

<https://doi.org/10.34142/2709-7986.2025.30.1.08>

EXPLORING CHALLENGES IN SPOKEN COURSES EVALUATION IN AUTHENTIC ASSESSMENT AT WOLLO UNIVERSITY IN ETHIOPIA

ВИВЧЕННЯ ПРОБЛЕМ ЕКСПЕРТИЗИ УСНИХ КУРСІВ В ОЦІНЮВАННІ В УНІВЕРСИТЕТІ ВОЛЛО В ЕФІОПІЇ

Received: 19/01/2025


Accepted: 24/02/2025

Ayalew YIMER¹, & Taye REGASSA²



¹Foreign Languages and Literature, Lecturer (M.A) HLJC, Wollo University, Dessie, Ethiopia.


✉ E-Mail: ayalewyimer665@gmail.com

 <https://orcid.org/0009-0009-8661-0574>



²Ph.D., Foreign Language and Literature, Assistant Prof. (Ph.D.) HLJC, Addis Ababa University, Addis Ababa, Ethiopia.

✉ E-Mail: taye12regassa@gmail.com

 <https://orcid.org/0009-0002-2837-4912>

How to Cite (APA Style):

Yimer, A., & Regassa, T. (2025). Exploring challenges in spoken courses evaluation in authentic assessment at Wollo University in Ethiopia. *Educational Challenges*, 30(1), 119–137. <https://doi.org/10.34142/2709-7986.2025.30.1.08>

ABSTRACT

Purpose. The current study aims to explore problems with the introduction of authentic speaking assessments in English as a Foreign Language (EFL) courses at Wollo University, Ethiopia.

Methodology. Methods data were collected from 25 ELT teachers, and 30 EFL students at the end of their ELT class via structured questionnaires, semi-structured interviews, and document reviews. Quantitative data

Мета. Це дослідження розглядає виклики впровадження надійного оцінювання усного мовлення в аудиторіях з англійської як іноземної мови (EFL) в Університеті Волло (Ефіопія).

Методи. Використовуючи змішаний методологічний підхід, було зібрано дані від 25 викладачів англійської мови та 30 студентів випускного курсу за допомогою структурованих анкет, напівструктурованих інтерв'ю та аналізу

© Ayalew Yimer, & Taye Regassa, 2025

Exploring challenges in spoken courses evaluation in authentic assessment at Wollo University in Ethiopia © 2025 by Ayalew Yimer Assen, & Taye Regassa Seda is licensed under **Attribution-NonCommercial 4.0 International**. To view a copy of the license, visit <http://creativecommons.org/licenses/by-nc/4.0/>

were statistically analyzed, and qualitative data were thematically analyzed to identify deficiencies in assessment practice.

Results. The results identified challenges including the absence of task-relevant context, variable criteria of assessment, insufficient resources, limited teacher professionalization, and factors influencing students (e.g., anxiety, lack of feedback). On the other hand, teachers have a huge challenge to ensure that the different learners have their assessment needs satisfied, in terms of the relationship between assessment procedures and real-life communication. Meanwhile the students also have limited opportunities to speak and limited chances to receive valuable feedback. Challenges include time limitations, professional development limitations, and the absence of classroom resources. Qualitative findings identify the critical need for culturally appropriate assessments i.e., to recognize the language plurality in Ethiopia.

Conclusions. The mismatch between the assessment method and students' requirement for communicative course work also is limiting. Due to the demand of genuine, equitable and reliable assessment tools to enhance students' spoken English, the current study recommends specific teacher training to address these issues, including life size assessment in assessment, and effective exploitation of resources.

Keywords: authentic speaking assessment, traditional assessment, EFL challenges, inclusive approaches, linguistic differentiation.

документів. Кількісні дані аналізувалися статистичними методами, а якісні – тематично, щоб визначити прогалини в оцінюванні.

Результати. Результати дослідження підкреслюють такі виклики, як відсутність контекстно релевантних завдань, непослідовні критерії оцінювання, нестача ресурсів, обмеженість дій викладачів і студентські проблеми, зокрема тривожність та недостатній зворотний зв'язок. Викладачі стикаються з труднощами в оцінюванні реальних комунікативних потреб, тоді як студенти відчують брак практичних можливостей для усного мовлення та неефективний зворотний зв'язок. Серед основних викликів – обмеженість часу, недостатній фаховий розвиток і відсутність належного обладнання в аудиторіях. Якісні результати наголошують на необхідності змін у процесі оцінювання з урахуванням культурного й лінгвістичного різноманіття Ефіопії.

Висновки. Невідповідність між методами оцінювання та комунікативними потребами студентів обмежує їхню готовність до взаємодії в реальному світі. Для покращення мовленнєвих навичок студентів англійською мовою необхідні достовірні, автентичні, справедливі та надійні інструменти оцінювання. У дослідженні рекомендовано цілеспрямоване навчання викладачів, інтеграцію життєво-ситуативних сценаріїв у процес оцінювання та ефективніший розподіл ресурсів.

Ключові слова: автентичне оцінювання мовлення, традиційне оцінювання, виклики викладання англійської як іноземної, інклюзивні практики, лінгвістичне різноманіття.

INTRODUCTION

The evaluation of a student's speaking ability in English as a Foreign Language (EFL) speaking classrooms is complicated by teaching methods, cultural environments, and resource constraints. Comprehensive assessment should consider fluency, accuracy, prosody, vocabulary, and interactional abilities, yet they may be too subjective. English fluency is critically important at Wollo University in Ethiopia. It leads to students' academic and occupational success; thus, effective assessment of English fluency is a pressing issue.

Unfortunately, the current approaches are subject to rote memorization at the expense of faculty-student interaction and undermine the development of students' real-life communication skills. As reported by Harris (2018) and Alderson, Clapham, and Wall (2019), assessment tasks must be realistic to be valid and reliable. Yet, by virtue of its

linguistic heterogeneity and uneven distribution of resources in Ethiopia, implementing certain practices raises some challenges. Inclusive and fair tests are required to meet the students' multiple linguistic backgrounds.

