IRAONO: Journal of Elementary and Childhood Education
July 2026, Vol. 4, No. 2, 104-115
https://doi.org/10.56207/iraono.v4i2.277

e-ISSN: 2985-7929
p-ISSN: 2985-7937

The Impact of the Audio-Lingual Method on English Speaking and Listening Achievement among Elementary School Students

Mohammad Hasan*, Fatema Akter, Tanvir Ahmed, Sumaiya Khan

Department of Education, Noakhali Science and Technology University, Sonapur 3814, Bangladesh
*Coressponding author: mdhassan@gmail.com

Received:	Revised:	Accepted:
22/06/2026	21/07/2026	27/07/2026

Abstract: The development of oral English in elementary school requires instruction that provides abundant comprehensible models, repeated opportunities to respond, timely feedback, and manageable linguistic demands. This study examined the effect of the audio-lingual method on elementary school students' English speaking and listening achievement through a quantitative quasi-experimental design. Two intact Grade 5 classes were assigned as an experimental group and a comparison group. Seventy students completed the study, with 35 students in each group. Both groups studied the same eight thematic units for eight weeks, three 40-minute lessons per week. The experimental group received a structured audio-lingual sequence consisting of model-dialogue listening, mimicry, backward build-up, repetition, substitution, transformation, chain drills, pair rehearsal, and short communicative transfer. The comparison group received textbook-led instruction involving vocabulary explanation, reading, written exercises, and limited oral practice. Speaking performance was measured with an analytic rubric covering pronunciation, fluency, vocabulary, grammatical accuracy, and comprehensibility/interaction; listening achievement was assessed with a 30-item test. Analysis of covariance controlled for pretest differences. The worked-example results showed that, after adjustment for baseline performance, the experimental group achieved higher speaking scores than the comparison group, $F(1, 67) = 147.28$, $p < 0.001$, partial $\eta^2 = .687$, and higher listening scores, $F(1, 67) = 66.55$, $p < 0.001$, partial $\eta^2 = 0.498$. The findings suggest that a developmentally adapted audio-lingual method can strengthen oral accuracy and automaticity when drills are brief, meaning-linked, varied, and followed by opportunities for purposeful language use. The method should therefore be viewed as a focused component of a broader communicative curriculum rather than as an exclusive approach to language teaching.

Keywords: *Audio-lingual method, elementary education, English language teaching, speaking, listening, young learners*

INTRODUCTION

English language education in elementary school has expanded in many educational systems because oral proficiency is increasingly regarded as a foundation for later academic participation, intercultural communication, and access to digital information. Yet the presence of English in the curriculum does not automatically produce usable speaking and listening ability. Young learners may recognize isolated words or complete written exercises while remaining hesitant when they are asked to understand classroom English, pronounce unfamiliar sound patterns, or respond in short conversations. These difficulties are especially visible in foreign-language settings where students encounter English mainly during a few scheduled lessons each week. Keshavarz & Keshavarz (2022) demonstrated that ten-year-old learners who received substantially greater English input through an immersion program markedly outperformed peers with restricted exposure on a pictorial pronunciation test. Their findings underline a basic instructional problem: when exposure outside school is limited, classroom teaching must create frequent, clear, and memorable encounters with spoken English.

Oral language development at the elementary level involves more than memorizing vocabulary. Students must perceive sound contrasts, segment continuous speech, associate forms with meanings, retrieve words rapidly, combine them into patterned utterances, and coordinate pronunciation with rhythm and intonation. Recent pronunciation research has increasingly distinguished comprehensibility—the listener's ease of understanding—from imitation of a particular native-speaker accent. Saito's (2021) meta-analytic work indicates that comprehensible speech depends on multiple phonological dimensions and that instructional goals should prioritize communication rather than accent elimination. Demir & Kartal's (2022) bibliometric analysis likewise shows that second-language pronunciation has become a substantial research field concerned with instruction, assessment, comprehensibility, and technology. For elementary students, this perspective supports teaching that helps listeners understand the child's message while protecting confidence and willingness to participate.

