (Vol. 25). Language, Learning and Technology.
- Kiani, G. R., Alibakhshi, G., & Akbari, R. (2009). *On the Consequential Validity of ESP Tests: A Qualitative Study in Iran*. Islamic Azad University, Tabriz Branch.
- Kim, Y., & Kellogg, D. (2007). *Rules out of roles: Differences in play language and their developmental significance*. (Vol. 28). (<https://doi.org/10.1093/applin/aml046>, Éd.) Applied Linguistics.
- King, J. (2002). *Preparing EFL Learners for Oral Presentations*.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.
- Kyle, K., A. Crossley, S., & McNamara, D. S. (2016). *Construct validity in TOEFL iBT speaking tasks: Insights from natural language processing* (Vol. 33). Language Testing .
- Larson-Hall, J. (2008). *Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation*. (Vol. 24). (<https://doi.org/10.1177/0267658307082981>, Éd.) Second Language Research.
- Lennon, P. (2000). *The lexical element in spoken second language fluency. Perspectives on fluency*. (n. H. (Ed.), Éd.) Ann Arbor, Michigan, USA: the University of Michigan Press.
- Lyster, R., & Ranta, L. (1997). *Corrective feedback and learner uptake: Negotiation of form in communicative classrooms*. (Vol. 19). (<https://doi.org/10.1017/S0272263197001034>, Éd.) Studies in Second Language Acquisition.

- Malone, M. E. (2003). *Research on the Oral Proficiency Interview: Analysis, Synthesis, and Future Directions*. Foreign Language Annals.
- May, L. (2000). *Assessment of oral proficiency in EAP programs : a case for pair interaction*. Singapore Tertiary English Teachers Society.
- Mazdayasna, G. (2012). *Objective Assessment of Oral Presentations and EFL Learners' Speaking Development*. (Vol. 1). Sheikhbahae University Press.
- McKellar, N. A., & Unruh, S. (2017). *Oral Language Issues and Assessment of Oral Language*. Springer, Cham.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Blackwell Publishing.
- Mehrens, W. A. (2005). *The Consequences of Consequential Validity*. Blackwell Publishing Ltd.
- Memon, M. A., Joughin, G., & Memon, B. (2010). *Oral assessment and postgraduate medical examinations: establishing conditions for validity, reliability and fairness*. (S. Netherlands, Éd.) Advances in Health Sciences Education.
- Messick, S. (1995). *Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning*. (Vol. 50). (<https://doi.org/10.1037/0003-066X.50.9.741>, Éd.) American Psychologist,.
- Morreale, S. P., Osborn, M. M., & Pearson, J. C. (2000). *Why communication is important: A rationale for the centrality of the study of communication*. (Vol. 29). Journal of the Association for Communication Administration,.
- Mosier, C. I. (1947). *A critical examination of the concepts of face validity*. (Vol. 7). (<https://doi.org/10.1177/001316444700700201>, Éd.) Educational and Psychological Measurement.
- Newton, J. (2013). *Incidental vocabulary learning in classroom communication tasks*. (Vol. 17). (<https://doi.org/10.1177/1362168812460814>, Éd.) Language Teaching Research,.
- Nunan, D. (2001). *The Cambridge Guide to Teaching English to Speakers of Other Languages: Second language acquisition*. Cambridge University Press.
- Piñana, A. (2009). *Self assessment of foreign language oral production: assessment accuracy, the role of proficiency level and implicit training*. Universitat Autònoma de Barcelona.
- Rosser, H. L. (1979). *Testing Oral Communicative Skills*. (F. L. Annals, Éd.) Blackwell Publishing Ltd.
- Royal, K. D., & Puffer, J. C. (2014). *The Consequential Validity of ABFM Examinations*. Journal of the American Board of Family Medicine.
- Rubin, D. L., & Lim, H.-W. (2006). *Assessment of First Language Proficiency*. Georgia: University of Georgia.

