

Evaluating Speaking Assessment Performance: a Case Study of an Accredited Training Center "Afak" at Skikda Province.

تقييم مهارات التحدث : دراسة حالة لمركز تدريب معتمد "Afak" في نواحي سكيكدة.

Received: 29/04/2025

Accepted: 26/06/2025

Published: 30/07/2025

Sofiane MECHTOUF *

Frère Mentouri Constantine1 University - (Algeria)

Email : mechtoufs@yahoo.fr



Journal of research and human studies
Issue 19 N°01 year 2025 P.P. 982-993
ISSN : 1112-8151 EISSN : 2588-2317

Abstract :

Assessing speaking skills is a key part of supporting language learners' development. This study explores how speaking is currently assessed at Afak Training Center in Skikda province. Using a descriptive approach, a questionnaire was given to all the English teachers at the center. The findings show that most teachers rely on a mix of holistic and analytical methods, with a strong focus on interactive and conversational tasks. However, some variation was noted in the assessment techniques, particularly in the use of checklists and technology. Based on these results, the study suggests providing teachers with further training in speaking assessment and creating standardized rubrics to help ensure fair and consistent evaluation.

Keywords: *Speaking assessment, assessment methods, testing criteria, evaluation practices.*

الملخص :

يعد تقييم مهارات التحدث جزءا أساسيا من دعم تطور المتعلمين في اللغة وتهدف هذه الدراسة إلى استكشاف كيفية تقييم مهارات التحدث في مركز أفاق للتدريب بولاية سكيكدة . تم استخدام المنهج الوصفي، حيث تم توزيع استبيان على جميع معلمي اللغة الإنجليزية في المركز. أظهرت النتائج أن معظم المعلمين يعتمدون على مزيج من الأساليب الشاملة والتحليلية، مع التركيز الكبير على المهام التفاعلية والحوارية. ومع ذلك، لوحظ بعض التباين في تقنيات التقييم، خاصة في استخدام قوائم المراجعة والتكنولوجيا. استنادا إلى هذه النتائج، تقترح الدراسة توفير المزيد من التدريب للمعلمين في مجال تقييم التحدث وإنشاء جداول تقييم موحدة للمساعدة في ضمان التقييم العادل والمتسق.

الكلمات المفتاحية: تقييم التحدث، أساليب التقييم، معايير الاختبار، ممارسات التقييم

1. Introduction:

Verbal communication predominantly constitutes the primary mode of human interchange. Hence, it is axiomatic for an individual's fluency and proficiency in a language to be equated with their capacity for oral expression. In this vein, Celce-Murcia (2001) underscores the primacy of spoken language in linguistic competence, arguing that "the ability to speak a language is synonymous with knowing that language since speech is the most basic means of human communication" (p.103). Added to the linguistic competence, oral communication is also an outstanding medium for cultural dissemination and self-expression, serving as a window into the speaker's background and worldview. As argued by Luma (2004), "Our personality, our self-image, our knowledge of the world and our ability to reason and express our thoughts are all reflected in our spoken performance in a foreign language" (p. 10). It is also an integrative skill that draws upon the symbiosis of other skills like listening, reading, and writing, which further points to its very significance among the four language skills. This very primacy is most evident in the fact that language users are often identified as or referred to as being its "speakers" (Ur, 2000).

Yet, despite its importance, its assessment poses considerable hurdles due to its very intricate nature as it involves a tangled web of lexical, phonological, grammatical, and pragmatic factors. Hence, to ensure the validity of its assessment, it is imperative to employ careful methods, develop rigorous rating criteria, and offer adequate training to evaluators (Pennington 1999).

2. Litertaure review

2.1 Paradigms for Evaluating Spoken Language Competence:

According to Clark (1979), language assessment methods can be classified into three distinct types: indirect, semi-direct, and direct. Direct methods which engage face-to-face interaction between the examinee and the interlocutors are especially appropriateand ideally suited for such assessment as they allow for evaluating language in life-like situations as best exemplified by the Oral Proficiency Interview (OPI).

Semi-direct assessment methods, while not requiring the physical presence of an interviewer, employ technology to tape examinees' responses to pre-recorded stimuli. Yet, in this method, it is crucial for the examinees to understand that their performance is being evaluated, not simply engaged in a social interaction (Clark, 1979).

Clark (1979) describes the **indirect method of speaking assessment** as an approach that evaluates oral proficiency *without* requiring direct spoken responses from learners; instead, it infers speaking ability through related skills like written tests (e.g., grammar, vocabulary), listening tasks, or even reading-based exercises (e.g., summarizing passages). This method assumes that competence in these correlated skills reflects underlying speaking ability, offering a practical but less authentic alternative to direct oral exams, as it avoids logistical challenges like one-on-one testing—though it may overlook nuances of fluency, pronunciation, and spontaneous interaction.