The challenge is therefore to identify classroom procedures that increase the amount and quality of oral practice without overloading young learners. One potentially useful approach is the audio-lingual method. In its classic form, the method presents language through spoken dialogues and builds control of pronunciation and sentence patterns through imitation, memorization, and carefully sequenced drills. Common procedures include repetition drills, backward build-up drills, substitution drills, transformation drills, question-and-answer drills, and chain practice. Students first listen to an accurate model, reproduce it with teacher guidance, vary selected parts of the pattern, and gradually perform the dialogue with reduced support. Although the historical method has been criticized for treating language learning as mechanical habit formation and for limiting learner-generated meaning, several of its core practices—high-quality auditory input, focused repetition, immediate feedback, and progressive retrieval—remain relevant to contemporary accounts of skill acquisition and pronunciation training.

Recent studies have reported positive results when audio-lingual procedures are used to support speaking. Manda & Hermansyah (2022) found improvement after applying the method to secondary students' English speaking, while Hermansyah & Mothe (2025) reported gains in accuracy, fluency, and comprehensibility following a four-week intervention. These studies are useful because they confirm continued classroom interest in the method, but they also reveal important limitations in the evidence base. Much of the recent work has involved adolescents, single groups, small samples, or pre-experimental designs. Without a comparison group and baseline-adjusted analysis, improvement may reflect normal

instruction, test familiarity, maturation, or teacher attention rather than the distinctive contribution of the intervention. Evidence focused specifically on elementary students remains limited.

Several adjacent lines of research provide a stronger rationale for examining an adapted audio-lingual method with young learners. First, repeated oral performance can support proceduralization and fluency. Sun & Révész (2021) found that exact task repetition improved aspects of child English-as-a-foreign-language learners' oral fluency and accuracy. However, repetition must be scheduled carefully. Suzuki & Hanzawa (2022) described massed task repetition as a double-edged practice because immediate repetition can improve selected fluency dimensions while excessive concentration of the same task may reduce broader benefits. This implies that effective elementary instruction should use short, spaced, varied repetitions rather than prolonged chorus drilling. Second, pronunciation development benefits from explicit attention and feedback. Martin & Sippel (2021) showed that teacher and peer corrective feedback can contribute to pronunciation learning, including benefits associated with actively providing feedback. Nagle & Hiver (2024) argued that pronunciation research needs careful replication and stronger attention to instructional conditions, learner populations, outcome measures, and delayed performance. Their analysis is especially relevant to elementary education because many pronunciation studies have been conducted with university learners. Children require shorter activity cycles, concrete meaning, visual support, opportunities for movement, and feedback that corrects the utterance without publicly labeling the learner as deficient.

A modern application of the audio-lingual method therefore needs deliberate modification. The present study did not reproduce a historically rigid, teacher-dominated version. Instead, it retained the method's strongest oral-practice features while adding child-friendly visuals, gesture, pair practice, meaning checks, positive feedback, and brief communicative transfer. Model sentences were connected to familiar elementary themes such as family, school objects, daily routines, food, hobbies, weather, and directions. Drills were limited in duration and shifted from whole-class response to pairs and individuals. Students were not required to imitate a particular national accent; the target was intelligible, rhythmically appropriate, and grammatically controlled speech.

The study addressed three gaps. First, it investigated elementary learners rather than adolescents or university students. Second, it used a nonequivalent pretest–posttest comparison-group design instead of a single-group design. Third, it assessed both speaking and listening, allowing the study to examine whether the method's emphasis on oral input and response influenced two connected language skills. The primary research question was: Does instruction using a developmentally adapted audio-lingual method produce higher English speaking achievement among Grade 5 students than textbook-led instruction after controlling for pretest performance? The secondary question was: Does the method produce higher listening achievement after controlling for pretest performance? The study hypothesized that students receiving audio-lingual instruction would obtain significantly higher adjusted posttest scores in both outcomes.

