About this sample

This is a teaching sample of a completed action research. The study is invented: no school, teacher, parent or learner took part, and every number was generated to behave like a real class's. The DepEd rules and the research it cites are real. Delete this page when you adapt the report for your own study.

This sample is dedicated to Antipolo National High School, Antipolo City, where I performed in a speech choir in my fourth year of high school.

Mark Anthony Llego

Speech Choir, Oral Delivery and Speaking Confidence of Grade 10 Learners

An Action Research

[Copy for blind review: the researcher's name, position, and school are removed. The final copy lists them here.]

Conducted from June 17, 2026 to September 4, 2026

Abstract

This action research tested whether a speech choir, built into the poetry lessons of Grade 10 English, would improve how learners read aloud and how calm they feel speaking English in class. One section of 42 learners studied film, prose and drama in the usual way for five weeks and then rehearsed a speech choir arrangement of Edgar Allan Poe's "The Bells" in 14 rehearsals of 20 minutes over the next five weeks, in Term 1 of School Year 2026–2027. Each learner read two unseen passages aloud at three occasions five weeks apart: two pretests before the speech choir and a posttest after it. Two raters scored all 236 recordings in one shuffled pool, without knowing whose they were or when they were made, on the Multidimensional Fluency Scale and a 9-point comprehensibility scale; learners also answered the short-form Foreign Language Classroom Anxiety Scale. The question was whether learners changed more in the speech choir weeks than in the same number of usual weeks before them. For the 38 learners with all three occasions, fluency rose 0.36 points in the usual weeks and 1.51 points in the speech choir weeks, a difference of 1.15 points of 16, 95% CI [0.65, 1.65], $t(37) = 4.65$, $p < .001$, $d_z = 0.75$. Comprehensibility rose 0.55 points more than its drift and anxiety fell 2.05 points more (Holm-adjusted ps = .034 and .034). Expression and phrasing changed most; smoothness and pace did not differ reliably. Interviews showed that the group let anxious learners try a louder voice, while some still felt alone when recorded. The speech choir will be kept, and the next cycle will add a solo, unprepared speaking task.

Keywords: speech choir, choral speaking, oral reading fluency, comprehensibility, speaking anxiety, double pretest, Grade 10 English

Acknowledgment

I thank the school head, who approved the study; the two English teachers who scored every recording, and the Master Teacher in English who served as third rater; the school's ICT coordinator, who coded and shuffled the recordings and kept the key; the parents who gave their consent; and the learners of the section, who rehearsed "The Bells" until the iron bells sounded like iron.

Speech Choir, Oral Delivery and Speaking Confidence of Grade 10 Learners

Context and Rationale

Many of the learners in this Grade 10 section were in my Grade 9 English class in School Year 2025–2026. When they read their pieces aloud in the oral interpretation task of the last quarter, my notes showed the same pattern again and again: most read in an even, flat voice whatever the line said, rushed through the punctuation, and let their voices drop at the ends of lines, and nine asked to read from their seats instead of in front of the class. They could name a poem's sound devices on paper; they did not let anyone hear them.

The Grade 10 English curriculum asks for both. Grade 10 follows the MATATAG curriculum from School Year 2026–2027 (DepEd, 2024a), and its Term 1 is world literature, with "POETRY, PROSE, DRAMA, & FILM" as its text foci, and its items include "rhyme and meter", "tone and mood" and "figures of speech and sound devices" (DepEd, 2026a). The content standard speaks of learners' "communicative competence" in evaluating world literature. The budget of work lists no separate speaking competency for Grade 10, so this study treats reading a text aloud with its meaning as part of responding to literature, which is where Term 1 puts it.

Speaking is also where Filipino learners report the most anxiety. In a study of 242 first-year teacher education students, speaking activities were the highest source of language anxiety and "the only significant predictor of English proficiency" (Jugo, 2020, p. 1). Anxiety and a flat reading voice feed

each other: a learner who fears being heard reads quickly and quietly, and a learner who has never heard herself read well has little reason to feel calmer.