2.2 Methods for Assessing Oral Proficiency:

Oral language testing approaches can be broadly categorized into two chief types: holistic and analytic. Holistic assessment, which is rooted in the Gestalt principle of weighing the whole rather than its separate constituents, prioritizes the overall impression of the speaker's performance as implied by its very name (Hornby, 2000). In this assessment paradigm, the skill is viewed as an indivisible whole, rather than a set of discrete and isolated components. As argued by Xi et al (2006),

In holistic scoring, raters consider the combined impact of delivery, language use, and topic development, and make a judgment about a person's performance on a particular task. During this process, raters attempt to weigh the impact of different dimensions on the overall effectiveness of communication to come up with a holistic score (p. 32).

This perspective posits that the constituents of speaking, such as fluency, accuracy, vocabulary, and pronunciation, are all intertwined, mutually dependent and should be evaluated as one entity, prioritizing the overall impact of the speaker's performance.

While the holistic approach considers the overall performance, the analytic approach centers rather on evaluating those very separate components such as vocabulary, accuracy, fluency and pronunciation, assigning scores to each (Hughes,1989). As Xi and Mollaun (2006) suggest, analytic scoring is an assessment approach that involves the evaluation of specific aspects of performance, such as linguistic accuracy, linguistic accuracy, presentational clarity and thematic coherence. This approach involves a more granular analysis of the speaker's performance, allowing for a more nuanced assessment.

2.3 Oral Performance Rating Rubrics:

A scale, as outlined by Crocker and Algina (1986), delineates the range of values associated with different levels of achievement. Speaking assessments commonly involve the categorization of students into hierarchical levels of performance, with scores assigned to denote their relative performance. As previously discussed, these scales are categorized as either holistic or analytic.

Holistic scales typically span from one to five, with higher scores indicating superior performance. The assessor's overall evaluation of the examinee's performance is the primary determinant of the assigned score. The evaluation scale spans a range of performance levels, from excellent, marked by strong command of the subject matter, clear organization, effective delivery, and accurate language use, to poor, characterized by limited knowledge, poor organization, ineffective delivery, and limited linguistic competence.

On the other hand and as previously said, analytic scales offer a structured approach to speaking assessment by breaking down the skill into its constituent components, such as pronunciation, grammar, vocabulary, and fluency. To ensure both the reliability and validity of speaking assessments, a

number of institutions have developed evaluation scales. International frameworks, such as IELTS, ACTFL OPI, and the CEFR offer granular guiding descriptions for language assessment that can be customized to design specific rating scales. For example, The ALTE Can-Do statements offer a framework for language abilities across the CEFR levels, ranging from basic communication skills at A1 to advanced proficiency and complex language use at C2. While C2-level speakers exhibit advanced language proficiency, allowing them to participate in sophisticated conversations and address complex issues, lower-level speakers (A1-B1) are only limited to basic communication and simple interactions (Council of Europe, 2001, pp. 249, 257).

2.4 Test Frameworks

Test frameworks serve as a blueprint for the design and implementation of language assessments. By specifying the purpose, primary audience, task types, time constraints, rating criteria, and other pertinent aspects, the frameworks ensure that assessments are consistent with specific learning objectives and offer dependable and valid measures of language proficiency. As noted by Richards (2012), while not all of the aforementioned specifications may be relevant in every testing context, it is essential to establish clear guidelines. By developing specific criteria, teachers can ensure that speaking assessments align with course objectives and learning outcomes, shaping the direction, approach, and evaluation standards of the assessment.

3. The Methodological Framework

3.1 Participants

This study was conducted with a group of ten (10) English language teachers from Afak Training School. These teachers were selected based on their experience in teaching and assessing speaking skills within the institution. They come from diverse professional backgrounds, with varying years of experience and exposure to different learner levels. Their collective insights provide a rich and realistic view of how speaking skills are assessed at Afak, offering a solid foundation for understanding both the practices and challenges encountered in real classroom settings.

3.2 Methodology

To achieve the aims of this study, a descriptive research method was used. The goal was to closely examine and describe how speaking skills are evaluated by teachers, without influencing or altering their usual practices. Data were collected through a questionnaire, carefully designed to gather both factual information and personal perspectives. The questionnaire combined closed-ended questions, which allowed for clear and measurable data collection, and open-ended questions, which encouraged participants to express their experiences and opinions more freely.

The descriptive approach was chosen because it provides an accurate, detailed snapshot of current assessment methods, making it possible to draw meaningful conclusions without the interference of experimental manipulation.

