

The Effectiveness of the English Listening Step by Step Application in Learning Listening Skills at the Eighth-Grade Students MTs N 1 Padangsidimpuan

Della Mandalika Hasibuan ¹, Khairunnisah ², Shoufi Nisma Dewi ³

^{1,2,3} Universitas Muhammadiyah Tapanuli Selatan, Padangsidimpuan

Article Info

Article history:

Received 14 Juni 2026

Revised 15 Juni 2026

Accepted 24 September 2026

Keywords:

Effectiveness;

Mobile Application;

Listening Skills;

Mobile-Assisted Language

Learning;

Quasi-Experiment;

Abstract

This study investigates the effectiveness of the English Listening Step by Step application in improving listening skills of eighth-grade students at MTs N 1 Padangsidimpuan in the 2025/2026 academic year. A quasi-experimental design with a Non-equivalent Control Group was employed, involving 64 students divided into an experimental class (8.1) and a control class (8.2). Data were collected through pre-tests, post-tests, and Likert-scale questionnaires. Since data were not normally distributed, the Wilcoxon Signed Rank Test and Mann-Whitney U Test were applied. Results showed the experimental group's mean score increased by 22.81 points, compared to only 2.81 points in the control group. The Mann-Whitney U Test yielded Asymp. Sig. (2-tailed) = 0.026 < 0.05, confirming a significant difference between groups. The experimental class N-Gain score reached 0.60 (medium category), and questionnaire responses were overwhelmingly positive, reflecting high student motivation and satisfaction.

Corresponding Author: Della Mandalika Hasibuan, dedeshasibuan90@gmail.com

1. Introduction

Listening skills form the primary foundation in English as a Foreign Language (EFL) learning. According to Richards (2020), listening is an active process involving the segmentation of sound streams into meaningful units, the identification of lexical items, and the simultaneous interpretation of grammatical structures. Vandergrift and Goh (2018) add that listening is a complex cognitive and metacognitive activity, in which listeners construct meaning through the interaction between oral input and their pre-existing schematic knowledge.

Nevertheless, listening skills often constitute the main barrier for EFL learners. Field (2019) notes that connected speech—with its phenomena of elision, assimilation, and catenation—poses difficulties for students accustomed to hearing words pronounced in isolation. Vandergrift and Goh (2018) further point out that speech rates exceeding students' cognitive processing capacity frequently cause cognitive overload. Moreover, Wilson (2021) asserts that listening anxiety is a significant affective barrier, particularly when instructional media are monotonous and non-interactive.

Pre-research findings at MTs N 1 Padangsidimpuan revealed that 87.5% of eighth-grade students failed to meet the Minimum Competency Criteria (KKM). Students encountered difficulties comprehending oral information due to the limited variety of instructional media used by teachers. This condition underscores the need to integrate digital technology—particularly mobile-based applications—as a more interactive instructional medium suited to the Digital Learning era of 2025/2026.

The concept of Mobile-Assisted Language Learning (MALL) has proven effective in supporting language learning. Atikah Sari (2021) found that podcast applications enhanced students' learning independence due to their portability and ease of use. Rossa (2023) demonstrated that the Learn English By Listening application had a significant effect on listening comprehension scores at SMA Muhammadiyah 2 Bandar Lampung. Lestari and Putri (2025) also showed that digital media such as ELLLO significantly

optimized learning outcomes ($t_{\text{obtained}} > t_{\text{critical}}$). However, no study has specifically examined the effectiveness of the English Listening Step by Step application in the MTsN context in North Sumatra.

The English Listening Step by Step application (developed by Xuan Trung Truong) offers a multi-level system (Beginner to C2), accent variations (American, British, International), synchronized transcripts, interactive quizzes with immediate feedback, and a replay feature that allows students to control their own learning pace. These features directly address the cognitive and affective challenges faced by EFL students.