Reading aloud together is one way to break that circle. In whole-class choral reading, a teacher leads the class in reading a text in unison, so every learner practices with the support of the others. Sixth graders who did whole-class choral reading improved more in oral reading fluency than a comparison group (d = 0.64; Paige, 2011), and a meta-analysis of Readers' Theatre, in which learners rehearse and perform a script aloud, found large gains in reading skills (Mastrothanasis et al., 2023). A speech choir goes further than choral reading: the lines are divided among solo voices and groups of light, medium and dark voices, and each line is marked for how loud and how fast it is said, so learners must decide what a line means before they can say it. I found no journal study that measured a speech choir's effect on speaking or on anxiety.

I chose the piece by five criteria: free to copy (in the public domain), rich in sound devices, with clear shifts in tone, long enough for every learner to have lines, and free of anything that mocks a people. Poe's "The Bells" (1849) met all five. Its four movements move from silver sleigh bells to golden wedding bells, brazen alarm bells and iron funeral bells, and its repetitions and sound words ("tinkle, tinkle, tinkle", "clang, and clash, and roar") let learners hear rhyme, meter and onomatopoeia, the Term 1 items, in their own voices.

The study falls under Theme 1 of the Basic Education Research Agenda, Teaching and Learning, sub-theme Instruction (DepEd, 2016). It asked whether rehearsing a speech choir would change how learners read aloud and

how calm they felt, beyond what five weeks of ordinary lessons would have done anyway.

Innovation, Intervention, and Strategy

In this study, a speech choir is a class performing a literary piece together, its lines divided among solo voices and voice groups and marked for loudness, pace and pause. The innovation was a speech choir arrangement of "The Bells", rehearsed in part of the English periods while the class studied the poem.

The Arrangement

I divided the class by how high or low their voices sounded when they read a line together, not by sex, into light, medium and dark voices of about 13 learners each, and gave each solo line to a different learner. Each of the poem's 113 lines has one part (a solo, a voice group, or all) and one loudness mark from very soft to very loud. The parts follow the poem's sound and mood: the light voices take the sleigh bells' "tinkle, tinkle, tinkle", the dark voices keep the beat on "Keeping time, time, time", a single voice sounds the alarm in the third movement before the whole choir answers, and the fourth movement slows and softens to its last line. The arrangement, with a note beside the lines whose sound explains the choice, is Appendix D. A learner who asked not to have a solo was not given one.

The Rehearsals

The speech choir took 20 minutes of three English periods a week, 14 rehearsals from July 24, 2026 to August 24, 2026, 280 minutes in all. The rest of each 45-minute period (one of the English class programs DepEd

allows: DepEd, 2024b, para. 2.a.i) was the term's poetry lessons, which used the arrangement as their text: a movement's sound devices, its tone and mood, and its rhyme and meter were found on the page first and then rehearsed aloud. Each rehearsal followed five steps:

1. Warm-up (3 minutes): breathing, and tongue twisters for clear consonants.
2. Meaning (5 minutes): the day's movement read for sense; learners marked its sound devices and where its mood turns.
3. Sectionals (5 minutes): each voice group read its own lines, then answered the others.
4. Run (5 minutes): the whole choir read the movement through, then again with the teacher's one correction.
5. Exit slip (2 minutes, after rehearsals 3, 7 and 11): "What did you notice about your voice today?"

The class performed the whole arrangement in its English period on August 25, 2026, the close of the poetry lessons. The performance was not graded. No costume was bought (the class wore the school uniform), and nothing was rehearsed outside the English periods. The speech choir here was the poetry lessons' own work, not a co-curricular event; a school's co-curricular activities during the Instructional Block are held "only after class hours" (DepEd, 2026c).

The Usual Weeks

In the five weeks between the two pretests, the class studied film, following the budget of work's own five-week sample, and then prose and

drama, with the oral activities those lessons always include and no choral work. The usual weeks were the comparison: what the learners' reading and anxiety did in five weeks without the speech choir.

Criteria for Success

Before the study, I set the criteria that would decide whether to keep the speech choir and share it with the other Grade 10 English teachers: a mean contrast on the fluency scale of at least 1.00 point of 16 (one level gained on one of the four dimensions, beyond the drift of the usual weeks) with its 95% confidence interval above zero; and comprehensibility and anxiety moving in the expected direction. One point, one level on one of the four dimensions, was the smallest change I judged worth the class time.

