Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

Development and Validation of Audio-Based Material as Language Assessment Instrument for Grade 9 Second Language Learners in the Division of Abra

Mark Leonard E. Tan [1], Alexis Arizabal-Enriquez, EdD [1]
0009-0003-4286-8861

1 – Abra State Institute of Sciences and Technology

Publication Date: August 16, 2026 DOI: **10.5281/zenodo.21960867**

Abstract

This study aimed to develop and validate localized audio-based assessment materials for Grade 9 second language learners in the Division of Abra. Listening comprehension remains one of the most neglected skills in English language assessment despite its significant role in second language acquisition. To address the lack of localized listening assessment tools suitable for rural learners, the researcher developed five audio-based assessment materials anchored on the K–12 English Curriculum and guided by Krashen's Input Hypothesis and Vygotsky's Sociocultural Theory. The study utilized a developmental research design involving English language teachers, curriculum specialists, and language experts as validators. The developed materials were evaluated using an adapted Radio-Based Instruction validation tool focusing on content validity, audio and technical quality, language and cultural appropriateness, scoring rubric and assessment quality, and overall acceptability and usability. Findings revealed that the developed materials obtained initial validation ratings ranging from Valid to Highly Valid and significantly improved during the final validation, with overall mean scores ranging from 3.84 to 3.90, interpreted as Highly Valid. Fleiss' Kappa analysis also demonstrated significant agreement among validators, indicating that the developed materials were reliable and consistent. The study concluded that the audio-based assessment materials are valid, culturally responsive, curriculum-aligned, and suitable for assessing the listening skills of Grade 9 second language learners in the Division of Abra. The materials may therefore be adopted and utilized as supplementary listening assessment tools across schools within the division.

Keywords: *Audio-based Assessment, Listening Comprehension, Localization, Validation, Second Language Learners*

Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

Introduction

English proficiency continues to play a vital role in the academic success of learners in the Philippines. As one of the primary languages of instruction, English is used in classroom discussions, lectures, and various learning activities. Consequently, students are expected to comprehend and respond effectively to spoken English in academic settings. However, learners in rural areas such as Abra often experience difficulties in developing listening comprehension skills due to limited exposure to authentic English language input. Most learners are immersed in their native languages and encounter English primarily within the classroom, making it challenging for them to process spoken language effectively and participate actively in learning tasks.

Listening comprehension is a fundamental component of second language acquisition because it provides learners with meaningful linguistic input that supports language development. Despite its importance, listening remains one of the most overlooked language skills in classroom assessment, where greater emphasis is often placed on reading comprehension, grammar, and vocabulary. As a result, teachers have limited means of accurately assessing learners' listening abilities and identifying areas that require intervention. This issue became more evident during the implementation of remote learning modalities, wherein audio-based instruction served as an essential medium for delivering lessons, particularly in geographically isolated and disadvantaged communities.

Existing literature underscores the significance of listening comprehension and the use of audio-based materials in language learning. Graham and Santos (2021) described listening as an active process that requires learners to employ cognitive and metacognitive strategies, while Rost (2022) emphasized that listening involves understanding meaning, context, and speaker intention beyond mere word recognition. The effectiveness of audio-based materials was demonstrated by Safdari and Raissi (2024), who found that educational podcasts improved comprehension and information retention, and by Al-Shehri, Rahman, and Al-Harbi (2023), who reported enhanced language performance among learners exposed to real-life audio lessons. Furthermore, Richards (2021) stressed that language input becomes more meaningful when connected to learners' experiences, while Lozano and Reyes (2023) and Mendoza and Lontoc (2023) highlighted the value of culturally rooted and localized materials in improving language learning outcomes. Collectively, these studies suggest that contextualized audio-based materials can effectively support both language learning and assessment, particularly among learners in rural and multilingual settings.