This study aims to: (1) measure the effectiveness of the English Listening Step by Step application in improving the listening skills of eighth-grade students at MTs N 1 Padangsidempuan, (2) test the significance of outcome differences between the experimental and control classes, and (3) describe students' responses to the use of the application as a recommendation for future listening instructional media.

2. Method

This study employed a quantitative approach with a Non-equivalent Control Group quasi-experimental design. This design was chosen because class selection could not be performed through pure random assignment; instead, it was based on naturally formed classes in the school. Two classes were used: the experimental class (VIII-1), which received the treatment of using the English Listening Step by Step application, and the control class (VIII-2), which continued using conventional methods.

The research population comprised all eighth-grade students at MTs N 1 Padangsidempuan in the 2025/2026 academic year, consisting of four classes with a total of 128 students. A sample of 64 students was determined through purposive sampling, comprising 32 students from class VIII-1 as the experimental class and 32 students from class VIII-2 as the control class.

The research instruments consisted of three components. First, listening tests (pre-test and post-test) comprising 10 audio-based multiple-choice items with a maximum score of 100, covering the aspects of Main Idea (3 items), Specific Information (5 items), and Vocabulary in Context (2 items). Second, a Likert-scale questionnaire with 10 statements (scores 1–5) measuring ease of use, motivation, feature utility, and application recommendation. Third, a semi-structured interview guide directed at selected teachers and students to enrich the qualitative data.

The data collection procedure was carried out in three stages. The first stage was the pre-test to measure both groups' initial ability. The second stage was the implementation of the treatment over six meetings (1 pre-test + 4 treatment sessions + 1 post-test), in which the experimental class used a Pre-Listening, While-Listening, and Post-Listening pedagogical cycle assisted by the application, while the control class used lecture methods and conventional audio media. The third stage consisted of the post-test and questionnaire distribution.

The data analysis technique employed non-parametric tests, as the results of the Shapiro-Wilk normality test indicated that the data were not normally distributed ($\text{Sig.} < 0.05$) even after logarithmic transformation. The Wilcoxon Signed Rank Test was used to analyze the pre-test and post-test differences within each class, while the Mann-Whitney U Test was used to compare outcomes between the experimental and control classes. Treatment effectiveness was measured using the N-Gain formula (Hake, 1998) with the following interpretation: $\text{N-Gain} \geq 0.70$ (high), $0.30 \leq \text{N-Gain} < 0.70$ (medium), and $\text{N-Gain} < 0.30$ (low).

3. Results

Research Data Description

Based on the pre-test and post-test data processing, the experimental class showed a highly significant improvement. The experimental class's mean pre-test score was 62.50, which increased to 85.31 on the post-test, a gain of 22.81 points. Conversely, the control class experienced only a marginal improvement from 71.25 to 74.06, an increase of approximately 2.81 points. This striking difference indicates a substantial treatment effect on the experimental class.

Table 1. Comparison of Pre-test and Post-test Results

Group	Mean Pre-test	Mean Post-test	Improvement	N-Gain
Experimental	62,50	85,31	22,81	0.60 (Medium)
Control	71,25	74,06	2,81	0.06 (Low)

Source: Primary Research Data (2026)

Normality Test

Before determining the appropriate statistical test, a normality test was conducted using the Shapiro-Wilk method because the sample size for each class was less than 50 ($n = 32$). The test results indicated that only the control class pre-test data were normally distributed, with a significance value of $0.092 > 0.05$. The other three data groups were not normally distributed ($\text{Sig.} < 0.05$). Attempts to transform the data using $\log_{10}$ and $\ln$ also failed to produce a normal distribution. This condition is attributed to the complexity of the cognitive processing involved in listening skills, which generates high individual score variation (Vandergrift & Goh, 2018). Therefore, non-parametric tests were employed for all subsequent inferential analyses.