Action Research Questions

1. What were the learners' oral reading fluency, comprehensibility, and speaking anxiety at the first pretest, the second pretest, and the posttest?
2. Did the learners' oral reading fluency change more during the five speech choir weeks than during the five usual weeks before them?
3. Did the learners' comprehensibility and speaking anxiety, and each of the four fluency dimensions, change more during the speech choir weeks than during the usual weeks?
4. How did the learners describe their experience of the speech choir, and how do their accounts explain their scores?

Question 2 tested the null hypothesis that the mean contrast on the fluency scale (the change from the second pretest to the posttest minus the

change from the first pretest to the second) is zero. Question 3 tested the same null hypothesis for comprehensibility, anxiety, and each fluency dimension. Questions 1 and 4 are descriptive and carry no hypothesis.

Action Research Methods

Research Design

The study used a one-group pretest–posttest design with a double pretest. Measuring a class only before and after an intervention cannot tell its effect from what the weeks would have done anyway: learners grow older, get used to being tested, and settle into a new class, and "improvements can result from the test itself" (Marsden & Torgerson, 2012, p. 585). A second pretest measures that drift before the intervention begins: "The plausibility of maturation and regression threats is reduced by adding a second pretest prior to the first pretest", and "the two pretests function as a 'dry run' to clarify the biases that might exist in estimating the effects of treatment" (Shadish et al., 2002, p. 110). The three occasions were equally spaced, 5 weeks apart: the first pretest on June 17 and 18, 2026, the second pretest on July 22 and 23, 2026, and the posttest on August 26 and 27, 2026, on class days kept free of the term's tests and holidays (DepEd, 2026c).

The effect of the speech choir was estimated for each learner as the contrast $O3 - 2{\cdot}O2 + O1$: the change during the speech choir weeks ($O3 - O2$) minus the change during the usual weeks ($O2 - O1$). A contrast of zero means the speech choir weeks added nothing to the drift. The design does not rule out an event that fell only in the speech choir weeks, and it assumes that the

drift would have continued at the same pace; both are discussed with the results.

Participants and Other Sources of Data and Information

The participants were the 42 learners of one Grade 10 section that I taught in a public secondary school, Term 1 of School Year 2026–2027. All took part in the lessons and the speech choir, which were the section's English lessons. Data were gathered only from learners whose parents consented and who themselves assented: 40 of the 42. Of these, 38 had all three occasions and were analyzed; the parent of one learner declined, one learner declined to take part, one learner joined the class on July 7, 2026, after the first pretest, and one was absent on both posttest days. The learners were 15 or 16 years old.

Six learners were interviewed, chosen from the scores by a rule set before the study: the two with the largest contrasts on the fluency scale (a tie broken in favor of the learner with the lower second pretest), the two with the lowest, and, among the rest, the two most anxious at the first pretest. The rule picks learners whose accounts can explain both ends of the results and the learners the speech choir was most meant to help. The analyzed learners wrote the exit slips at the rehearsals they attended.

Sample Size

Before the study, I planned for 38 learners with all three occasions, the class less about one in ten expected to miss one. With 38 learners, a one-sample *t* test of the contrast at $\alpha = .05$ (two-sided) has a power of .85 to detect

a contrast of 0.5 standard deviations (d_z), a medium effect; 34 is the fewest that would have reached .80.

Data Gathering Methods

Oral Reading Recordings

At each occasion every learner read two unseen passages aloud. The six passages were Aesop fables in two American versions for young readers, both in the public domain, chosen to be easy to read (Table 1) so that unfamiliar words would not hide how a learner used her voice. Moser et al. (2014) stress "finding texts at students' independent reading levels" and recommend at least two passages per occasion for reliable ratings. The six formed three sets of two, matched on length, sentences and readability, and each learner read the sets in one of three orders drawn by lot (A-B-C, B-C-A, C-A-B), a third of the class in each, so that each set appeared at each occasion about equally often.