Preparing for speaking assessments is further complicated by the lack of teacher preparation and the usage of outmoded assessment techniques. The absence of resources (e.g., materials and tools) additionally restricts the educator and the student. Furthermore, students suffer from anxiety and insufficient feedback, both of which have detrimental effects on their performance. Studies highlight how it is essential, among other things, to design positive learning and teaching environments that help alleviate stress to achieve better results.

Current assessment models in Ethiopia are inadequate with respect to contemporary communicative needs because they lack genuine speaking opportunities and show cultural bias. The aim of this study is to collect opinions from Wollo University (students and teachers), to analyze spoken assessment as it is being conducted, to identify gaps, and suggest an improvement.

Contextualized evaluation paradigms, in which local factors are integrated with global ones, are critical for promoting communicative competence. Through the enhancement of teacher education, the creation of real tasks in a natural environment, and by addressing the shortages of resources, teachers and staff at Ethiopian universities institutions will be better able to provide students with speaking skills which will support success in their studies and careers.

The current study aims to explore problems with the introduction of authentic speaking assessments in English as a Foreign Language courses at Wollo University, Ethiopia. The goals for the research are as follows.

1. To describe the major challenges of EFL teachers and learners in the assessment of speaking abilities at Wollo University.
2. To investigate EFL teachers and learners' perceptions about the utility and impartiality of present speaking ability assessment procedures.
3. To offer strategies regarding the authenticity and efficacy of speaking skill assessment in EFL education of Wollo University.

LITERATURE REVIEW

Language assessment is a highly important component of assessing linguistic ability and informing practice. Researchers have also explored various aspects of language assessment, such as validity, reliability, communicative competence or as a function of contextual factors. This review summarizes major theories by influential researchers in methods of language assessment, background of theoretical foundations, and problem areas of implementation.

Bachman (1990) built a foundational framework for language testing centering on questions of validity and reliability in assessment design. He claimed that tests should measure not only linguistic knowledge, but also cognitive and contextual aspects influencing performance. Bachman and Palmer (2010) also extended this approach by defining the construct of test usefulness, which consists of reliability, construct validity, authenticity, interactiveness, effect, and practical significance. They argued that an

assessment of language should, for this reason, also to consider these issues in such a way as to ensure fairness and meaningful assessment. Their research continues to have a very important role in the development of current assessment practices but for these practices to favour an holistic approach which “counts” authentic use of language.

Communicative competence has been the core issue in the context of language assessment studies. Canale and Swain (1980) introduced a model of communicative competence comprising grammatical, sociolinguistic, discourse, and strategic competencies. Their model influenced later assessment practices that prioritize real-world communication over rote memorization. Brown and Abeywickrama (2010) have also put these propositions into words, and stressed the value of performance-based tests, such that students are assessed in real communicative situations rather than on isolated units of language. Similarly, McNamara (2009) argued that traditional testing approaches often fail to capture the dynamic nature of language use and advocated for performance assessments that mirror actual communicative interactions.

The effect of assessment on learning has also been extensively investigated. Black, and Wiliam (1998) highlighted the value of formative assessment for improving the learning of students. They argued that assessment for true learning must be continuous and should not just give them a mark but should give them some useful feedback. Their research demonstrated that formative assessment techniques, such as self-assessment and peer feedback, contribute to deeper learning and improved language proficiency. This aligns with the perspective that authentic assessments are essential for evaluating students through meaningful, contextually relevant tasks, enabling them to apply their language skills beyond the classroom (Sakuda, 2023; Dolgin et al. 2010).

The difficulties in the evaluation of speaking ability have been a topic of much debate. Luoma (2004) also noted that, by their very subjective design and the lack of standardization of evaluation criteria, speaking evaluations are inherently complex. She proposed with the application of specific rubrics and multiple raters in the assessment of speaking tests that speaking tests would be improved in reliability. Hughes (2003) further addressed this, suggesting formalized rating scales in order to promote homogeneity in the process of oral language assessment. Orr (2002) also studied the ability of speaking tests and stressed the importance of high-quality training for assessors of exactly how to discern fluency, coherence and pronunciation characteristics in spoken tests.

Cultural and contextual factors further complicate language assessment. Shohamy (2001) also critiqued the power relations embedded in language measurement, pointing out the tendency among standardized tests to further disadvantage language minority speakers (i.e. Hofstede, 2001) supported this idea also when discussing how cultural influences shape expectations and results of language assessment. Kumaravadivelu (2008) also went on to extend these notions to argue for postmethod assessment methods to recognize such different linguistic backgrounds and communicative contexts. These authors jointly emphasize the need for culturally sensitive assessment approaches considering linguistic variation and the tendency to perpetuate systemic inequities.

Overall, language assessment research has changed from the traditional dominant testing model to one that includes communicative, formative, and culturally relevant models. Bachman and Palmer (2010), Canale and Swain (1980), and Black and Wiliam (1998) have influenced the field by emphasizing assessments that are valid, reliable, and

of value to learning. Nevertheless, issues still exist in achieving fairness, particularly for speaking evaluations and cultures. Additional research should further look into new assessment approaches that are integrating technology and are sensitive to the evolving needs of global communication

– **Theoretical Foundations of Authentic Assessment**

Authentic tasks, based on constructivist theories, put knowledge application at the center of a functional process (Nunan, 2004). Methods like role plays and interviews enable students to demonstrate communicative abilities in problem solving. These approaches align with active learning principles, acknowledging the inseparability of speech and learning. Nevertheless, attaining validity and reliability of speaking assessments continues to be an issue. Validity guarantees that the tools can measure communicative competence, while reliability guarantees a consistent and uniform scoring. Speaking evaluations are subjective, and standardization is also challenging (Weir, 2005). Further, resource limitations, large class sizes, and restrictions on individualized assessments present major challenges.