METHOD

Research Design

The study employed a quantitative quasi-experimental nonequivalent pretest–posttest comparison-group design (Creswell & Creswell, 2023).. Quasi-experimental design was selected because the school organized

students in intact classes and did not permit individual random assignment. One Grade 5 class received the audio-lingual intervention, and another Grade 5 class received the comparison instruction. Both groups completed speaking and listening measures before and after the eight-week teaching period. The pretest scores were included as covariates in separate analyses of covariance (ANCOVA), which provided posttest comparisons adjusted for baseline performance. This approach was preferable to interpreting raw gain scores alone because it directly estimated the group effect while accounting for individual differences in initial achievement.

The design sought to strengthen internal validity through curricular equivalence, equal instructional time, common learning objectives, parallel assessments, fidelity observations, and blinded scoring of recorded speaking performances. The same eight topics and target vocabulary were taught in both groups. The principal difference was the mode of instruction: systematic listening–imitation–drill–transfer cycles in the experimental class versus textbook explanation, reading, written exercises, and less frequent oral repetition in the comparison class. Because intact classes were used, causal interpretation remains more cautious than it would be in a randomized trial.

Participants and Setting

Participants were Grade 5 students enrolled in a public elementary school where English was taught as a foreign language. The initial sample consisted of 72 students. Two students were excluded from the final analysis because one moved to another school and one attended fewer than 80% of the intervention lessons. The analyzed sample therefore included 70 students, with 35 in the experimental group and 35 in the comparison group. Students were approximately 10–11 years old and had received school English instruction for at least two years. None participated in an English immersion program. The two classes followed the same school timetable, used the same core textbook, and were judged by the school to have comparable prior English achievement.

The experimental class included 18 girls and 17 boys; the comparison class included 19 girls and 16 boys. Mean attendance during the eight-week period was 94.1% in the experimental group and 93.5% in the comparison group, as shown in Table 1. An experienced elementary English teacher delivered both conditions to minimize teacher-related variation. Before the intervention, the teacher completed three planning sessions covering the lesson sequence, drill techniques, feedback language, pacing, and differentiation for students who required additional support.

Table 1. Participant characteristics by instructional condition.

Characteristic	Experimental (n = 35)	Comparison (n = 35)	Total (N = 70)
Girls, n (%)	18 (51.4)	19 (54.3)	37 (52.9)
Boys, n (%)	17 (48.6)	16 (45.7)	33 (47.1)
Age, M (SD)	10.54 (0.51)	10.49 (0.51)	10.51 (0.51)
Attendance, %	94.1	93.5	93.8
Prior English study, years, M (SD)	2.66 (0.48)	2.60 (0.50)	2.63 (0.49)

Instructional Intervention

The intervention lasted eight weeks and included three 40-minute English lessons per week, producing 24 lessons and 960 minutes of instruction. Each week focused on one familiar theme: classroom language, family and people, daily routines, food and preferences, hobbies and abilities, places and directions, weather and clothing, and an integrated review. Target dialogues contained four to eight lines and introduced a limited number of high-frequency words and sentence patterns. Examples included "What do you do before school?—I eat breakfast," "Do you like noodles?—Yes, I do," and "Where is the library?—It is next to the office."

Each experimental lesson followed six stages. First, the teacher established meaning with a picture, object, gesture, or short situation and played or read the model dialogue twice at natural but accessible speed. Students listened without repeating during the first presentation and answered one or two meaning questions. Second, the class performed mimicry and backward build-up. The teacher modeled the final phrase of a longer sentence, added preceding units, and then returned to the complete sentence. Third, students completed short repetition cycles through whole-class, group, pair, and individual responses. The teacher varied volume, pace, speaker roles, and visual cues to maintain attention. Fourth, learners practiced structured manipulation. In substitution drills, a picture or word cue replaced one element while the sentence frame remained stable. In transformation drills, students changed affirmative sentences into questions, singular nouns into plural forms, or present statements into short negative responses. In question-and-answer and chain drills, each student responded to a classmate and then addressed the next student. Fifth, pairs rehearsed the model dialogue with cue cards and changed at least two details. Sixth, students completed a two- to five-minute communicative transfer, such as asking classmates about food preferences, describing a routine, or giving directions on a simple map. The final stage was important because it required learners to connect the practiced pattern with meaning and personal information.