Wilcoxon Signed Rank Test Results

The Wilcoxon Signed Rank Test was used to examine the significance of differences between pre-test and post-test scores within each class. The test results for the experimental class yielded Asymp. Sig. (2-tailed) = $0.000 < 0.05$. Accordingly, H_0 was rejected and H_a was accepted: a significant difference exists between the experimental class's pre-test and post-test scores. This improvement can be explained through the bottom-up and top-down processing mechanisms facilitated by the application's features. The audio feature helps students recognize language sounds, while the transcript feature and material variety support contextual comprehension (top-down processing). In line with the MALL concept (Atikah Sari, 2021; Rossa, 2023), the self-directed learning flexibility offered by the application enables students to control their learning pace according to their individual abilities.

Mann-Whitney U Test Results

The Mann-Whitney U Test was used to compare post-test scores between the experimental and control classes. The test yielded Asymp. Sig. (2-tailed) = $0.026 < 0.05$ with a Z-value of -2.223. Since the significance value is less than 0.05, H_0 was rejected and H_a was accepted: a significant difference exists between the learning outcomes of the experimental and control classes. The negative Z-value indicates that the experimental class had a higher rank than the control class. This difference arose because the experimental class received a more interactive learning experience through audio repetition, transcript access, and practice exercises with immediate feedback, as opposed to the one-directional conventional instruction.

N-Gain Score Analysis

Treatment effectiveness was measured using the N-Gain formula (Hake, 1998). In this analysis, three students who obtained perfect scores (100) on the pre-test were excluded from the calculation as they did not show representative improvement, so the analysis was conducted on 29 students in each class. The experimental class obtained a mean N-Gain of 0.60 ($SD = 0.34$), classified as medium category. The minimum N-Gain value was 0.00 and the maximum was 1.00. Meanwhile, the control class obtained a mean N-Gain of 0.06 ($SD = 0.89$), classified as low category, with a minimum value of -3.00 indicating a decline in learning outcomes for some students. The significant difference between the two groups (0.60 vs. 0.06) confirms that the use of the English Listening Step by Step application produced a greater positive impact than conventional methods.

Table 2. N-Gain Score Analysis Results

Parameter	N	Mean	Std. Deviation	Category
N-Gain Experimental	29	0,60	0,34	Medium
N-Gain Control	29	0,06	0,89	Low

Source: Primary Research Data (2026)

Student Responses to the Application

The Likert-scale questionnaire distributed to 32 students in the experimental class revealed overwhelmingly positive responses. The majority of students' total scores fell in the range of 35–45 (Very Positive category), with the exception of one respondent (S14), who gave the minimum score for all

statements. Interestingly, S14's pre-test and post-test data showed a dramatic improvement from a score of 20 to 80, indicating that the application's systematic structure still successfully delivered the necessary linguistic input even when the student's affective engagement was low.

Table 3. Distribution of Student Questionnaire Responses (Scale 1–5)

Statement	STS	TS	N	S	SS
Ease of use of the application	3%	0%	0%	38%	59%
Increased motivation to learn listening	3%	0%	3%	72%	22%
Transcript feature aids comprehension	3%	0%	0%	34%	63%
Quiz scores motivate repeated attempts	3%	0%	13%	47%	38%
Overall learning experience satisfaction	3%	0%	9%	53%	34%
Increased confidence in listening	3%	3%	16%	44%	34%
Suitable for independent study at home	3%	0%	3%	53%	41%
Varied material, less likely to get bored	3%	0%	9%	44%	44%
Willing to recommend to friends	3%	0%	13%	56%	28%
Suitable as a long-term learning medium	3%	0%	0%	47%	50%

Note: STS=Strongly Disagree, TS=Disagree, N=Neutral, S=Agree, SS=Strongly Agree

4. Discussion

The findings of this study consistently support several existing theories and previous research. The significant score improvement in the experimental class aligns with Mayer's (2001) Cognitive Theory of Multimedia Learning, which posits that students learn more deeply when words are combined with images or interactivity rather than words alone. The application's ability to integrate audio, synchronized visual transcripts, and instant quiz feedback enables better information encoding in students' cognitive systems.