– **Complexities in Speaking Assessment**

Fluency, accuracy, pronunciation, and discourse management are assessed in speaking (McNamara, 2009). Creating balanced rubrics has also been a formidable task in resource poor contexts (Brown & Abeywickrama, 2010). The lack of qualified teachers negatively impacts the variability of assessment processes (Hughes, 2003). These challenges are particularly acute in the Global South, where both a shortage of time, limited resources, and deficient infrastructure complicate robust evaluation.

– **Learners' Perspectives**

For undergraduate students, authentic learning tasks such as role plays and discussions are meaningful and engaging. The exercises fill the gap between lesson work and real-world communication (Veretiuk & Chetveryk, 2024, p. 139). Yet, high-stakes authentic evaluations also have the potential to cause anxiety and consequently decrease student achievement (Shohamy, 2001). The openness and unpredictability of real-world situations in assessments could lead to this anxiety (McNamara, 2009).

– **Bridging Theory and Practice**

Despite their value, authentic assessments face barriers such as large class sizes, limited technology, and inadequate teacher training. Cultural norms and institutional policy often conflict with CLT principles. Making them work is difficult. These challenges emphasize the need for culturally based, individualized assessment approaches that consider practical limitations.

Authentic assessment research in Ethiopian universities holds promise in addressing and potentially overcoming these challenges. When addressing resource constraints, training gaps, and contextual barriers, teachers can further assess practices to promote communicative competence and to reduce biases.

METHODOLOGY

– **Research Design and Methodology**

This research used a descriptive explanatory survey designed to investigate the difficulties and avenues of EFL teachers and students' assessment of speaking ability at

Wollo University. This design is suitable since it is possible to conduct a thorough investigation of the subject matter by comprehensively including quantitative and qualitative data collection techniques. The mixed-method design consists of systematically administered questionnaires, semi-structured interviews, a focus group, and document analysis.

A sample of 30 EFL teachers and 25 undergraduate students was selected through comprehensive sampling to ensure representation of individuals with relevant experiences in speaking skill assessment.

Instruments for data collection are designed based on the aims of the study, for example, pointing out difficulties, facilitating fairness and effectiveness, and looking into how to improve. The questionnaires used contain both quantitative and qualitative questions. Moreover, semi-structured interviews with teachers and focus group discussions with students provide deeper reflections on subjective experiences and difficulties encountered with speaking skill assessments. Quantitative data are described and tested with descriptive and inferential statistics, whereas qualitative data are investigated with thematic analysis to find rare themes and repeated patterns.

Ethical considerations are prioritized throughout the research. Participation is voluntary, informed consent is acquired, and privacy is protected to foster trust and integrity. The research seeks to deliver practical advice on how to best assess speaking competence in Ethiopian EFL settings to contribute to culturally appropriate, representative practices that take into account Ethiopia's linguistic diversity, ensuring fairness for all learners

– **Data Collection Instruments and Procedures**

The study employs multiple instruments for data collection. Questionnaires with five-point Likert scales measure quantitative data related to perceptions, behavior, and barriers. Semi-structured interviews with five teachers and five students add to this, providing qualitative detail. Document analysis of three years of English exam papers provides further evidence, evaluated against a standardized checklist aligned with the study's goals.

Ethical approval was secured from the university's review committee. Data collection involved administering questionnaires over two weeks, conducting audio-recorded interviews with consent, and performing discreet document analysis. Triangulation of data sources (i.e., questionnaires, interviews, and document analysis) provides for validity and reliability. Pilot testing of the instruments further enhances their robustness. Quantitative data are coded and analyzed using SPSS, while qualitative data undergo thematic analysis, ensuring a comprehensive exploration of the research objectives.

RESULTS AND DISCUSSION

The results are anticipated to generate valuable interventions for teachers to improve their own strategies and for students to reduce anxiety and enhance motivation, leading to overall academic improvement. Secondly, the paper provides relevant suggestions to policymakers and curriculum makers, concerning the assessment practices matching with the Ethiopian national education objectives and the international communication requirements. In this study's effort to fill research voids within the Ethiopian EFL context, this paper makes academic contributions to the literature and offers, as per best

practices, concrete suggestions for enhancing the effectiveness and efficiency of English language teaching.

– **Quantitative Analysis**

In order to analyze the survey data, descriptive analysis (frequency, percentages, and mean) was performed to summarize the data while inferential statistics was used to assess using SPSS (version 26).

– **Qualitative Data Analysis**

Regarding the qualitative data obtained from the interviews and observations, thematic analysis was used to identify common themes and patterns within the transcribed data. The coding process involved several stages: open-coding, axial coding, and selective coding. After coding, the data were categorized into various themes based on similarities among the codes. This was supported by NVivo (version 11). The results were arrived at by triangulating the results of the quantitative and qualitative data.

– **Results and discussion from the quantitative data**

Table 1

Perception of teachers about the assessment of speaking skills.

Nº	Question	Strongly Agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly Disagree (%)
1	The teaching resources available at the university are sufficient for effective instruction.	5 (20%)	10 (40%)	7 (28%)	3 (12%)	0 (0%)
2	I have received adequate training to implement modern teaching methodologies.	8 (32%)	9 (36%)	5 (20%)	3 (12%)	0 (0%)
3	The syllabus aligns with the needs and interests of students.	6 (24%)	12 (48%)	4 (16%)	3 (12%)	0 (0%)
4	There is enough time allocated to cover all the topics in the syllabus.	2 (8%)	8 (32%)	6 (24%)	7 (28%)	2 (8%)
5	Students actively participate in classroom activities.	5 (20%)	13 (52%)	4 (16%)	3 (12%)	0 (0%)
6	Assessments accurately measure students' language skills.	4 (16%)	12 (48%)	6 (24%)	2 (8%)	1 (4%)
7	The professional development	3 (12%)	11 (44%)	5 (20%)	4 (16%)	2 (8%)

	opportunities provided are beneficial.					
8	There is adequate support from administration to address teaching challenges.	2 (8%)	9 (36%)	8 (32%)	5 (20%)	1 (4%)
9	The integration of technology in the classroom has improved learning outcomes.	5 (20%)	13 (52%)	4 (16%)	2 (8%)	1 (4%)
10	Collaborative practices among teachers enhance teaching quality.	6 (24%)	14 (56%)	3 (12%)	2 (8%)	0 (0%)

The results of the survey provide insight into various aspects of teaching resources, training, curriculum alignment, time management, student participation, assessment accuracy, professional development, administrative support, technology integration, and collaborative teaching practices.