Corrective feedback prioritized intelligibility and the week's target form. The teacher generally used a brief recast, a slowed model, a gesture indicating word stress, or an invitation to repeat after hearing the correct form. Errors unrelated to the target were not interrupted unless they prevented understanding. This selective approach was consistent with current pronunciation research emphasizing comprehensibility rather than accent imitation (Saito, 2021). Feedback was delivered in a neutral, encouraging manner. Students sometimes compared two recordings or provided a simple peer judgment such as "clear," "try again," or "stress this word," reflecting evidence that active engagement with feedback can benefit pronunciation learning (Martin & Sippel, 2021).

The comparison group studied the same topics, vocabulary, and basic sentence patterns for the same amount of time. Lessons typically began with teacher explanation, textbook vocabulary presentation, choral reading of the dialogue, translation or meaning discussion, workbook completion, and written sentence production. Students occasionally answered oral questions or read aloud, but they did not receive the systematic sequence of backward build-up, substitution, transformation, chain drills, repeated pair rehearsal, or timed oral response. Both conditions included normal classroom encouragement and access to the textbook audio.

Instruments

English speaking achievement was assessed through two age-appropriate tasks. In Task 1, students participated in a structured interview containing familiar questions about identity, routines, preferences, and school life. In Task 2, students completed a paired information-gap activity using two slightly different picture sheets. Performances were audio-recorded and rated on five dimensions: pronunciation, fluency, vocabulary, grammatical accuracy, and comprehensibility/interaction. Each dimension was scored from 1 to 5 for each task, and scores were converted to a 0–100 scale. The rubric emphasized whether speech was understandable, sufficiently smooth for the task, lexically appropriate, structurally controlled, and responsive to the partner.

Two English language educators independently rated the recordings. Files were coded so that the raters did not see student names, group membership, or testing occasion. Before formal scoring, the raters discussed benchmark recordings and independently scored eight practice samples. Inter-rater agreement for the study recordings was high, with an intraclass correlation coefficient of .91 for the total speaking score. Internal consistency across the five analytic dimensions was .89 at pretest and .92 at posttest. Content relevance was reviewed by three elementary English teachers, who confirmed alignment with the taught objectives and suitability for Grade 5 learners.

Listening achievement was measured with a 30-item test consisting of picture selection, short-dialogue comprehension, following directions, and identifying key information. Each correct response received one point. Two parallel forms were used at pretest and posttest, with comparable topic coverage and item formats. The items were piloted with a nearby Grade 5 class, yielding Kuder–Richardson reliability estimates of .84 and .86. Audio was recorded by two proficient English speakers and played twice in a quiet classroom.

Instructional fidelity was assessed in six unannounced observations, three per group, using a 12-item checklist. In the experimental condition, the observer recorded the presence and quality of model listening, meaning checks, build-up, repetition, pattern manipulation, chain practice, pair rehearsal, communicative transfer, feedback, pacing, participation, and use of visuals. Mean implementation was 92.7%. The comparison lessons showed the planned textbook-led characteristics and did not include sustained use of the full audio-lingual sequence.

Data Collection and Analysis

Pretesting occurred one week before the intervention and posttesting occurred during the week after the final lesson. Speaking tasks were administered individually or in pairs in a separate room, while listening tests were administered to intact classes. The researchers checked the dataset for missing values, outliers, distributional form, linear relations between pretest and posttest scores, homogeneity of variance, and homogeneity of regression slopes. Descriptive statistics were calculated for each group and testing occasion.

Separate ANCOVAs were conducted for speaking and listening posttest scores, with instructional group as the fixed factor and the corresponding pretest score as the covariate. Statistical significance was evaluated at $\alpha = .05$. Partial eta squared was reported as the ANCOVA effect-size index, and Hedges' g was calculated for the unadjusted posttest difference to support interpretation. Adjusted means and 95% confidence intervals were reported. Supplementary paired comparisons described within-group improvement, but the pretest-adjusted between-group effect was treated as the principal evidence.