The experimental class's success is also supported by the balance between bottom-up and top-down processing. Clear audio recordings help students decode sounds, phonemes, and word boundaries (bottom-up), while the transcript feature and material variety allow students to draw on prior knowledge to predict and understand context (top-down). As stated by Vandergrift and Goh (2018), effective listening instruction must guide students through the integrated processes of perception, parsing, and utilization.

From a psychological perspective, the overwhelmingly positive questionnaire results reflect an increase in students' sense of autonomy and competence. This is consistent with Self-Determination Theory (Deci & Ryan, 2017), in which students who feel competent and autonomous tend to exhibit higher intrinsic motivation. The immediate score feature in the quizzes provides instant validation that encourages students to repeat exercises until achieving a perfect score.

These findings strengthen and extend previous research (Atikah Sari, 2021; Rossa, 2023; Lestari & Putri, 2025) that has demonstrated the effectiveness of digital applications in listening instruction. The distinctiveness of this study lies in its use of the English Listening Step by Step application with a structured multi-level system in the MTsN context in North Sumatra, confirming the generalizability of MALL in Islamic-based secondary school settings in Indonesia.

5. Conclusion

Based on the data analysis and discussion conducted, this study produced three main conclusions. First, there is a significant effect of using the English Listening Step by Step application on the listening skills of eighth-grade students at MTs N 1 Padangsidempuan, as evidenced by the Wilcoxon Signed Rank Test result with Asymp. Sig. (2-tailed) = 0.000 < 0.05. Second, the application proved effective at a medium level of effectiveness (N-Gain = 0.60), far surpassing the control class (N-Gain = 0.06). The experimental class's mean score increased by 22.81 points (62.50 → 85.31), compared to the control class's improvement of only

2.81 points. Third, students responded very positively to the use of the application, with the majority of questionnaire scores falling in the range of 35–45 (Very Positive), reflecting high levels of motivation, confidence, and learning satisfaction.

Based on the conclusions above, recommendations are directed to four parties. First, English teachers are advised to integrate the English Listening Step by Step application as a supplementary medium in listening instruction to overcome the limitations of conventional methods by providing diverse, interactive, and repeated input. Second, students are encouraged to use the application for independent practice outside of class hours, as listening skills require consistent exposure. Third, the school administration is expected to support digital transformation by providing facilities and policies that enable student access to digital learning platforms. Fourth, future researchers are advised to expand the study with larger samples, longer treatment durations, and to explore the application's impact on other language skills such as speaking and pronunciation.

References

- Atikah Sari, D. (2021). The effectiveness of using podcasts to improve students' listening comprehension [Undergraduate Thesis]. State Islamic University.
- Deci, E. L., & Ryan, R. M. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press.
- Fayzullayeva, N. (2023). The importance of listening skills in the educational process. *International Journal of Pedagogics*, 3(05), 45–51.
- Field, J. (2019). *Listening in the language classroom*. Cambridge University Press.
- Floyd, J. J. (2011). *Listening: A life skill* (2nd ed.). Allyn & Bacon.
- Goh, C. C. (2022). *Teaching and learning second language listening: Metacognition in action*. Routledge.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74.
- Hargie, O. (2022). *Skilled interpersonal communication: Research, theory and practice*. Routledge.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Lestari, A., & Putri, D. A. (2025). The effectiveness of ELLLO application toward students' listening comprehension at MA Mathalibul Huda Mlonggo. *Journal of English Language Teaching*, 10(1).
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.
- Richards, J. C. (2020). *The practice of English language teaching*. Pearson Education.
- Rossa, A. (2023). Improving students' listening skills through "Learn English by Listening" application [Unpublished Thesis].
- Rost, M. (2016). *Teaching and researching: Listening* (3rd ed.). Routledge.
- Sugiyono. (2019). *Quantitative, qualitative, and R&D research methods*. Alfabeta.
- Vandergrift, L., & Goh, C. C. (2018). *Teaching and learning second language listening: Metacognition in action*. Routledge.
- Wilson, J. (2021). *How to teach listening*. Pearson