3.3 Data Analysis

Once the questionnaires were collected, the responses were carefully analyzed. For the closed-ended questions, the results were organized into percentages, which helped in identifying patterns and trends across the participants' answers. By comparing these percentages, it became possible to highlight the most common practices, spot any noticeable differences, and better understand the overall tendencies among the teachers.

Based on these comparisons, conclusions and deductions were made, providing a clearer picture of how speaking skills are currently assessed at Afak Training School. This method of analysis allowed the research to remain both objective and directly linked to the study's main questions.

Evaluating Speaking Assessment Performance: a Case Study of an Accredited Training Center “Afak” at Skikda Province.

Survey Data:

Questions	Frequency	Percentage	Questions	Frequency	Percentage
Gender			Element of speaking ***		
Female	05	50%	Pronunciation	00	00%
Male	05	50%	Grammar	00	00%
Academic qualification			Vocabulary	00	00%
Licence	04	00%	Fluency	01	10%
Master	06	20%	All of them	09	90%
		80%	Approach to rating		
Phd	00	00%	Holistic	03	30%
Area of expertise			Analytical	05	50%
Literature	02	20%	Both	02	20%
			Use of assessment checklists		
Linguistics	04	40%	Always		
Civilization	02	20%	Often	02	20%
Others	02	20%	Sometimes	06	60%
Training			Rarely	01	10%
Yes	05	50%	Never	01	10%
No	05	50%	Emphasis during evaluation		
Teaching experience in speaking			Pronunciation		
1-3	06	60%	Vocabulary	02	20%
3-6	00	00%	Flow	02	20%
6-9	03	30%	Correctness	03	30%
9-12	01	10%	G. impression	03	30%
Significance of speaking			Assessment modality		
Very significant	05	50%	Subject of choice	00	00%
significant	01	10%	With interaction	04	40%
Equally significant	04	40%	Conversation like	06	60%

Not significant	00	00%	Computer utilization***		
Mode of oral instruction			Yes	03	30%
	Interaction	08	No	07	70%
	Performance	01			
	Transaction	01			
	Others				

Nb. A question followed by an asterisk (*) indicates an open-ended query that necessitates an explanatory response.**

3.4 Data Analysis and Discussion

The gender distribution among the participants was perfectly balanced, with 50% male and 50% female teachers. Such parity ensures that the data interpretation is free from gender-related bias and that the findings represent the perspectives of both male and female teachers equally. This balance strengthens the credibility and inclusiveness of the research outcomes, particularly in a field like language assessment where personal perceptions and experiences can vary significantly across demographics.

In terms of academic background, the participants displayed a relatively high level of education: 40% (4 out of 10) held a Licence degree, 60% (6 out of 10) had obtained a Master's degree, and 0% had a PhD.

The absence of PhD holders suggests that the teaching body is largely composed of early to mid-career educators, which is typical for a training school environment. The predominance of Master's degree holders indicates a good academic standing, ensuring that participants have an adequate theoretical foundation for informed practices in assessing speaking skills.

The teaching experience of participants varied, which provided a valuable diversity of viewpoints: 1–3 years of experience: 4 teachers (40%), 3–6 years: 3 teachers (30%), 6–9 years: 2 teachers (20%), 9–12 years: 1 teacher (10%).

The majority of teachers (70%) had less than 6 years of teaching experience, suggesting that relatively early-career perspectives dominate the sample. This could influence the assessment practices, as less experienced teachers might rely

more on structured guidelines (such as rubrics) or institutional policies rather than purely personal judgment. However, the presence of a few more experienced educators (20% with 6–9 years, 10% with 9–12 years) adds depth and practical wisdom to the data.

When asked about their training in teaching and assessing speaking skills: **50%** reported having received formal training, while the other 50% indicated they had not.

This finding is significant. It highlights a split within the teaching staff between those with specialized preparation and those who assess speaking based largely on general teaching experience or intuition. The equal division underlines a crucial point for institutional improvement: while half of the staff is formally equipped, there remains a substantial need for ongoing professional development to standardize assessment practices and ensure fairness.

Participants were asked which elements of speaking they prioritized when assessing students: Fluency alone was chosen by 10% (1 teacher), while 90% (9 teachers) reported assessing all components together: pronunciation, accuracy, fluency, and vocabulary.

This overwhelming emphasis on a holistic view of speaking reflects a mature understanding of communicative competence. Rather than isolating one skill, the majority of teachers acknowledge the multifaceted nature of speaking. However, the small percentage focusing only on fluency raises questions about the potential undervaluation of other critical elements such as accuracy or pronunciation for that one participant.

When describing their general approach to speaking assessment: 3 teachers (30%), reported using a holistic approach, 5 teachers (50%) preferred an analytical approach, and 2 teachers (20%) employed both methods.