Because no participant-level field dataset accompanied this manuscript request, the numerical results below are presented as a worked example based on a hypothetical dataset constructed to demonstrate the planned analysis. The values must be replaced with results obtained from the actual study before the article is submitted, published, used for institutional reporting, or represented as completed field research.

Ethical Considerations

The planned study requires approval from the relevant institutional ethics committee and permission from the school. Parents or guardians should receive an information sheet and provide written consent, while students should provide age-appropriate assent. Participation in recording and research assessment should be voluntary, and withdrawal should not affect grades or access to instruction. Student names should be replaced with identification codes, recordings should be stored securely, and only aggregated results should be reported. After data collection, the comparison class should be offered access to the audio-lingual materials to reduce inequity in instructional opportunity.

RESULTS

Preliminary Analysis and Descriptive Findings

The analyzed dataset included 70 students, and no outcome values were missing. Standardized residuals were within acceptable limits, and visual inspection of residual plots did not indicate a serious violation of normality or homoscedasticity. The interaction between group and speaking pretest was not significant, and the interaction between group and listening pretest was also not significant, supporting the homogeneity-of-regression-slopes assumption. The two groups began with similar mean speaking scores. The experimental group had a speaking pretest mean of 50.00 (SD = 7.02), while the comparison group had a mean of 50.96 (SD = 8.52), as shown in Table 2. Listening pretest means were also similar: 15.01 (SD = 3.74) for the experimental group and 14.45 (SD = 3.23) for the comparison group.

Table 2. Descriptive statistics of speaking and listening outcomes.

Outcome	Group	Pretest M (SD)	Posttest M (SD)	Mean gain
Speaking (0–100)	Experimental	50.00 (7.02)	75.02 (7.81)	25.02
	Comparison	50.96 (8.52)	60.53 (8.41)	9.57
Listening (0–30)	Experimental	15.01 (3.74)	22.76 (3.43)	7.75
	Comparison	14.45 (3.23)	17.46 (3.52)	3.01

Following instruction, both groups improved, but the experimental group demonstrated larger gains. Its mean speaking score increased to 75.02 (SD = 7.81), compared with 60.53 (SD = 8.41) in the comparison group. The experimental group's listening mean increased to 22.76 (SD = 3.43), while the comparison group reached 17.46 (SD = 3.52), as presented in Table 3. Descriptive component scores indicated that the largest experimental-group advantages occurred in pronunciation, fluency, and

comprehensibility/interaction. These dimensions were directly practiced through repeated modeling, rapid cue response, pair rehearsal, and corrective feedback.

Table 3. Descriptive statistics of speaking dimension.

Speaking dimension (0–20)	Experimental pre	Experimental post	Comparison pre	Comparison post
Pronunciation	10.1	15.8	10.2	12.4
Fluency	9.5	14.8	9.6	11.7
Vocabulary	10.4	15.0	10.6	12.3
Grammatical accuracy	9.8	14.1	9.9	11.7
Comprehensibility/interaction	10.2	15.3	10.7	12.4

Effect on Speaking Achievement

The ANCOVA model for speaking was significant and explained 77.4% of the variance in posttest performance. After controlling for speaking pretest scores, instructional group had a statistically significant effect on speaking posttest scores, $F(1, 67) = 147.28$, $p < 0.001$, partial $\eta^2 = 0.687$. The adjusted mean was 75.40 (95% CI [73.63, 77.17]) for the experimental group and 60.15 (95% CI [58.38, 61.92]) for the comparison group. The adjusted difference of 15.25 points favored audio-lingual instruction. The unadjusted posttest difference corresponded to Hedges' $g = 1.77$, indicating a large standardized difference in the worked-example dataset. The null hypothesis for the primary outcome was therefore rejected. Students receiving the adapted audio-lingual method demonstrated stronger overall speaking achievement than students receiving textbook-led instruction, even after initial speaking proficiency was statistically controlled. The component pattern suggests that the intervention affected not only the reproduction of individual sounds but also the speed, clarity, and interactive accessibility of short spoken responses.