Regarding the availability of teaching resources, 60% of respondents (20% strongly agree, 40% agree) believe that the resources provided by the university are sufficient for effective instruction. However, 28% remained neutral, and 12% disagreed, indicating that while most instructors find the resources adequate, a sizable portion perceives room for improvement. Similarly, 68% (32% strongly agree, 36% agree) reported receiving adequate training in modern teaching methodologies, though 20% were neutral and 12% disagreed, suggesting a need for further professional development opportunities.

In terms of curriculum alignment, 72% of respondents (24% strongly agree, 48% agree) agreed that the syllabus aligns with student needs and interests. However, 16% were neutral and 12% disagreed, which may indicate that some areas of the curriculum require revision to better address diverse learning needs. Time management emerged as a significant concern, with only 40% agreeing that enough time is allocated to cover all topics, while 24% were neutral and 36% (28% disagree, 8% strongly disagree) expressed dissatisfaction. This finding highlights potential challenges in curriculum pacing and instructional time allocation.

Student engagement in classroom activities was positively perceived, with 72% of respondents (20% strongly agree, 52% agree) acknowledging active students' participation. However, 16% remained neutral, and 12% disagreed, which suggests that engagement strategies could be further enhanced. Additionally, 64% (16% strongly agree, 48% agree) believed that assessments effectively measure students' language skills, though 24% were neutral and 12% (8% disagree, 4% strongly disagree) expressed concerns about the accuracy of assessment tools.

Professional development opportunities received mixed reviews, with 56% (12% strongly agree, 44% agree) finding them beneficial, but 20% neutral and 24% (16% disagree, 8%

strongly disagree) expressing concerns. This suggests that while training is available, its effectiveness may vary among faculty members. Administrative support was also seen as an area for improvement, with only 44% agreeing that adequate support is provided, while 32% were neutral and 24% (20% disagree, 4% strongly disagree) disagreed, indicating a potential gap in institutional support mechanisms.

Technology integration was viewed favorably, with 72% (20% strongly agree, 52% agree) agreeing that it has improved learning outcomes, though 16% were neutral and 12% disagreed. This finding suggests that while technology is beneficial, its implementation may not be uniform across all classrooms. Lastly, collaborative teaching practices were seen as highly beneficial, with 80% (24% strongly agree, 56% agreeing) agreeing that they enhance teaching quality, while only 12% were neutral and 8% disagreed. This underscores the importance of fostering a collaborative teaching environment to improve instructional effectiveness.

Overall, the findings highlight several strengths, such as student engagement, curriculum alignment, and technology integration, while also identifying areas for improvement, including time management, assessment accuracy, professional development, and administrative support. Addressing these concerns could enhance teaching and learning outcomes within the institution.

Table 2
Relevance, Effectiveness, and Course Materials

Objective	Positive Responses (No. & %)	Neutral/Negative Responses (No. & %)	Mean
Relevance of English Courses	20 (83.33%)	5 (16.67%)	1.23
Effectiveness of Teaching Methods	18 (70%)	7 (30%)	1.47
Adequacy and Usefulness of Materials	15 (60%)	10 (40%)	2.00

The results presented in Table 2 provide insights into the perceived relevance of English courses, the effectiveness of teaching methods, and the adequacy and usefulness of course materials. While most respondents provided positive feedback, there are notable areas where improvements may be needed.

The findings indicate that 83.33% of respondents perceive English courses as relevant, with a mean score of 1.23. This suggests that the courses align with students' academic and professional needs. Previous studies highlight the importance of curriculum alignment with students' real-world applications, emphasizing that English for Specific Purposes (ESP) courses should be designed to meet industry and academic demands (Hutchinson & Waters, 1987). However, 16.67% of respondents who provided neutral or negative feedback suggest that some learners may find certain aspects of the curriculum less applicable to their needs. This aligns with research indicating that generic language courses may fail to fully address students' specific requirements (Basturkmen, 2010). Therefore, continuous curriculum evaluation and adaptation are necessary to ensure that course content remains relevant and applicable.

The effectiveness of teaching methods received a positive response from 70% of participants, with a mean score of 1.47. However, 30% of respondents expressed neutral or negative opinions, suggesting that some teaching strategies may not fully engage students or facilitate effective learning. Research has shown that student-centered and interactive teaching methodologies, such as communicative language teaching (CLT), can significantly enhance learning outcomes (Richards & Rodgers, 2014). Traditional lecture-based approaches, while still widely used, may not be as effective in promoting active learning, critical thinking, and language retention (Brown, 2007). Therefore, integrating more interactive and student-centered strategies, including task-based learning and technology-enhanced instruction, could help improve student engagement and overall effectiveness.

The lowest positive response rate (60%) was observed for the adequacy and usefulness of course materials, with 40% of respondents expressing neutral or negative views and a mean score of 2.00. This suggests that while some learners find the materials beneficial, a sizable portion believes there is room for improvement. According to Tomlinson (2012), well-designed instructional materials play a crucial role in language acquisition by providing meaningful, engaging, and contextually relevant content. The findings indicate that some students may perceive the current materials as outdated, insufficient, or lacking in relevance to their learning needs. Research supports the need for updated and authentic materials that reflect real-world usage, incorporating multimedia resources and digital platforms to enhance engagement (Mishan & Timmis, 2015). Therefore, revising course materials to include more contemporary and context-specific resources could enhance their effectiveness.