Table 1

The Six Passages, by Set

Set	Passage	Source	Words	Sentences	Flesch Reading Ease	Flesch–Kincaid grade
A	The Crab and His Mother	Stickney (1915), p. 8	119	7	92.2	4.6
A	The Crow and the Pitcher	The Aesop for Children (1919), p. 34	114	7	87.9	5.0
B	The Woman and Her Hen	Stickney (1915), p. 36	110	7	91.7	4.4
B	The Leap at Rhodes	The Aesop for Children (1919), pp. 37–38	115	8	88.5	4.5
C	The Boy and the Nettle	Stickney (1915), pp. 169–170	118	7	88.6	5.1

Set	Passage	Source	Words	Sentences	Flesch Reading Ease	Flesch–Kincaid grade
C	The Gnat and the Bull	The Aesop for Children (1919), p. 25	112	8	88.4	4.4

Note. Readability from the standard Flesch formulas, with syllables from the CMU Pronouncing Dictionary. Stickney (1915) and *The Aesop for Children* (1919) are both in the public domain.

Each learner read in a quiet corner of the library, one at a time, with a school-owned tablet on a stand at the same distance and the same settings at every occasion; no personal phone was used. The learner read each passage silently for one minute and then aloud once. Before scoring, the school's ICT coordinator, who did not rate, renamed every one of the 236 recordings with a random five-character code and put them in one folder in random order, keeping the key. The raters therefore never knew whose recording they heard, or whether it came from a pretest or the posttest; blind scoring is "optimal for any study" (Gersten et al., 2005, p. 160).

The raters were two English teachers of the school who did not teach the section. In a one-hour calibration, they scored ten practice recordings, read by volunteers of another section with their parents' consent and deleted afterward, and discussed every score that differed by more than one level. They then scored all 236 recordings in one session on September 2, 2026, each alone. A third rater, a Master Teacher in English of the school, scored a whole recording again when the two raters were two or more levels apart on any fluency dimension or three or more points apart on comprehensibility; the final score on each dimension was then the middle of the three ratings, and otherwise the mean of the two. A learner's score at an occasion was the mean of her two passages.

The recordings were scored on two scales. The Multidimensional Fluency Scale rates four dimensions of reading aloud, expression and volume, phrasing, smoothness, and pace, each from 1 to 4, for a total of 4 to 16 (Rasinski, 2004, p. 19, adapted from Zutell & Rasinski, 1991). With two raters and several passages it has reached generalizability coefficients of .92 to .98 (Moser et al., 2014, as cited in Morrison & Wilcox, 2020), and .91 to .94 in a second study (Smith & Paige, 2019), both with younger children. Comprehensibility, how easy a reading is to understand, was rated on a 9-point scale from 1 = very hard to understand to 9 = very easy to understand, the direction Isaacs and Thomson (2013) used; Munro and Derwing's (1995) original scale runs the other way (Thomson, 2018). Individual listeners agree on comprehensibility only modestly (Kendall's W = .39 to .53 within groups of 20 raters; Isaacs & Thomson, 2013), so this outcome was expected to be the least reliable.

Speaking Anxiety

Learners answered the short-form Foreign Language Classroom Anxiety Scale (S-FLCAS), eight items such as "I can feel my heart pounding when I'm going to be called on" and "I feel confident when I speak," each from 1 (strongly disagree) to 5 (strongly agree), with the two positively worded items (4 and 5) reversed, for a total of 8 to 40; higher means more anxious (Botes et al., 2022). Its items come from the full scale of Horwitz et al. (1986), and its validation sample gave α = .89 (Botes et al., 2022). Because English is these learners' second language, not a foreign one, "FL class" was worded "English class"

throughout; that is the study's one change to the scale. Learners answered it on the first day of each occasion, before any recording.

Reflection Entries and Interviews

Learners wrote a two-minute exit slip after rehearsals 3, 7 and 11, those present at each. The six learners chosen by the rule were interviewed on September 3 and 4, 2026, after the recordings had been scored, one at a time, for about 15 minutes, in English or Filipino as each chose. The guide asked what the rehearsals were like, what the learner noticed about her voice, how the recording felt, and whether anything from the choir helped when reading alone (Appendix E). The interviews were recorded with the learners' assent, transcribed, and the recordings deleted.