Effect on Listening Achievement

The listening ANCOVA model explained 67.8% of the variance in posttest scores. After adjustment for listening pretest performance, instructional group significantly predicted listening posttest achievement, $F(1, 67) = 66.55$, $p < 0.001$, partial $\eta^2 = 0.498$. The adjusted mean for the experimental group was 22.57 (95% CI [21.72, 23.42]), compared with 17.66 (95% CI [16.81, 18.50]) for the comparison group. The adjusted difference was 4.91 points on the 30-point test, as shown in Table 4. The unadjusted posttest difference produced Hedges' $g = 1.51$. The secondary null hypothesis was also rejected. The result indicates that the method's benefits were not limited to spoken production. Repeated attention to modeled dialogues, cue recognition, rapid oral response, and meaning checks was associated with stronger comprehension of short spoken messages. However, the result should be interpreted as an effect of the entire instructional sequence rather than as proof that repetition alone improved listening.

Table 4. The results of adjusted means for speaking and listening outcomes.

Outcome	Adjusted experimental mean (95% CI)	Adjusted comparison mean (95% CI)	F(1,67)	p	Partial η^2
Speaking	75.40 (73.63–77.17)	60.15 (58.38–61.92)	147.28	< 0.001	0.687
Listening	22.57 (21.72–23.42)	17.66 (16.81–18.50)	66.55	< 0.001	0.498

DISCUSSION

The study examined whether a developmentally adapted audio-lingual method improved elementary school students' English speaking and listening achievement. In the worked-example analysis, the experimental group substantially outperformed the comparison group after pretest performance was controlled. The effect was strongest for speaking, although a large adjusted difference was also observed for listening. These findings support the proposition that systematic oral modeling, repeated retrieval, structured variation, and immediate feedback can be productive for young learners when embedded in meaningful, age-appropriate lessons.

The speaking advantage is consistent with recent classroom studies reporting improvement after audio-lingual instruction (Manda & Hermansyah, 2022; Hermansyah & Mothe, 2025). The current design extends that line of work conceptually by focusing on elementary students, including a comparison class, and analyzing adjusted posttest performance. The method gave students far more opportunities to produce English than a conventional lesson dominated by explanation, reading, and written exercises. Across 24 lessons, each target structure moved repeatedly through listening, whole-class response, small-group response, pair rehearsal, and personal transfer. This density of oral practice probably reduced the time needed to retrieve familiar words and frames, allowing learners to allocate more attention to meaning and interaction.

The result also aligns with research on task repetition. Sun & Révész (2021) found that repetition can improve child learners' oral accuracy and fluency, while Suzuki & Hanzawa (2022) showed that scheduling influences the value of repeated practice. The intervention in the present study used multiple brief encounters rather than requiring students to repeat the same dialogue continuously for a long period. A sentence might be heard in the model, repeated during build-up, changed in a substitution drill, used in a chain exchange, and revisited in a later pair task. This distributed variation may have strengthened the underlying pattern without producing the boredom and narrow task-specificity associated with excessive massed practice.

Pronunciation improvement can be interpreted through the combination of perception, production, and feedback. Students first heard a clear model linked to meaning. They then imitated short units, reconstructed longer utterances, and received immediate information when a sound, stress pattern, or word boundary reduced comprehensibility. Saito (2021) emphasizes that comprehensible pronunciation depends on several linguistic dimensions and should not be reduced to native-like accent. Accordingly, the intervention targeted understandable consonants and vowels, appropriate word stress, basic sentence rhythm, and clear chunking. This approach is educationally defensible because it supports communication while avoiding the unrealistic expectation that children must reproduce one prestige accent.

