EXPLORING ENGLISH TEACHERS' CHALLENGES AND STRATEGIES IN ASSESSING STUDENTS' SPEAKING SKILLS AT SENIOR HIGH SCHOOL

Ina[1]
Muh Nur Akbar Rasyid[2]
Nur Aliyah Nur[3]

[123]Universitas Islam Negeri Alauddin Makassar
E-mail: damingina@gmail.com

ABSTRACT

Speaking assessment is one of the most challenging aspects of English language teaching due to its real-time, interactive, and multidimensional nature. This qualitative case study explored the challenges and strategies of two English teachers in assessing students' speaking skills at MA Madani Alauddin Pao-Pao, South Sulawesi, Indonesia. Data were collected through classroom observation, semi-structured interviews, and document analysis, then analyzed using Miles et al.s (2014) interactive model. Findings revealed three interconnected challenge categories: technical, practical, and pedagogical. Teachers responded with adaptive strategies including multiple assessment formats, task-based activities, analytical rubrics, and continuous formative assessment. The study highlights the persistent gap between speaking assessment theory and classroom practice, and calls for institutional support in the form of professional development, standardized rubrics, and manageable class sizes.

Keywords: speaking assessment, EFL, challenges, strategies, Madrasah Aliyah.

1) INTRODUCTION

Speaking is considered one of the most demanding skills in English language learning, particularly in the context of English as a Foreign Language (EFL) education. In the Indonesian educational context, English is a mandatory subject at all levels, and speaking skills represent one of four fundamental language competencies students must master. Despite its significance, assessing speaking skills presents unique challenges that distinguish it from evaluating other language competencies due to its dynamic, interactive, and multifaceted nature (Fulcher, 2003).

Speaking assessment in English contexts requires teachers to evaluate multiple dimensions simultaneously including fluency, pronunciation, interaction, and coherence — making the process less straightforward and more prone to subjectivity (Brown & Abeywickrama, 2019). Effective language assessment must adhere to fundamental principles such as practicality, reliability, validity, authenticity, and washback. These principles are particularly complex to implement when assessing

oral communication skills, as speaking assessment involves real-time evaluation of spontaneous language production that cannot be easily standardized like written assessments.

Previous studies, including Suwandi, (2023), Putri (2022), Salsabilla (2023), Husna (2020), and Fajar (2023), have examined speaking assessment challenges and strategies in various EFL contexts, but most focus on general classrooms, higher education, or young learners. Few address the specific realities of senior high school teachers, particularly in Islamic secondary schools (Madrasah Aliyah). This study fills that gap by investigating how English teachers at MA Madani Alauddin Pao-Pao navigate speaking assessment challenges and what strategies they employ. The research addresses two questions: (1) What challenges do English teachers face in assessing students' speaking skills? (2) What strategies do they employ to overcome these challenges?

2) METHODS

This study used a descriptive qualitative case study design (Arikunto, 2019), appropriate for exploring complex phenomena in their natural context. The research was conducted at MA Madani Alauddin Pao-Pao, Gowa Regency, South Sulawesi, during the 2025/2026 academic year. Two English teachers were selected through purposive sampling, the school's only two English teachers, chosen for their professional experience and direct involvement in speaking assessment.

Purposive sampling was used to select two English teachers as research participants. The reason for selecting only two teachers was that the school has only two English teachers available. These teachers were chosen based on their professional experience in teaching English and their active involvement in assessing students' speaking skills. They were considered to possess sufficient knowledge and firsthand experience of the challenges and strategies involved in speaking assessment. The study did not focus on the number of participants but rather on the depth and richness of the data obtained.

Data were collected via three instruments: (1) non-participant classroom observation using field notes; (2) semi-structured interviews conducted in Indonesian to facilitate depth and authenticity (Creswell, 2014), audio-recorded with consent; and (3) document analysis of rubrics, lesson plans, and assessment records. Data were analyzed using Miles et al.s (2014) interactive model through three stages: condensation, display, and conclusion drawing. Trustworthiness was ensured through source triangulation, cross-checking interview data, observation notes, and documents.