Ethics

The study followed DepEd's research ethics: the confidentiality of participants' data, consent forms, and the Child Protection Policy as the reference for working with learners (DepEd, 2012, 2017). Before the study began, each parent received a letter explaining the study and asking for written consent, and each learner was asked to sign an assent form, as DepEd's quality tools for completed research ask (DepEd, 2022, Item 13); 40 of the 42 gave both. Voice recordings and scores are personal information, and information about a learner's education is sensitive personal information under the Data Privacy Act of 2012 (Republic Act No. 10173, Secs. 3[g], 3[l]), so they were kept on the school's tablet and one password-protected computer, identified by codes, used only for this study, and the recordings were deleted once the final scores had been checked. No learner data went

into an AI tool (DepEd, 2026b). The recordings, questionnaires and the performance were not graded, a rule of this study; formative assessment is not graded in any case (DepEd, 2026d, para. 15). Learners whose data were not gathered took part in every lesson and in the performance like everyone else.

Data Analysis

For Question 1, the means and standard deviations at each occasion were computed for the 38 analyzed learners. For Questions 2 and 3, each learner's contrast O3 − 2·O2 + O1 was tested against zero with a one-sample t test, with Cohen's d_z (the mean contrast divided by the standard deviation of the contrasts) and its 95% confidence interval from the noncentral t distribution. The drift (O2 − O1) and the gain (O3 − O2) were tested the same way and reported for context. The normality of the contrasts was checked with the Shapiro–Wilk test, and the Wilcoxon signed-rank test was run as a check that does not assume it. The fluency total was the primary outcome, tested at $\alpha = .05$; the six secondary outcomes (comprehensibility, anxiety, and the four fluency dimensions) were adjusted together by Holm's (1979) procedure.

The raters' agreement was computed before any result: the intraclass correlation for absolute agreement, ICC(A,1) for one rater and ICC(A,2) for the mean of two, from the two-way mean squares (McGraw & Wong, 1996), on the fluency totals and on comprehensibility, with the raters' exact and within-one agreement on each dimension. The anxiety scale's reliability was Cronbach's alpha at each occasion. The analysis was done in Python; an Excel workbook

repeats the scores, the *t* tests, Holm's correction, the agreement statistics and Cronbach's alpha as formulas (Appendix A).

For Question 4, the exit slips and the interview transcripts were analyzed by thematic analysis in Braun and Clarke's (2006) six phases: becoming familiar with the data, coding, generating themes, reviewing them, defining and naming them, and writing them up. The themes were then set beside each interviewee's scores in a joint display (Guetterman et al., 2015) to see where the accounts and the numbers agreed and where they did not.

Discussion of Results and Reflection

The Raters' Agreement

The two raters agreed well on the fluency totals: ICC(A,1) = .85 for one rater and ICC(A,2) = .92, 95% CI [.83, .95], for the mean of the two, which is the score analyzed. On single dimensions they gave the same level on 59.3% (phrasing) to 61.0% (pace) of recordings and were within one level on 97.9% (pace) to 99.6% (phrasing and smoothness). Rater B scored a little higher than Rater A, by 0.72 points of 16 on average; because both raters scored every recording from all three occasions in one shuffled pool, that leniency is the same at every occasion and cancels out of the contrast. A third rater was needed for 20 recordings (8.5%). Comprehensibility agreed less, as expected: ICC(A,1) = .63 and ICC(A,2) = .77, 95% CI [.70, .83]. The anxiety scale was reliable at each occasion, α = .92, .90 and .92.

Scores at the Three Occasions

Table 2 answers Question 1. On the fluency scale the class moved from 10.00 at the first pretest to 10.36 at the second, after the usual weeks, and

to 11.88 at the posttest, after the speech choir weeks. Comprehensibility moved the same way, and anxiety fell a little in the usual weeks and more in the speech choir weeks.