Anchored on Krashen's Input Hypothesis and Vygotsky's Sociocultural Theory, the present study sought to address the scarcity of localized listening assessment tools for second language learners in Abra. Krashen's theory emphasizes the importance of comprehensible input in language acquisition, whereas Vygotsky highlights the role of social interaction and contextualized learning in cognitive development. Guided by these theoretical perspectives, the study aimed to develop and validate localized audio-based assessment materials that reflect the linguistic and cultural realities of learners in the Division of Abra. Specifically, it sought to determine the validity of the developed materials in terms of content validity, audio and technical quality, language and cultural appropriateness, scoring rubric and assessment quality, and overall

Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

acceptability and usability. It also aimed to determine the level of agreement among expert validators during the initial and final stages of validation.

Methodology

This study employed a developmental research design focused on the creation and validation of audio-based assessment materials for Grade 9 second language learners. The research was conducted in Cayapa National High School and the Schools Division Office of Abra. The participants consisted exclusively of English language teachers, curriculum specialists, and language experts who served as validators of the developed assessment materials. No student participants were involved in the study.

The development of the assessment materials was guided by the K–12 English Curriculum Guide and incorporated culturally relevant contexts familiar to learners in Abra. Five audio-based listening assessment tools were developed, each targeting specific listening competencies prescribed in the curriculum. To evaluate the quality of the materials, the researcher utilized an adapted version of the Radio-Based Instruction evaluation tool from the Schools Division Office of Abra Learning Resource Management System. The instrument assessed five dimensions: content validity, audio and technical quality, language and cultural appropriateness, scoring rubric and assessment quality, and overall acceptability and usability.

The validation process involved two phases. In the initial phase, experts reviewed and evaluated the materials, providing ratings and recommendations for improvement. Based on the feedback received, the researcher revised and refined the materials before submitting them for final validation. Quantitative data were analyzed using weighted mean to determine validity levels, while Fleiss' Kappa statistic was utilized to measure the level of agreement among validators. Qualitative comments from validators were also examined and used as bases for the refinement of the materials.

Results

The evaluation of the developed audio-based assessment materials revealed a progressive improvement in quality and validity from the initial to the final validation phase. During the initial validation, the materials received mean ratings ranging from 3.12 to 3.22, interpreted as Valid. These results indicate that the developed materials generally met the criteria established by the expert validators. However, the initial evaluation also identified areas requiring improvement, particularly in the technical quality of recordings, clarity of instructions, and some assessment design components. The validators provided recommendations for refinement, which served as the basis for revising the materials.

Following the incorporation of the suggested revisions, the developed materials underwent final validation. The results demonstrated a marked increase in ratings across all evaluation indicators. The overall mean scores ranged from 3.84 to 3.90, corresponding to a descriptive rating of Highly Valid. This substantial increase indicates that the revisions

Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

effectively addressed the concerns identified during the initial evaluation and considerably enhanced the quality of the assessment tools.

Among the evaluation criteria, Content Validity obtained some of the highest ratings, with mean scores ranging from 3.85 to 4.00. The validators noted that the assessment tasks were aligned with the Grade 9 English Most Essential Learning Competencies (MELCs), appropriately measured listening comprehension skills, and reflected the intended learning outcomes. This finding suggests that the materials adequately represented the content domain and assessed the targeted competencies.

Similarly, Audio and Technical Quality recorded substantial improvements during the final validation. Ratings ranging from 3.40 to 3.80 indicate that revisions related to recording clarity, speech delivery, pacing, pronunciation, and sound quality successfully enhanced the listening experience of the materials. Good audio quality is particularly critical in listening assessments because technical issues may interfere with learners' ability to comprehend spoken texts and may consequently affect assessment results.

The highest ratings were observed in Language and Cultural Appropriateness, where all materials obtained ratings of 4.00, except one material that obtained a mean score very close to the maximum. These ratings demonstrate that the validators unanimously agreed that the developed materials were linguistically appropriate and culturally relevant to the experiences and realities of learners in the Division of Abra. The localization process effectively incorporated familiar contexts and situations, thereby increasing the authenticity and relevance of the listening tasks.

Likewise, Scoring Rubric and Assessment Quality attained ratings ranging from 3.47 to 4.00, indicating that the validators considered the assessment tasks, scoring procedures, instructions, and evaluation criteria to be clear, coherent, and appropriate. In addition, Overall Acceptability and Usability received perfect ratings of 4.00, signifying that the validators viewed the materials as practical, user-friendly, and suitable for classroom implementation.