The fact that half of the participants adopted an analytical approach suggests a clear inclination towards detailed, criterion-based assessment, where different aspects of performance are scored separately. Meanwhile, those who used a holistic approach possibly prioritize the overall communicative effectiveness

over individual linguistic features. Those using both approaches show a flexible and perhaps more context-sensitive assessment practice.

Overall, the predominance of the analytical method aligns well with the broader trend in language testing literature that emphasizes the need for clear, specific feedback to learners.

Regarding the use of checklists during evaluation: Often: 2 teachers (20%), Sometimes: 6 teachers (60%), rarely: 1 teacher (10%), Never: 1 teacher (10%).

A majority (60%) reported using checklists sometimes, indicating that while assessment tools are recognized, they are not consistently integrated into daily practice.

The fact that only 20% often rely on checklists suggests room for improvement in the use of structured assessment tools, which are known to enhance objectivity and reliability. The occasional or rare use of such tools might expose speaking evaluations to greater subjectivity, depending on the teacher's intuition rather than standardized criteria.

Teachers were also asked which aspect they tend to emphasize most during evaluation:

Pronunciation: 2 teachers (20%), Vocabulary: 2 teachers (20%), Accuracy: 2 teachers (20%), General Impression: 4 teachers (40%).

Interestingly, while attention to specific linguistic aspects (pronunciation, vocabulary, and accuracy) was evenly distributed, a significant portion (40%) emphasized the general impression. This trend suggests that for many teachers, overall communicative effectiveness outweighs strict linguistic precision during assessment. While this can foster communicative confidence among learners, it may also risk neglecting certain structural competencies unless balanced carefully.

In terms of the type of speaking assessment activities used: Interaction-based tasks: 4 teachers (40%), Conversation-based tasks: 6 teachers (60%), Subject-choice tasks: 0 teachers (0%)

The dominance of conversation as a testing modality indicates a preference for realistic, spontaneous speaking situations where learners must demonstrate interactive competence.

Interaction tasks also feature prominently, supporting an emphasis on communication rather than monologic performance (like giving speeches or isolated descriptions).

The absence of subject-choice activities may point to a more controlled testing environment where topics are predetermined by the teacher, perhaps limiting learner autonomy to some extent.

When asked about the use of technology in speaking assessment: 3 teachers reported using computers, while 7 teachers said they did not.

Among those who used computers, it is plausible that they referred to using recording software for oral exams, online assessment platforms, or speech analysis applications to facilitate objective evaluation. For example, they might record students’ oral responses for later review, use pronunciation software to assess accuracy, or conduct oral interviews via online platforms.

The majority (70%) who do not use computers for speaking assessment might cite reasons such as lack of training, insufficient resources, or a preference for traditional face-to-face evaluation methods. This finding reflects a relatively low integration of technology into speaking assessment practices at Afak Training School, highlighting a potential area for future training and resource investment.

Conclusion

This study sheds light on the diverse approaches English language teachers at Afak Training School employ to assess speaking skills. While many educators demonstrate a nuanced understanding of communicative competence by evaluating multiple aspects of speaking, inconsistencies in assessment tools and methods reveal opportunities for refinement. The reliance on experiential judgment, coupled with limited integration of structured frameworks or technology, highlights the need for professional development initiatives to standardize practices.

To enhance the reliability and fairness of speaking assessments, the institution could prioritize teacher training programs, promote the use of rubrics or checklists, and explore technology-assisted evaluation. By fostering a more systematic approach, Afak Training School can better support both educators and learners in achieving consistent, objective, and effective speaking assessments. This study underscores the importance of aligning classroom practices with broader educational goals to ensure meaningful language development.

Bibliography List

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education.
- Celce-Murcia, M. (Ed.). (2001). *Teaching English as a second or foreign language* (3rd ed.). Heinle & Heinle.
- Clark, J. L. D. (1979). Direct vs. indirect language testing. In *Testing Language Proficiency*. Center for Applied Linguistics.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Hornby, A. S. (2000). *Oxford advanced learner's dictionary of current English* (6th ed.). Oxford University Press.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge University Press.
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). <https://doi.org/10.7275/8f8d-rj89>
- Richards, J. C. (2012). *Teaching speaking: Holistic approach*. Cambridge University Press.
- Taylor, L. (2011). *Examining speaking* (Vol. 30). Cambridge University Press. (*Series: Studies in Language Testing*)
- Xi, X., & Mollaun, P. (2006). Investigating the utility of analytic scoring for the TOEFL academic speaking test (TAST). *TOEFL Research Report No. RR-06-04*. Educational Testing Service