Table 3

Classroom Facilities, Speaking Practice, and Feedback

Nº	Question	Strongly Agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly Disagree (%)
1	The teaching resources available at the university are sufficient for effective instruction.	5 (20%)	10 (40%)	7 (28%)	3 (12%)	0 (0%)
2	I have received adequate training to implement modern teaching methodologies.	8 (32%)	9 (36%)	5 (20%)	3 (12%)	0 (0%)
3	The syllabus aligns with the needs and interests of students.	6 (24%)	12 (48%)	4 (16%)	3 (12%)	0 (0%)
4	There is enough time allocated to cover all the topics in the syllabus.	2 (8%)	8 (32%)	6 (24%)	7 (28%)	2 (8%)

5	Students actively participate in classroom activities.	5 (20%)	13 (52%)	4 (16%)	3 (12%)	0 (0%)
6	Assessments accurately measure students' language skills.	4 (16%)	12 (48%)	6 (24%)	2 (8%)	1 (4%)
7	The professional development opportunities provided are beneficial.	3 (12%)	11 (44%)	5 (20%)	4 (16%)	2 (8%)
8	There is adequate support from administration to address teaching challenges.	2 (8%)	9 (36%)	8 (32%)	5 (20%)	1 (4%)
9	The integration of technology in the classroom has improved learning outcomes.	5 (20%)	13 (52%)	4 (16%)	2 (8%)	1 (4%)
10	Collaborative practices among teachers enhance teaching quality.	6 (24%)	14 (56%)	3 (12%)	2 (8%)	0 (0%)

Table 3 presents data on classroom facilities, speaking practice, and feedback, highlighting key areas of strength and aspects requiring improvement. The findings suggest that while many aspects of teaching resources, instructional methods, and assessment practices are positively perceived, challenges remain in administrative support and time allocation for syllabus coverage.

The availability of sufficient teaching resources received a positive response, with 60% of participants (20% strongly agree, 40% agree) indicating adequacy. However, 28% remained neutral, and 12% disagreed. This finding aligns with previous research emphasizing the importance of well-equipped classrooms in enhancing student learning outcomes (Richards, 2015). Adequate teaching resources, including textbooks, multimedia tools, and access to digital platforms, are essential for effective instruction (Tomlinson, 2012). The presence of neutral and negative responses suggests that some instructors may experience resource limitations, which could hinder teaching effectiveness.

A significant majority (72%) agreed that the syllabus aligns with student needs, with a mean response leaning toward positive perceptions. Curriculum alignment plays a crucial role in student engagement and learning motivation (Nation & Macalister, 2010). However, the results also indicate that 16% were neutral, and 12% disagreed, implying that curriculum revisions may be necessary to better cater to diverse learner needs. Similarly, while 68% of respondents felt adequately trained in modern teaching methodologies, 20%

were neutral, and 12% disagreed. Research suggests that ongoing teacher training is critical for adapting to evolving pedagogical approaches (Freeman, 2017).

The results indicate strong agreement that students actively participate in classroom activities (72% positive responses). This supports the argument that interactive learning methods, such as communicative language teaching (CLT), enhance student engagement (Richards & Rodgers, 2014). However, 16% of participants remained neutral, and 12% disagreed, suggesting that participation levels may vary depending on class structure, teaching strategies, or student motivation.

Regarding assessment accuracy, 64% of respondents believed that current evaluation methods effectively measure students' language skills. However, 24% were neutral, and 12% (8% disagree, 4% strongly disagree) expressed concerns. Research highlights the importance of using diverse assessment tools, including formative and summative evaluations, to ensure comprehensive language proficiency measurement (Brown, 2004). The presence of neutral and negative responses suggests that certain assessment practices may need refinement to enhance fairness and reliability.

Adequate administrative support received one of the lowest approval ratings, with only 44% agreeing that sufficient support is provided. A significant 32% were neutral, while 24% expressed dissatisfaction. This finding aligns with studies emphasizing the critical role of institutional support in improving teaching effectiveness and reducing instructor workload (Farrell, 2018). Similarly, professional development opportunities were viewed positively by 56% of respondents, while 24% expressed dissatisfaction. Effective teacher training programs have been shown to significantly improve instructional quality and student learning experiences (Day, 2013). Institutions should consider expanding training programs and providing additional administrative support to address faculty concerns.

Technology integration in the classroom was viewed favorably by 72% of respondents, reinforcing previous research indicating that digital tools enhance language learning experiences (Chapelle & Sauro, 2017). However, 16% were neutral, and 12% disagreed, suggesting that access to or effective implementation of technology may still be a challenge. Finally, collaborative teaching practices received the highest approval rating, with 80% agreeing that they enhance teaching quality. Collaborative learning has been widely recognized as an effective method for professional development and knowledge sharing among educators (Lave & Wenger, 1991).

Table 4
Peer Discussions, Technology Support, Assessment Methods, and Overall Satisfaction

Objective	Positive Responses (%)	Neutral/Negative Responses (%)	Mean
Impact of Peer Discussions and Group Activities	22 (73.33%)	8 (26.67%)	1.33
Support from Technology Integration	21 (70%)	9 (30%)	1.47
Assessment Methods Reflecting Proficiency	19 (63.33%)	11 (36.67%)	1.87

Overall Satisfaction with English Courses	25 (83.33%)	5 (16.67%)	1.23
---	-------------	------------	------

Table 4 presents findings on peer discussions, technology support, assessment methods, and overall satisfaction with English courses. The results suggest that while students hold positive perceptions in these areas, assessment methods and technology integration could benefit from further enhancements.