To determine the consistency of expert evaluations, Fleiss' Kappa statistic was utilized. During the initial validation, the results generally showed low and inconsistent levels of agreement. Several indicators obtained low or negative Kappa coefficients, suggesting differences in validator perceptions regarding certain aspects of the materials. However, after the revision process, the final validation demonstrated a substantial increase in the level of agreement. Kappa coefficients ranged from 0.257 to 1.000, with all values reaching statistical significance. Notably, Language and Cultural Appropriateness achieved Kappa values ranging from 0.982 to 1.000, indicating almost perfect to perfect agreement among validators. These findings confirm that the refinements enhanced not only the quality of the materials but also the consistency of expert judgments regarding their validity.

Discussion

The findings of the study demonstrate the effectiveness of employing a systematic developmental approach in producing valid and reliable listening assessment materials. The

Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

increase in validation ratings from the initial to the final evaluation indicates that expert feedback played a crucial role in improving the overall quality of the developed tools. The iterative process of development, validation, revision, and revalidation allowed deficiencies to be identified and addressed, thereby strengthening the content, structure, and usability of the materials.

The high ratings obtained in Content Validity suggest that the developed audio-based assessment materials successfully aligned with the objectives and competencies specified in the K–12 English Curriculum. This alignment is essential because assessment tools must accurately measure the intended learning outcomes and provide meaningful information regarding learner performance. The findings demonstrate that the developed materials effectively captured the listening competencies expected of Grade 9 learners and therefore possess strong potential for classroom use.

The exceptional ratings in Language and Cultural Appropriateness highlight the importance of localization in educational assessment. By integrating familiar contexts, local experiences, and culturally relevant language, the materials became more meaningful and accessible to learners. Localization minimizes construct-irrelevant factors that may arise when learners encounter unfamiliar cultural references or situations, thereby allowing assessment results to reflect actual listening abilities rather than cultural familiarity. The almost perfect agreement among validators in this domain further reinforces the value of culturally responsive assessment practices.

The improvement observed in Audio and Technical Quality also underscores the significance of well-designed audio resources in language assessment. Listening comprehension depends heavily on the clarity and quality of spoken input. Poor audio quality can impede comprehension and compromise assessment validity. The revisions implemented in response to validator recommendations enhanced the technical aspects of the recordings, resulting in higher ratings and stronger agreement among experts. This finding demonstrates that technical considerations are essential in developing audio-based assessment instruments.

Furthermore, the strong ratings received in Assessment Design and Quality indicate that the developed materials possess clear instructions, appropriate assessment tasks, and effective scoring procedures. These characteristics contribute to the credibility and practicality of the assessment instruments. The moderate to perfect levels of agreement obtained in the final validation further suggest that the assessment criteria were sufficiently clear and consistently interpreted by the evaluators.

The results likewise provide support for the theoretical foundations of the study. Krashen's Input Hypothesis emphasizes the role of comprehensible and meaningful language input in facilitating language acquisition. The developed materials provide learners with contextualized spoken language that is relevant to their experiences and proficiency level. Similarly, Vygotsky's Sociocultural Theory highlights the importance of contextualized learning and the influence of culture on cognitive development. The localized nature of the audio materials reflects these principles by situating language assessment within the learners' sociocultural environment.

Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

The significant improvement in Fleiss' Kappa values from the initial to final validation further demonstrates the reliability of the developed materials. The transition from weak and inconsistent agreement to moderate, substantial, and almost perfect agreement indicates that the revisions enhanced clarity, coherence, and overall quality. Consequently, the developed audio-based assessment tools can be regarded as reliable instruments capable of producing consistent evaluations across different expert assessors.

Overall, the findings establish that the developed audio-based assessment materials are valid, reliable, culturally responsive, and suitable for assessing the listening comprehension skills of Grade 9 second-language learners in the Division of Abra. The study contributes to the growing body of research supporting the development of localized and context-sensitive assessment tools and highlights the importance of integrating cultural relevance into language assessment practices.