Table 2

Means (and Standard Deviations) at the Three Occasions

Outcome	First pretest	Second pretest	Posttest
Fluency scale total (4 to 16)	10.00 (2.48)	10.36 (2.42)	11.88 (2.32)
Comprehensibility (1 to 9)	5.18 (1.20)	5.42 (1.30)	6.21 (1.19)
Speaking anxiety (8 to 40)	24.76 (7.10)	24.39 (6.51)	21.97 (7.05)
Expression and volume (1 to 4)	2.22 (0.58)	2.33 (0.67)	2.81 (0.63)
Phrasing (1 to 4)	2.40 (0.62)	2.49 (0.64)	2.98 (0.60)
Smoothness (1 to 4)	2.59 (0.69)	2.72 (0.64)	3.02 (0.63)
Pace (1 to 4)	2.78 (0.73)	2.83 (0.58)	3.07 (0.64)

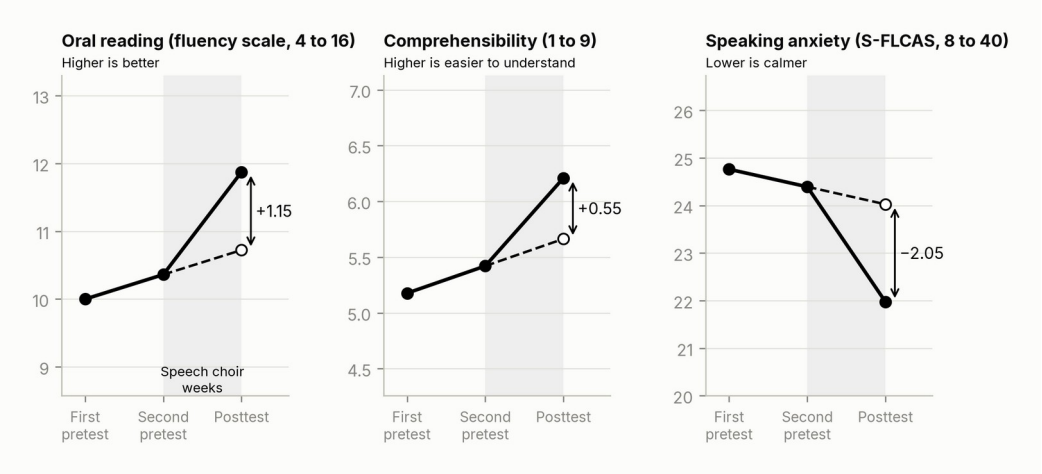
Note. n = 38. Higher is better on every scale except speaking anxiety, where lower is calmer. Fluency and comprehensibility are the means of two raters and two passages.

Did Fluency Change More in the Speech Choir Weeks?

Yes. In the five usual weeks the class's fluency rose 0.36 points, 95% CI [0.06, 0.66], a small but real drift (p = .019). In the five speech choir weeks it rose 1.51 points. The mean contrast was 1.15 points, 95% CI [0.65, 1.65], $t(37) = 4.65$, $p < .001$, $d_z = 0.75$, 95% CI [0.39, 1.11], so the null hypothesis was rejected (Figure 1). The contrasts did not depart from normality (Shapiro–Wilk $W = .95$, $p = .078$), and the Wilcoxon signed-rank test agreed ($V = 524$, $p < .001$). The first criterion for success was met: the mean contrast passed 1.00 point and its confidence interval lay above zero, though the interval's lower end, [0.65, 1.65], reaches below one point.

Figure 1

Class Means at the Three Occasions, With the Drift of the Usual Weeks Extended



The second pretest mattered. Compared with the second pretest alone, the posttest was 1.51 points higher, and compared with the first pretest 1.88 points higher; either figure would have credited the speech choir with weeks of ordinary growth. The double pretest kept 0.36 points of it out of the estimate.

Comprehensibility, Anxiety, and the Four Dimensions

Table 3 answers Question 3. Comprehensibility rose 0.55 points more than its drift, $t(37) = 2.72$, Holm-adjusted $p = .034$, $d_z = 0.44$, and speaking anxiety fell 2.05 points more, $t(37) = -2.78$, Holm-adjusted $p = .034$, $d_z = -0.45$; the second criterion for success was met. Of the four dimensions, expression and volume (0.38 of 4 points, adjusted $p < .001$) and phrasing (0.41, adjusted $p = .004$) changed most, the two a speech choir trains directly. Smoothness (0.18, adjusted $p = .099$) and pace (0.19, adjusted $p = .125$) did not differ reliably after the correction. That does not show they were unaffected; the study cannot tell a small change in them from none.