Most respondents (73.33%) agreed that peer discussions and group activities positively impact their learning experience, while 26.67% remained neutral or disagreed. The mean score of 1.33 indicates a favorable perception of collaborative learning. Research highlights that peer interactions play a crucial role in language acquisition, as they promote communication skills, social learning, and critical thinking (Vygotsky, 1978). Studies also emphasize that group-based learning fosters motivation and deeper engagement (Jacobs & Kimura, 2013). However, the presence of neutral or negative responses suggests that not all students may feel equally comfortable or benefit from group activities, due to differences in learning styles or participation levels.

A total of 70% of respondents found technology integration beneficial, while 30% expressed neutral or negative views, with a mean score of 1.47. The role of technology in language learning has been widely recognized, with digital tools such as online resources, learning management systems (LMS), and interactive applications enhancing the learning process (Chapelle & Sauro, 2017). However, mixed responses suggest that some students may not have equal access to technological resources, or they may not find them as effective as traditional learning methods. Research suggests that effective technology integration depends on its alignment with pedagogy and its ability to engage learners interactively (Hubbard, 2013).

The effectiveness of assessment methods received a lower approval rating, with 63.33% positive responses and 36.67% neutral or negative responses, resulting in a mean score of 1.87. This finding indicates some concerns regarding how well assessments measure students' language proficiency. Effective assessment practices should align with learning objectives, incorporating both formative and summative evaluations (Brown, 2004). The presence of a high percentage of neutral or negative responses suggests that some students may perceive existing assessment methods as inadequate in accurately reflecting their proficiency levels. Research indicates that traditional tests may not fully capture language competency, and alternative approaches, such as performance-based assessments and portfolio evaluations, could provide a more comprehensive measure of student learning (Bachman & Palmer, 2010).

The highest approval rating was observed in overall satisfaction with English courses, with 83.33% of respondents expressing positive views, while 16.67% were neutral or negative, resulting in a mean score of 1.23. This suggests that most students find the courses beneficial, aligning with research that well-structured language courses with engaging content led to greater student satisfaction (Richards, 2015). However, the existence of neutral or negative responses indicates that some students may feel that specific aspects, such as assessment, instructional strategies, or course materials, require improvement. Previous studies suggest that continuous curriculum evaluation and

student feedback mechanisms are essential for maintaining course relevance and effectiveness (Graves, 2008).

– **Qualitative Analysis of Teachers' Responses to Speaking Skills Assessment**

Teachers perceived that current speaking assessments are inauthentic and fail to reflect realistic communication. Limited resources, such as technology and materials, restrict the ability to develop functional tasks (Luoma, 2004). The institutional emphasis on linguistic efficiency over communicative ability complicates comprehensive evaluation.

Teachers highlighted the persistence of subjectivity despite structured rubrics, particularly in areas like accent and fluency. Output is often prioritized over interpersonal skills, and rigid rubrics fail to account for the variability in everyday communication (Fulcher, 2010).

Cultural norms affect student performance, with shyness and hesitation impacting confidence in formal assessments (Hofstede, 2001). Teachers requested more space for cultural sensitivity in the assessment process to alleviate anxiety and foster equity (Kumaravadivelu, 2008).

Teachers noted a lack of practical communication skills in current practices. While students are often proficient in linguistic precision, they frequently struggle with tasks requiring adaptation and interaction (Canale & Swain, 1980).

Recommendations included implementing dynamic tasks such as debates and role-plays, and incorporating cross-cultural contexts to prepare students for real-world situations. These changes emphasize the importance of practical language production for accuracy and relevance in real-world language use (Jones, 2007; Gass & Mackey, 2007a; Gass & Mackey, 2007b).

– **Qualitative Analysis of Students' Responses to Speaking Assessments**

Students highlighted anxiety and nervousness as key challenges, often influenced by cultural norms that discourage public speaking. These issues are particularly acute for rural or conservative students unfamiliar with formal assessment (McNamara, 2009; Shohamy, 2001).

Urban students reported higher confidence due to more opportunities to practice English, while those from under-resourced areas struggled, reflecting inequities in exposure and resources (Hughes, 2003).

Students complained that current assessments favor linguistic “correctness” over communicative competence, privileging students with higher English-speaking “acculturation”. Assessments often neglect conversational fluency, focusing instead on academic English (Wiggins, 1990; Nunan, 2004).

Students expressed concerns over the lack of practical tasks in assessments, which focus on structured, academic language rather than real-world skills. Research supports the use of activities like role-playing and dialogue to simulate real-life speech.

Students recommended active, participatory, and interactive strategies, such as group discussions and role-playing, aligning with Wiggins' (1990) concept of authentic assessment. They emphasized the need for communicative validity over grammatical correctness, with clear rubrics and feedback (Black & Wiliam, 1998).

– Document Analysis

The document analysis provides insights into the evolution of assessment practices over three academic years, highlighting shifts from a theoretical, accuracy-focused approach to a more communicative, practical, and inclusive assessment framework. The findings suggest gradual but meaningful progress in assessment design, feedback mechanisms, and the integration of cultural and technological elements.

In the first year, assessments primarily emphasized linguistic accuracy, with a strong focus on grammar and pronunciation rather than natural language use. This approach aligns with traditional assessment models that prioritize structural correctness over communicative effectiveness (Brown, 2004).

However, such methods have been criticized for their limited applicability to real-world communication, as they often fail to measure functional language use (Bachman & Palmer, 2010). Additionally, cultural sensitivity and assessment diversity were minimal, and feedback was limited, which may have restricted students' ability to reflect on and improve their performance.

In the second year, moderate progress was observed, as practical aspects of language learning were introduced. A notable shift was the prioritization of fluency over accuracy, aligning with research suggesting that communicative competence should take precedence in language acquisition (Richards & Rodgers, 2014). The introduction of culturally sensitive materials marked an effort to diversify content and enhance student engagement, addressing concerns about linguistic relativity and intercultural competence (Byram, 1997).

Additionally, the implementation of rubrics provided clearer evaluation criteria, while group work and presentations encouraged interactive learning. Feedback also became more specific, though opportunities for reassessment remained limited, which may have hindered students' ability to revise and improve their language skills.