3) RESULTS

Findings are organized into three challenge categories and four strategy categories, summarized in Table 1 and Table 2.

Table 1. Summary of Challenges

Challenges Category	Sub-Challenges	Frequency Level
Technical Challenges	Construct definition	High
	Multiple competencies	High
	Reliability issues	Medium*
	Subjectivity in scoring	Medium
Practical Challenges	Time constraints	High*
	Large class sizes	High*
	Limited resources and facilities	Medium
	Administrative burden	Medium
Pedagogical Challenges	Assessment alignment	Medium*
	Authenticity vs standardization	Medium
	Student anxiety	High*
	Effective and timely feedback	Medium

Technical Challenges

Teachers shared an understanding of speaking ability centered on fluency and pronunciation but differed in application: MA prioritized student confidence and communicative risk-taking, while CM focused primarily on pronunciation accuracy. This inconsistency in construct definition created validity concerns (Bachman & Palmer, 1996). Both teachers also struggled to evaluate multiple competencies simultaneously (grammar, vocabulary, fluency, and pronunciation) under real-time conditions, leading to selective attention and compensatory judgment. Reliability was further compromised by the absence of formal rater training (both teachers confirmed having received none) and by individual adjustments to rubric criteria based on perceived student ability levels.

Practical Challenges

Time constraints combined with class sizes of 30–50 students made individual speaking assessment logistically unfeasible within single class periods. Group-based tasks were adopted as

adaptations. Technology use was minimal MA used personal voice notes for delayed review; CM used no recording tools. Administrative responsibilities further reduced time available for detailed assessment planning and documentation.

Pedagogical Challenges

Speaking tasks were generally aligned with curricular objectives, though alignment was sometimes superficial, integration into multi-skill lessons reduced focused evaluation of oral competence. Student anxiety was a high-frequency challenge (particularly for MA): observations recorded students pausing for extended periods and appearing visibly nervous. Feedback was predominantly formative and immediate but generalized (e.g., "Good, but improve your pronunciation"), lacking specificity or follow-up mechanisms to maximize developmental impact.

Table 2. Summary of Strategies

Strategies Category	Sub-Strategies	Frequency Level
Assessment Design	Multiple assessment	High
	Task-based assessment approaches	High
	Authentic tasks	Medium*
	Integrating assessment	Medium*
Scoring and Evaluation	Analytical rubrics	High
	Multiple raters	Low*
	Peer-assessment	—
	Self-assessment	—
Implementation	Time management	High
	Technology use	Low*
	Clear criteria	High
	Rater training	—
Feedback and Improvement	Constructive feedback	High
	Practice refinement	High
	Continuous assessment	High
	Portfolio approaches	Low

Teachers used multiple assessment formats (individual oral tasks, role-plays, and group discussions) to capture different aspects of speaking competence. Task design incorporated authentic, personally relevant contexts alongside standardized textbook materials. Analytical rubrics focusing on fluency, pronunciation, and occasionally grammar were applied, though selectively. Continuous formative assessment embedded in instruction was the dominant approach. Peer and self-assessment were not implemented due to concerns about negative peer interactions; portfolio documentation was absent.

4) DISCUSSION

The most prominent challenges construct inconsistency, simultaneous multi-competency evaluation, student anxiety, and structural time/class-size constraints are consistent with Brown & Abeywickrama (2019) and Fulcher's (2003) characterizations of speaking assessment as inherently complex. The variation in construct definition between the two teachers extends Suwandi's (2023) findings by showing how differing conceptual frameworks directly shape classroom evaluation practices within the same institution.

Reliability concerns arising from absent rater training echo Luoma's (2004) warning about rater variability as a primary source of measurement error. Importantly, this study shows that such problems occur equally in face-to-face settings, not only in remote teaching contexts, as Fajar (2023) discussed, arising from structural deficits in professional development rather than from technological limitations.