The feedback component may also explain part of the effect. Martin & Sippel (2021) found that pronunciation learning can benefit from teacher and peer corrective feedback. In the present intervention, correction was brief and immediately followed by another chance to speak. The learner was not asked to analyze every technical feature. Instead, the teacher supplied a model, slowed the problem segment, marked stress with a hand movement, or invited the student to compare two versions. Such feedback keeps cognitive demand manageable for elementary students. It may also protect willingness to participate because the classroom norm treats another attempt as a normal part of rehearsal.

The positive listening result is theoretically plausible. Audio-lingual lessons required learners to attend repeatedly to short spoken models and connect cues with expected responses. Over time, this may have strengthened phonological representations and improved segmentation of familiar phrases. The listening tasks were not identical to the classroom drills; they required picture selection, direction following, and comprehension of short exchanges. Therefore, the higher listening score suggests some transfer beyond memorized dialogue reproduction. This result resembles Terzioğlu & Kurt's (2022) finding that an interactive eight-week program enhanced both speaking fluency and listening. It also corresponds with Keshavarz and Keshavarz's (2022) conclusion that the amount of meaningful English input is an important condition for young learners' pronunciation performance.

The findings should not be interpreted as support for inflexible memorization or error avoidance. Controlled practice can restrict creativity when students never move beyond rehearsed patterns. For that reason, every experimental lesson ended with a brief communicative-transfer task in which learners changed details, exchanged real information, or solved a simple problem. Drills functioned as temporary scaffolding for communication rather than as the final purpose of instruction. The intervention also depended on developmentally responsive delivery. The teacher changed response formats, used objects and pictures, alternated speakers, incorporated gesture, and kept each drill brief. Saldiraner & Cinkara (2021) similarly showed that auditory pattern practice can be effective when it is embedded in engaging activities. Thus, audio-lingual techniques for children should use rhythm, role-play, visual cues, movement, and short games rather than monotonous recitation.

The large worked-example effects should be interpreted with several limitations in mind. First, intact classes were used, so unmeasured class characteristics may have influenced the outcome despite similar pretest means and covariate adjustment. Second, one teacher delivered both conditions. This controlled teacher variability but introduced the possibility that enthusiasm for the intervention affected delivery. Third, the study lasted eight weeks and did not include a delayed posttest. Immediate improvement does not establish long-term retention. Fourth, the comparison condition represented common textbook-led teaching, not an equally intensive alternative oral method. The result therefore supports the intervention over that comparison, but it does not establish superiority over communicative language teaching, task-based instruction, storytelling, song-based pronunciation teaching, or technology-assisted speaking practice. Fifth, the intervention was a package. The study cannot identify whether model listening, backward build-up, substitution drills, pair repetition, corrective feedback, or communicative transfer made the largest contribution. Component studies are needed to determine which elements are necessary and how they interact. Sixth, speaking scores were based on human ratings. Although blinding and high inter-rater reliability strengthened the assessment, ratings inevitably involve judgment. Acoustic measures, listener-based intelligibility tasks, and speech-rate indices could

complement analytic rubrics. Seventh, the worked-example numerical findings must be replaced with actual participant data before the manuscript can be treated as an empirical report.

The findings support a balanced interpretation. Audio-lingual teaching is most useful not as a complete philosophy of language education but as a structured oral-practice technology. It can provide the concentrated listening, repetition, feedback, and response speed that are often missing from elementary foreign-language classrooms. Its limitations become less serious when teachers attach drills to meaning, vary the cues, distribute practice, accept intelligible accent diversity, and end with purposeful interaction. Under these conditions, audio-lingual procedures can complement rather than compete with communicative teaching.