- Sánchez, L. (2006). *Bilingualism/Second-Language Research and the Assessment of Oral Proficiency in Minority Bilingual Children*. (I. Lawrence Erlbaum Associates, Éd.) *Language Assessment Quarterly*.
- Sato, S. (1985). *Toward an Objective Assessment of Oral Proficiency* (Vol. 14).
- Shohamy, E. (1983). *The stability of oral proficiency assessment on the oral interview testing procedures*. Blackwell Publishing Ltd.
- Siddiek, D. A. (2010). *The Impact of Test Content Validity* (Vol. 6). Asian Social Science.
- Stinson, M., & Winston, J. (2011). *Drama education and second language learning: A growing field of practice and research*. (Vol. 5). Scenario: Journal for Performative Teaching, Learning, Research,.
- Surface, E. A., & C.Dierdorff, E. (2003). *Reliability and the ACTFL Oral Proficiency Interview: Reporting Indices of Interrater Consistency and Agreement for 19 Languages*. *Foreign Language Annals*.
- Swain, M., & Lapkin, S. (1998). *Interaction and second language learning: Two adolescent French immersion students working together*. (Vol. 82). (<https://doi.org/10.1111/j.1540-4781.1998.tb01209.x>, Éd.) *Modern Language Journal*,.
- Thu, D. M. (2019). *A review on validating language tests*. Vietnam National University Journal of Science.
- Tseng, T. J.-L. (2009). *On Measuring Communicative Competence: A Case Study on the Construct Validity of the Assessment Practice of an EFL Program* (Vol. 6).
- Walt, C. v., Wet, F. d., & Niesler, T. (2008). *Oral proficiency assessment: the use of automatic speech recognition systems* (Vol. 26). (T. & Group, Éd.) *Southern African Linguistics and Applied Language Studies*.
- Wolf, M. K., Farnsworth, T., & Herman, J. L. (2008). *Validity Issues in Assessing English Language Learners' Language Proficiency* (Vol. 13). (T. & Group, Éd.) *Educational Assessment*.
- Yorozuya, R., & Oller, J. W. (1980). *Oral proficiency scales: construct validity and the halo effect* (Vol. 30). (L. Learning, Éd.) Blackwell Publishing Ltd.
- Zhao, Z. (2013). *An Investigation into the Consequential Validity of a Diagnostic College English Speaking Test*. *English Language Teaching* .
- Zhao-hui, D. (2011). *A Study of the Reliability of Computerized Oral Proficiency Interview*. *Computer-assisted Foreign Language Education* .



**Appendices**

## Appendices

### Student Questionnaire on Oral Proficiency Assessment (Likert Scale)

#### Instructions:

Please indicate your level of agreement with each statement by ticking the appropriate choice, where:

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

#### Section 1: Oral Proficiency Components

**1. I can speak English fluently without hesitation.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**2. I use correct grammar when speaking English.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**3. My pronunciation in English is clear and understandable.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**4. I can express my ideas in English using a wide range of vocabulary.**

Strongly Disagree. Disagree. Neutral. Agree. Strongly Agree.

**5. I feel confident participating in English conversations.**

Strongly Disagree. Disagree. Neutral. Agree. Strongly Agree.

## **Section 2: Teaching and Learning Methods**

**6. Role-playing activities help me improve my speaking skills.**

Strongly Disagree. Disagree. Neutral. Agree. Strongly Agree.

**7. Watching English TV shows/movies improves my oral proficiency.**

Strongly Disagree. Disagree. Neutral. Agree. Strongly Agree.

**8. Group discussions in class enhance my communication skills.**

Strongly Disagree. Disagree. Neutral. Agree. Strongly Agree.

**9. I receive useful feedback from my teacher on my speaking performance.**

Strongly Disagree. Disagree. Neutral. Agree. Strongly Agree.

**10. I feel motivated to practice speaking English in class.**

Strongly Disagree. Disagree. Neutral. Agree. Strongly Agree.