Table 3

Drift, Gain, and Contrast for Each Outcome

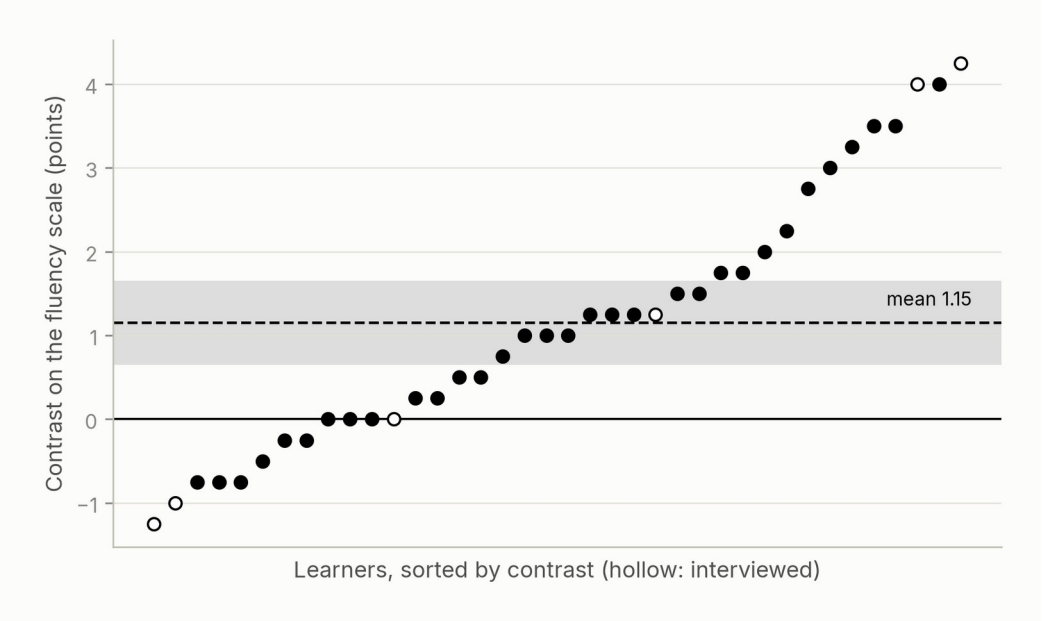
Outcome	Drift (O2 − O1)	Gain (O3 − O2)	Contrast [95% CI]	$t(37)$	p	Holm p	d_z [95% CI]
Fluency scale total (4 to 16)	0.36	1.51	1.15 [0.65, 1.65]	4.65	< .001	(primary)	0.75 [0.39, 1.11]
Comprehensibility (1 to 9)	0.24	0.79	0.55 [0.14, 0.95]	2.72	.010	.034	0.44 [0.11, 0.77]
Speaking anxiety (8 to 40)	−0.37	−2.42	−2.05 [−3.55, −0.56]	−2.78	.009	.034	−0.45 [−0.78, −0.11]
Expression and volume (1 to 4)	0.11	0.48	0.38 [0.20, 0.55]	4.22	< .001	< .001	0.68 [0.33, 1.04]
Phrasing (1 to 4)	0.09	0.49	0.41 [0.18, 0.64]	3.62	< .001	.004	0.59 [0.24, 0.93]
Smoothness (1 to 4)	0.13	0.30	0.18 [0.00, 0.35]	2.03	.049	.099	0.33 [0.00, 0.65]
Pace (1 to 4)	0.05	0.24	0.19 [−0.06, 0.44]	1.57	.125	.125	0.25 [−0.07, 0.58]

Note. n = 38. Contrast = O3 − 2·O2 + O1. Holm's correction covers the six secondary outcomes; the fluency total was the primary outcome. For anxiety, a negative contrast means learners became calmer.

Not every learner gained (Figure 2). Of the 38 contrasts, 26 were positive, four were zero (no more change in the speech choir weeks than in the usual weeks) and eight were negative. Attendance did not explain the difference: learners attended 95.1% of the rehearsals on average, none fewer than 11 of 14, and attendance and contrast were unrelated (Spearman's ρ = .12, p = .481).

Figure 2

Each Learner's Contrast on the Fluency Scale



What the Learners Said

Three themes came from the exit slips and the interviews.