By the third year, assessments underwent significant improvements, shifting towards a communicative, real-world-oriented approach. Tasks became more diverse and pragmatic, simulating scenarios such as interviews and debates, which align with best practices in task-based language teaching (Ellis, 2003). The assessment rubrics evolved to emphasize communicative skills, moving beyond grammatical accuracy to include fluency, coherence, and interaction, thus fostering a more holistic evaluation framework (Skehan, 2014).

Moreover, feedback mechanisms improved, incorporating multiple opportunities for reassessment, which supports formative assessment practices that enhance learner autonomy (Black & Wiliam, 1998). Technology also played an increasing role in assessment, with virtual tools and multimedia resources enriching the evaluation process. Research suggests that digital tools can enhance engagement and provide more dynamic assessment methods, allowing students to demonstrate their language skills in varied contexts (Chapelle & Sauro, 2017).

The **discussive** issues may have different results in the country Ethiopia (Habtamu, & Belay, 2023) as well as for teaching English speaking in different countries (Ptushka, 2024; Kostikova, Miasoiedova, Razumenko, Chernenko, & Pochuieva, 2019).

CONCLUSIONS

The findings from the various analyses underscore the importance of continuous improvement in English language education to meet the evolving needs of students. The major challenges of EFL teachers and learners in the assessment of speaking abilities at Wollo University are described. EFL teachers and learners' perceptions about the utility and impartiality of present speaking ability assessment procedures are investigated. The strategies regarding the authenticity and efficacy of speaking skill assessment in EFL education of Wollo University are offered.

While English courses are perceived as relevant and effective, there is a clear need for enhancements in teaching methodologies, course materials, administrative support, and assessment practices. Modernizing instructional approaches to incorporate interactive and communicative strategies, updating materials to ensure contextual relevance, and integrating technology effectively are critical steps toward improving student engagement and learning outcomes.

Additionally, addressing challenges related to assessment diversity, fairness, and alignment with real-world communication demands is essential for fostering students' academic and professional growth. Future research should focus on identifying specific barriers to implementing these improvements and evaluating targeted interventions to optimize language education. By prioritizing these areas, institutions can create more inclusive, effective, and student-centered learning environments that better prepare learners for success in a globalized world.

Moreover, the role of faculty development and institutional support cannot be overstated. Providing educators with ongoing professional development opportunities, access to modern teaching resources, and platforms for collaboration can empower them to implement innovative practices effectively. Institutions should also prioritize creating feedback loops between students, instructors, and administrators to ensure that course designs and assessments remain responsive to learners' needs. By fostering a culture of continuous reflection and adaptation, English language programs can not only address current limitations but also anticipate future challenges, ensuring sustained relevance and impact in an ever-changing educational landscape.

ACKNOWLEDGEMENTS

We would like to extend our deepest gratitude to all those who supported and guided us throughout this study. Our sincere thanks go to Wollo University for providing the necessary resources and flexibility, which enabled us to carry out this research. Their contribution was crucial in achieving the objectives of this research. Lastly, we acknowledge the support of our colleagues and friends for their encouragement and guidance throughout this journey.

CONFLICT OF INTERESTS

The authors declare that there are no conflicts of interest regarding the publication of this paper.

FUNDING

The authors declare that this study received no specific financial support.

REFERENCES

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press. <https://doi.org/10.1093/applin/ams032>
- Basturkmen, H. (2010). *Developing courses in English for specific purposes*. Palgrave Macmillan. <https://doi.org/10.1057/9780230290518>
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education.
- 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. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/l.1.1>
- Chapelle, C. A., & Sauro, S. (2017). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell. <https://doi.org/10.1002/9781118914069>
- Day, C. (2013). The New Lives of Teachers. In M.A. Flores, A.A. Carvalho, F.I. Ferreira, M.T. Vilaça (Eds.), *Back to the Future*. SensePublishers. https://doi.org/10.1007/978-94-6209-240-2_4
- Dolgin, J., Kelly, K., & Zelkha, S. (2010). *Authentic assessments for the English classroom*. National Council of Teachers of English
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Farrell, T. S. C. (2018). *Reflective language teaching: Practical applications for TESOL teachers*. Bloomsbury Publishing.
- Freeman, D. (2017). *Educating second language teachers: Theorizing and practicing teacher education*. Oxford University Press.
- Freeman, L. (2017). The importance of explicitly teaching language and literacy to English language learners. *Practical Literacy: The Early and Primary Years*, 22(2), 37–40.
- Fulcher, G. (2010). *Practical language testing*. Routledge. <https://doi.org/10.4324/980203767399>
- Gass, S. M., & Mackey, A. (2007a). *Data elicitation for second and foreign language research*. Routledge. <https://doi.org/10.4324/9780203826102>
- Gass, S. M., & Mackey, A. (2007b). Input, interaction, and output in second language acquisition. In B. Vanpatten, & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175-200). London: LEA.