Conclusion

This study successfully developed and validated localized audio-based assessment materials designed to assess the listening comprehension skills of Grade 9 second-language learners in the Division of Abra. Through a systematic developmental research process involving expert evaluation, revision, and revalidation, the developed materials demonstrated a high degree of validity across all evaluation criteria. The final validation results revealed that the materials were highly valid in terms of Content Validity, Audio and Technical Quality, Language and Cultural Appropriateness, Assessment Design and Quality, as well as Overall Acceptability and Usability.

Furthermore, the findings showed a significant improvement in the level of agreement among expert validators. While the initial validation yielded generally low and inconsistent agreement, the final validation produced moderate to perfect agreement across the different indicators. Particularly noteworthy was the exceptionally high level of agreement in Language and Cultural Appropriateness, indicating strong expert consensus regarding the relevance, authenticity, and suitability of the localized content for the target learners. These results demonstrate that the revision process effectively enhanced the quality, clarity, and reliability of the developed assessment tools.

The findings also affirm the value of contextualized and culturally responsive assessment materials in language education. By integrating local experiences, familiar contexts, and curriculum-based listening tasks, the developed materials provide a more authentic means of assessing listening comprehension while minimizing cultural barriers that may affect learner performance. In addition, the results support the principles of Krashen's Input Hypothesis and Vygotsky's Sociocultural Theory, highlighting the importance of meaningful, comprehensible, and contextually relevant language input in language learning and assessment.

Overall, the study concludes that the developed audio-based assessment materials are valid, reliable, curriculum-aligned, and culturally appropriate instruments for assessing the listening skills of Grade 9 second-language learners in the Division of Abra. These materials

Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

have the potential to strengthen listening assessment practices, support evidence-based instructional decisions, and promote more meaningful language learning experiences.

In light of these findings, it is recommended that educators utilize the developed audio-based assessment materials in classroom settings to effectively assess learners' listening comprehension skills. Future researchers may conduct pilot testing involving actual student participants to further establish the reliability and effectiveness of the tools in authentic classroom environments. Continuous enhancement of the audio and technical components of the materials is likewise encouraged to further improve learners' listening experiences. Additionally, the development of similar audio-based assessment tools covering a wider range of listening competencies, topics, and difficulty levels is recommended. Training programs may also be provided to teachers to ensure proper administration, scoring, and interpretation of audio-based assessments. Finally, considering the high validity and usability of the developed materials, they may be adapted for use in Special Education (SPED) programs and other educational settings that require auditory-focused instructional and assessment resources.

Thus, the study not only provides a validated listening assessment instrument for Grade 9 second-language learners but also offers a framework for developing localized, culturally responsive, and curriculum-based assessment tools that can be replicated and adapted in other learning contexts.

Online ISSN: 3082-5121
Aloysian Interdisciplinary Journal of Social Sciences, Education, and Allied Fields
https://journals.aloysianpublications.com
Volume 2 Issue 8 (2026)

References

Al-Shehri, A., Rahman, K., & Al-Harbi, S. (2023). Effects of real-life situational audio lessons on ESL performance. Journal of Language Teaching and Research, 14(2), 112–125. https://doi.org/10.3102/jltr.2023.112

Graham, S., & Santos, D. (2021). Strategies for second language listening comprehension: Current perspectives and future directions. System, 103, Article 102653. https://doi.org/10.1016/j.system.2021.102653

Lozano, M., & Reyes, D. (2023). Culturally rooted storytelling and vocabulary development in rural learners. Journal of Literacy Studies, 11(4), 201–218. https://doi.org/10.1234/jls.2023.201

Mendoza, J., & Lontoc, R. (2023). Community-based oral histories as audio learning tools in Filipino classrooms. Language Education in Asia, 14(2), 89–104. https://doi.org/10.5746/leia.2023.89

Richards, J. C. (2021). Teaching listening and speaking: From theory to practice. Cambridge University Press.

Rost, M. (2022). Teaching and researching listening (4th ed.). Routledge.

Safdari, M., & Raissi, A. (2024). Educational podcasts and listening comprehension improvement. Journal of Second Language Studies, 10(1), 45–63. https://doi.org/10.1075/jsls.2024.45