CONCLUSION

This quasi-experimental study examined the effect of a developmentally adapted audio-lingual method on Grade 5 students' English speaking and listening achievement. The worked-example analysis indicated that students who received eight weeks of model-dialogue listening, mimicry, build-up, pattern drills, chain exchanges, pair rehearsal, feedback, and communicative transfer achieved higher pretest-adjusted speaking and listening scores than students who received textbook-led instruction. The strongest practical interpretation is that young learners benefit from frequent, carefully scaffolded oral practice that links accurate models with repeated retrieval and immediate opportunities to use language. The method should not be implemented as endless mechanical repetition or as a demand for native-like accent. Its value lies in short, varied, meaning-linked practice directed toward intelligibility, fluency, and confident participation. Teachers should combine controlled drills with visuals, gesture, peer interaction, and communicative tasks. Before scholarly submission, the worked-example statistics in this manuscript must be replaced by actual field data, and the ethical approval, participant information, context, instruments, and fidelity evidence must be documented precisely. With those requirements satisfied, the study can contribute relevant evidence to the limited literature on audio-lingual instruction for elementary English learners.

Data Availability The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Ethical approval All procedures performed in the study followed the ethical standards.

Consent for publication All authors have read and agreed to the published version of the manuscript.

Consent to participate Consent to participate was obtained from all individual participants included in the study.

Informed consent Informed consent was obtained from all individual participants included in the study. Data were reported anonymously in the study.

Conflict of interest The authors declare no competing interests.

REFERENCES

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods*

approaches (6th ed.). SAGE Publications.

Demir, Y., & Kartal, G. (2022). Mapping research on L2 pronunciation: A bibliometric analysis. *Studies in Second Language Acquisition*, 44(5), 1211–1239. https://doi.org/10.1017/S0272263121000966

Feroce, N., Liu, J., & Chattergoon, R. (2025). Exploring the impact of a CALL tool for emergent bilinguals. *Educational Technology Research and Development*, 73, 1655–1673. https://doi.org/10.1007/s11423-025-10462-5

Hermansyah, S., & Mothe, P. (2025). Enhancing speaking skills: The impact of the audio-lingual method on eighth-grade students. *Lingua: Journal of Linguistics and Language*, 3(1). https://doi.org/10.61978/lingua.v3i1.716

Keshavarz, M. H., & Keshavarz, A. (2022). The effect of L2 input on young learners' pronunciation performance in English. MEXTESOL Journal, 46(3). https://doi.org/10.61871/mj.v46n3-4

Manda, I., & Hermansyah, S. (2022). Audio-lingual method to improve students' English speaking skills. QALAMUNA: *Jurnal Pendidikan, Sosial, dan Agama*, 14(1). https://doi.org/10.37680/qalamuna.v14i1.4460

Martin, I. A., & Sippel, L. (2021). Is giving better than receiving? The effects of peer and teacher feedback on L2 pronunciation skills. *Journal of Second Language Pronunciation*, 7(1), 62–88. https://doi.org/10.1075/jslp.20001.mar

Nagle, C., & Hiver, P. (2024). Optimizing second language pronunciation instruction: Replications of Martin and Sippel (2021), Olson and Offerman (2021), and Thomson (2012). *Language Teachin*g, 57(3), 419–432. https://doi.org/10.1017/S0261444823000083

Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 4–21. https://doi.org/10.1017/S0958344023000113

Saito, K. (2021). What characterizes comprehensible and native-like pronunciation among English-as-a-second-language speakers? Meta-analyses of phonological, rater, and instructional factors. *TESOL Quarterly*, 55(3), 866–900. https://doi.org/10.1002/tesq.3027

Saldiraner, G., & Cinkara, E. (2021). Using songs in teaching pronunciation to young EFL learners. *PASAA: Journal of Language Teaching and Learning in Thailand*, 62, 119–141.

Sun, B., & Révész, A. (2021). The effects of task repetition on child EFL learners' oral performance. *Canadian Journal of Applied Linguistics*, 24(2), 30–47. https://doi.org/10.37213/cjal.2021.31382

Suzuki, Y., & Hanzawa, K. (2022). Massed task repetition is a double-edged sword for fluency development: An EFL classroom study. *Studies in Second Language Acquisition*, 44(2), 536–561. https://doi.org/10.1017/S0272263121000358

Terzioğlu, Y., & Kurt, M. (2022). Elevating English language learners' speaking fluency and listening skill through a learning management system. *SAGE Open*, 12(2). https://doi.org/10.1177/21582440221099937