- Graves, K. (2008). The language curriculum: A social contextual perspective. *Language Teaching*, 41(2), 147–181. <https://doi.org/10.1017/S0261444807004867>
- Habtamu, N., & Belay, B. (2023). Examining English Language Instructors' Practice of Reflective Teaching Components in Ethiopia. *Educational Challenges*, 28(2), 75–96. <https://doi.org/10.34142/2709-7986.2023.28.2.06>
- Harmer, J. (2007). *The practice of English language teaching* (4th ed.). Pearson Education.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Sage Publications.
- Hubbard, P. (2013). *Computer-assisted language learning: Critical concepts in linguistics*. Routledge.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.
- Jacobs, G. M., & Kimura, H. (2013). *Cooperative Learning and Teaching*. TESOL International Association
- Jones, L. (2007). *The student-centered classroom*. Cambridge University Press.
- Kostikova, I., Miasoiedova, S., Razumenko, T., Chernenko, A., & Pochuieva, O. (2019). Teaching English speaking for FCE: using Facebook as a tool of instructional practice. *Amazonia Investiga*, 8(22), 719–727. <https://amazoniainvestiga.info/index.php/amazonia/article/view/825/773>
- Kumaravadivelu, B. (2008). *Understanding language teaching: From method to postmethod*. Routledge. <https://doi.org/10.4324/9781410615725>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511733017>
- McNamara, T. F. (2009). *Language testing*. Oxford University Press.
- Mishan, F., & Timmis, I. (2015). *Materials development for TESOL*. Edinburgh University Press. <https://doi.org/10.1515/9780748691371>
- Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design*. Routledge.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667336>
- Orr, M. (2002). The FCE speaking test: Using rater reports to help interpret test scores. *System*, 30(2), 143–154. [https://doi.org/10.1016/S0346-251X\(02\)00003-2](https://doi.org/10.1016/S0346-251X(02)00003-2)
- Ptushka, A. (2024). Level Characteristics of Foreign Language communicative competence development of students for technical specialties. *Educational Challenges*, 29(1), 124–134. <https://doi.org/10.34142/2709-7986.2024.29.1.09>
- Richards, J. C. (2015). *Key issues in language teaching*. Cambridge University Press.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.

Sakuda, S. (2023). Creating authentic assessments for the language classroom. *CELE Journal, 31*, 83–91.

Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Pearson Education.

Skehan, P. (2014). *Individual differences in second language learning*. Routledge.

Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching, 45*(2), 143–179. <https://doi.org/10.1017/S0261444811000528>

Veretiuk, T., & Chetveryk, V. (2024). Developing critical thinking in language and literary education. In *Scientific exploration: Bridging theory and practice* (pp. 138–141). Berlin. <https://dspace.hnpu.edu.ua/handle/123456789/16506>

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan.

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation, 2*(2), 1–3. <https://doi.org/10.7275/ffb1-mm19>

ABSTRACT [in Amharic]:

በወሎ ዩኒቨርሲቲ በእውነተኛ ግምገማ(አተንቲክ) የንግግር ኮርሶች ግምገማ ውስጥ የሚያጋጥሙ ተግዳሮቶችን መርምር

ይህ ጥናት በወሎ ዩኒቨርሲቲ፣ ኢትዮጵያ ውስጥ እንግሊዝኛን እንደ የውጭ ቋንቋ (EFL) በሚማሩበት ኮርስ ውስጥ እውነተኛ የንግግር ግምገማን ለማስተዋወቅ የሚኖሩ ችግሮችን ያስሳል። የውሂብ ስብሰባው ከ25 የእንግሊዝኛ ትምህርት (ELT) መምህራን እና ከ30 የእንግሊዝኛ ተማሪዎች ጋር በኮርሱ መጨረሻ ተካሂዷል። የተዋቀሩ ጥያቄዎች፣ ክፍል-የተዋቀሩ ቃላት መጠይቆች እና ሰነዶች ግምገማ ዘዴዎች ጥቅም ላይ ውለዋል። ቁጥራዊ ውሂብ በስታቲስቲካዊ ሁኔታ ተተንትኗል፣ እንዲሁም ጥራዊ ውሂብ በገጽታ ትንታኔ በመጠቀም በግምገማ ልምድ ውስጥ ያሉ እጥረቶች ተለይተዋል። ውጤቶቹ ከተግባር ጋር የሚገናኙ አውዶች አለመኖር፣ የተለያዩ የግምገማ መስፈርቶች፣ በቂ ያልሆኑ መሰሪያዎች፣ የመምህራን ሙያዊ እድገት ገደቦች እና የተማሪዎች ምክንያቶች (ለምሳሌ፣ ፍርሃት፣ ግብረመልስ አለመኖር) የመሳሰሉ ተግዳሮቶችን ያመለክታሉ። በተጨማሪ፣ መምህራን የተለያዩ ተማሪዎች የግምገማ ፍላጎቶቻቸው እንዲሟሉ፣ በግምገማ ሂደቶች እና በእውነተኛ ሕይወት ግንኙነት መካከል ያለው ግንኙነት እንዲጠበቅ ማድረግ አስችጋሪ ሆኖ ተገኝቷል፤ ተማሪዎችም የንግግር እድሎች እና ጠቃሚ ግብረመልስ ለማግኘት ውሳኔን እድሎች አሏቸው። ተጨማሪ ተግዳሮቶች የጊዜ ገደቦች፣ የሙያዊ እድገት እጥረቶች እና በክፍል ውስጥ የመሰሪያዎች እጥረትን ያካትታሉ። ጥራዊ ውጤቶች ባህላዊ ተስማሚ የግምገማ ዘዴዎች ወሳኝ እንደሆኑ ያሳያሉ፤ ማለትም፣ በኢትዮጵያ ያለው የቋንቋ ብዙሀን ግምት ውስጥ ማስገባት አስፈላጊ ነው። በግምገማ ዘዴው እና በተማሪዎች የግንኙነት ኮርስ ሥራ ፍላጎቶች መካከል ያለው አለመስማማት ገደቦችን ያሳያል። የተማሪዎችን የንግግር እንግሊዝኛ ብቃት ለማሳደግ እውነተኛ፣ ፍትሃዊ እና አስተማማኝ የግምገማ መሰሪያዎች በመጠቀም፣ ይህ ጥናት እንደ ተጨማሪ የመምህራን ስልጠና፣ በእውነተኛ ሕይወት ሁኔታዎች ውስጥ የሚደረግ ግምገማ እና የመሰሪያዎች ውጤታማ አጠቃቀም ያሉ አስፈላጊ ምክሮችን ያቀርባል።

ዋና ቃላት:- እውነተኛ የንግግር ግምገማ፣ ባህላዊ ግምገማ፣ የእንግሊዝኛን እንደ የውጭ ቋንቋ የመማር ተግዳሮቶች፣ ሁሉን አቀፍ አቀራረቦች፣ የቋንቋ ልዩነት።