Student anxiety as a mediating variable in assessment validity extends Horwitz et al.'s (1986) theory to the Indonesian secondary school context. The pedagogical dilemma is clear: interventions that reduce anxiety (note allowances, reassurance) may alter task demands and reduce construct validity. Teachers must continuously balance emotional support with assessment rigor.

Teachers' adaptive strategies demonstrate contextual intelligence, prioritizing feasibility while preserving instructional value. However, the absence of peer/self-assessment, portfolio documentation, multiple raters, and rater training represents a significant gap between theory and practice. These are not failures of awareness but of structural support: teachers lack the institutional conditions (time, training, resources) needed to implement more comprehensive assessment systems.

This study is limited to two teachers at one school during a single academic year and excludes student perspectives. Future research should involve larger, more varied samples; incorporate student voices; examine the effect of targeted interventions such as rater training; and employ longitudinal designs to track how speaking assessment practices evolve over time.

5) CONCLUSION

This study found that English teachers at MA Madani Alauddin Pao-Pao face interconnected technical, practical, and pedagogical challenges in assessing speaking skills, most critically, inconsistent construct definition, structural constraints from large classes and limited time, and student anxiety. In response, teachers employed adaptive strategies centered on multiple assessment formats, task-based design, analytical rubrics, and integrated formative assessment. The findings reveal a persistent gap between speaking assessment theory and classroom reality, rooted in systemic, institutional factors rather than individual teacher shortcomings.

Teachers are encouraged to systematically document student progress and gradually introduce peer and self-assessment under structured conditions. Educational institutions should provide professional development in speaking assessment, standardized rubrics, and manageable class sizes. Policymakers should develop practical assessment frameworks that reflect classroom realities. These steps are essential to improve the fairness, reliability, and developmental impact of speaking assessment in similar EFL contexts.

REFERENCES

Al., H. et. (1986). Foreign Language Classroom Anxiety. The Modern Language Journal, Vol. 70, N.

Arikunto, S. (2019). "Research Procedures". Jakarta: PT Rineka Cipta.

Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice: Designing and developing useful language tests. Oxford University Press., Vol. 1.

Brown, H. D., & Abeywickrama, P. (2019). Language Assessment: principles and classroom practices (Third Edit). Pearson.

Creswell, J. W. (2014). Reseach Design: Pendekatan, Kualitatif, Kuantitatif, dan Mixed. Yogyakarta: Pustaka pelajar.

Fajar, L. (2023). An Analysis on The Difficulties of Assessing English Speaking Skills in Emergency Remote Teaching. https://repository.uinjkt.ac.id/dspace/handle/123456789/73672

Fulcher, G. (2003). Testing second language speaking. England: Pearson Longman.

Husna, N. (2020). Exploring the Assessment in Speaking Classroom (Vol. 2507, Issue February). https://repository.ar-raniry.ac.id

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

Miles, M.B, Huberman, A.M, &, & Saldana, J. (2014). Qualitative Data Analysis, A. Methods Sourcebook, Edition 3. In Sustainability (Switzerland) (Vol. 11, Issue 1). USA: Sage Publications.

Putri, R. E. (2022). Teachers' Strategies in Conducting Speaking Assessments for Young English Learners At Mis Nurul Huda, Tanjung Morawa. Warta Dharmawangsa, 16(3), 286–298. https://doi.org/10.46576/wdw.v16i3.2224

Salsabilla, A. (2023). Teachers Chalenges in Assessing English Young Learners Speaking Skills Thesis Antasari State Islamic Universities. https://idr.uin-antasari.ac.id/23928/

Suwandi. (2023). the Assessment of Students' Speaking Skills in Indonesian Classrooms. International Journal of Social Science, 2(6), 3465–3470. https://doi.org/10.53625/ijss.v2i6.6103