### **Section 3: Assessment Methods**

**11. Oral presentations are an effective way to assess my speaking skills.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**12. I feel nervous during oral exams or speaking tests.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**13. Pair/group activities (e.g., dialogues, debates) help me perform better in speaking assessments.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**14. I prefer being assessed through informal methods (e.g., storytelling, picture description) over formal exams.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**15. The grading criteria for oral tests are clear and fair.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

### **Section 4: Challenges and Suggestions**

**16. I lack opportunities to practice English outside the classroom.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**17. Cultural differences make it harder for me to communicate in English.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**18. I would like more technology (e.g., apps, recordings) to practice speaking.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**19. My teacher's feedback helps me identify my weaknesses in speaking.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**20. I need more interactive activities to improve my oral proficiency.**

Strongly Disagree.    Disagree.    Neutral.    Agree.    Strongly Agree.

**Teachers' Questionnaire on Oral Proficiency Assessment**

**Instructions:**

This questionnaire aims to gather your insights and experiences regarding oral proficiency assessment in language learning. Your responses will remain confidential and will be used for research purposes only. Please answer all questions honestly and thoroughly.

## **Section 1: Demographic Information**

### **1. Gender:**

- Male
  
- Female

### **2. Years of Teaching Experience:**

- Less than 1 year
  
- 1-5 years
  
- 6-10 years
  
- More than 10 years

### **3. The average class size in your oral lessons::**

- Fewer than 20 students.
  
- 20 - 30 students.
  
- 30 - 40 students.

- More than 40 students.

**4. Language(s) You Teach: \_\_\_\_\_**

**Section 2: Current Assessment Practices (Likert Scale)**

Please rate the following statements on a scale of 1 (Strongly Disagree) to 5 (Strongly Agree).

<b>Statement</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<b>(Strongly Disagree)</b>	<b>(Disagree)</b>	<b>(Neutral)</b>	<b>(Agree)</b>	<b>(Strongly Agree)</b>
5. I regularly assess students' oral proficiency in my classes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. My institution provides clear guidelines for oral proficiency assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. I use a variety of assessment methods	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

<b>Statement</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<b>(StronglyDisagree)</b>	<b>(Disagree)</b>	<b>(Neutral)</b>	<b>(Agree)</b>	<b>(StronglyAgree)</b>

(e.g., interviews, role-plays, presentations).

**8.** I believe oral proficiency assessments accurately measure students' speaking skills.

	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

**9.** I receive adequate training on how to conduct oral assessments.

	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

**10.** I provide feedback to students after oral assessments.

	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

**11.** Students feel anxious during oral

	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

<b>Statement</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<b>(Strongly Disagree)</b>	<b>(Disagree)</b>	<b>(Neutral)</b>	<b>(Agree)</b>	<b>(Strongly Agree)</b>

assessments.

**12.** Technology (e.g., speech recognition software) improves oral assessment reliability.

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

**13.** Cultural factors influence students' performance in oral assessments.

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

**14.** Peer assessment is an effective method for evaluating oral proficiency.

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

**Section 3: Challenges and Strategies.**

**15. What challenges do you face when assessing students' oral proficiency?**

*(Select all that apply or add your own)*

- Lack of time for individual assessments
- Student anxiety or nervousness
- Difficulty in maintaining objectivity
- Limited resources (e.g., recording devices, software)
- Large class sizes
- Varying proficiency levels in the same class
- Lack of standardized rubrics
- Other: \_\_\_\_\_

**16. How do you ensure fairness and objectivity in oral assessments?**

*(Select all that apply or add your own)*

- Using standardized rubrics/scoring criteria
- Recording assessments for later review

- Multiple assessors (peer/co-teacher evaluations)
- Blind grading (assessing without knowing student identities)
- Providing clear instructions and practice beforehand
- Other: \_\_\_\_\_

**17. What strategies or tools have you found most effective in improving students' oral skills?**

*(Select all that apply or add your own)*

- Role-plays and simulations
- Group discussions/debates
- Storytelling activities
- Pronunciation apps (e.g., ELSA, Speechling)
- Video/audio recording for self-assessment
- Regular one-on-one speaking practice
- Other: \_\_\_\_\_

**18. Do you believe self-assessment helps students improve their oral proficiency?**

**Why or why not?**

*(Select one and elaborate if needed)*

- **Yes**, because it encourages reflection and ownership of learning.
- **No**, because students may overestimate or underestimate their skills.
- **Sometimes**, depending on student maturity and training in self-assessment.
- Additional comments: \_\_\_\_\_

#### **Section 4: Support and Recommendations.**

**19. What additional support or resources would help you conduct better oral assessments?**

*(Select all that apply or add your own)*

- Training workshops on oral assessment techniques
- Access to speech analysis software (e.g., AI tools)
- Pre-designed rubrics and assessment templates
- Smaller class sizes or teaching assistants

- More time allocated for assessments
  
- Other: \_\_\_\_\_

**20. Any other comments or suggestions regarding oral proficiency assessment?**

*(Optional free response)*

## الملخص

تستكشف هذه الدراسة الصدق والثبات في تقييمات الكفاءة الشفوية في سياق تعليم اللغة الإنجليزية كلغة أجنبية في ثانوية مولود قاسم في الجزائر. باستخدام منهجية البحث المختلط، جمعت الدراسة بين التقنيات الكمية والنوعية، بما في ذلك استبيانات مقياس ليكرت التي طُبقت على 170 طالبًا و10 معلمين. كشفت النتائج أنه بينما أظهر الطلاب تصورات متوسطة عن كفاءتهم الشفوية، إلا أنهم أبلغوا عن مستويات أعلى من القلق أثناء التقييمات الرسمية وفضلوا الأساليب غير الرسمية مثل سرد القصص وتمثيل الأدوار. من جانبهم، اعترف المعلمون بتحديات مثل كثافة الفصول الدراسية، وقلق الطلاب، وعدم وجود معايير تقييم موحدة، ومع ذلك رأوا أن مستويات الكفاءة الشفوية للطلاب جيدة بشكل عام. أكدت الدراسة على أهمية ضمان الصدق والثبات في التقييمات الشفوية من خلال استخدام أساليب متنوعة، وتوظيف التكنولوجيا، ومعايير تقييم واضحة. كما سلطت النتائج الضوء على الحاجة إلى الدعم المؤسسي، وتدريب المعلمين، واستراتيجيات تركز على الطالب لتحسين تقييم الكفاءة الشفوية. في الختام، تدعو الدراسة إلى اعتماد ممارسات تقييم متوازنة تعزز الكفاءة التواصلية مع معالجة التحديات السياقية في بيئات تعلم اللغة الإنجليزية كلغة أجنبية.

## الكلمات المفتاحية

تقييم الكفاءة الشفوية، تصورات الطلاب، تحديات المعلمين، استراتيجيات تعلم اللغة، الكفاءة التواصلية

## Résumé

Cette étude examine la validité et la fiabilité des évaluations de la compétence orale dans le contexte de l'enseignement de l'anglais langue étrangère (ALE) au lycée Mouloud Quassem en Algérie. Adoptant une approche mixte, la recherche a combiné des techniques quantitatives et qualitatives de collecte de données, incluant des questionnaires à échelle de Likert administrés à 170 élèves et 10 enseignants. Les résultats ont révélé que, bien que les élèves aient une perception modérée de leur compétence orale, ils ont exprimé une anxiété accrue lors des évaluations formelles et ont préféré des méthodes informelles telles que les récits et les jeux de rôle. Les enseignants ont reconnu les défis liés aux effectifs élevés, à l'anxiété des élèves et au manque de grilles d'évaluation standardisées, tout en estimant que le niveau de compétence orale des élèves était globalement satisfaisant. L'étude a souligné l'importance d'assurer la validité et la fiabilité des évaluations orales par des méthodes diversifiées, l'intégration technologique et des critères de notation transparents. Les résultats mettent en évidence la nécessité d'un soutien institutionnel, de formations pour les enseignants et de stratégies centrées sur l'apprenant pour améliorer l'évaluation de la compétence orale. En conclusion, la recherche plaide en faveur de pratiques d'évaluation équilibrées, privilégiant la compétence communicative tout en tenant compte des défis contextuels dans les environnements ALE.

### Mots-clés :

Évaluation de la compétence orale, Perceptions des élèves, Défis des enseignants, Stratégies d'apprentissage des langues, Compétence communicative.