Hiding, then heard. The group made it safe to try a louder voice. The learner with the largest contrast was also one of the most anxious at the start (39 of 40): "At first I only moved my mouth. In the iron bells part I was in the dark voices, and nobody could hear only me, so I tried it louder." Another anxious learner said, "it was everybody, so no one was looking at me," and the light voices' "tinkle, tinkle" part made the class laugh, "and after that it was not so scary."

One line, many ways. Learners found that loudness, pace and pauses are choices that follow meaning. "Ma'am made me say it soft first, then loud, then fast. I never thought a line could be said so many ways," said a learner who had a solo; another now asks, "what is the feeling here." An exit slip after rehearsal 7 read: "When I say 'moaning and groaning' slowly it really sounds

sad." The learner who hid in the dark voices heard the change in his reading of the unseen passage: "I heard myself stopping at the commas like in the poem."

Alone again. For some, the choir's support did not follow them into the recording. "In the choir there is a rhythm you follow, but in the passage there is no rhythm, so I just read," said a learner whose contrast was below zero; she had "copied the girl beside me." Another, who missed rehearsals when she was sick, said, "Mas gusto ko po 'yung may kasabay. [I prefer having others speaking with me.] But the recording is only you." One of the most anxious, who had asked for no solo, said, "Sa grupo, okay lang. Pero pag recording, parang exam. [In the group it is fine. But the recording feels like an exam.]"

Table 4 sets the six interviewees' scores beside what they said.

Table 4

Joint Display: Interviewees' Scores and Accounts

Learner	Chosen as	Fluency contrast	Anxiety, first pretest → posttest	What the learner said	Reading
L07	Largest contrast	4.25	39 → 39	Hid in the dark voices, then tried it louder; pauses at commas now	The group was a safe place to practice; the skill reached the unseen passage
L29	Largest contrast	4.00	14 → 15	A solo line said many ways; reads for the feeling	Calm from the start; gained from deciding how to say a line
L26	Lowest contrast	−1.25	23 → 17	Copied the girl beside her; no rhythm when alone	Followed the group without deciding; the choir's rhythm did not carry over
L09	Lowest contrast	−1.00	25 → 24	Missed rehearsals; one word behind; prefers others speaking with her	Missed practice and leaned on the group
L16	Most anxious at the start	1.25	40 → 37	No one was looking at her; still nervous, but knew what to do with her voice	Skill grew before calm did

Learner	Chosen as	Fluency contrast	Anxiety, first pretest → posttest	What the learner said	Reading
L11	Most anxious at the start	0.00	38 → 33	Fine in the group; asked for no solo; the recording felt like an exam	The group helped; facing a recording alone did not get easier

Note. Anxiety on the S-FLCAS, 8 to 40 (lower is calmer). Learners are identified by codes.

Reflection

The speech choir did what I hoped for most of the class: learners read aloud with more expression and better phrasing after five weeks of it than after five weeks of the usual lessons, and they felt calmer speaking English in class. The interviews explain why it worked for some and not for others. Learners who used the group as a place to try their own voice, or who had to decide how to say a solo line, carried something into reading alone; learners who followed the group without deciding, or missed rehearsals, did not, though across the class attendance and contrast were unrelated, so missed rehearsals explain one learner's account, not the class's results. The next cycle should give every learner at least one line to decide alone, and a reason to say it alone.

The study has limits that a reader should weigh. It is one class, taught by me, the researcher, and the learners knew they were in a study. The double pretest removes the drift of the usual weeks, but not an event that fell only in the speech choir weeks, and it assumes the drift would have continued at the same pace; if the drift was slowing, the contrast understates the effect, and if it was speeding up, it overstates it. Anxiety may fall fastest in the first weeks with a new class, which would make the usual weeks' drift in anxiety larger than later drift; I could not test that with two pretests. The posttest came a

day after the class performance and two days before the Term 1 Examination, which may have raised anxiety at the very occasion meant to show it falling. The rehearsals added 280 minutes of reading aloud that the usual weeks did not have, so the study cannot separate the speech choir from any added practice in reading aloud. The passages were easy, so the scores show how learners used their voices on simple text, not on Grade 10 literature, and the study measured reading aloud, not learners' own unprepared talk. Comprehensibility rested on two raters who agreed only moderately, and the anxiety scale's wording was adapted. Each of these points to the next cycle.

Action Plan

Activity	Persons involved	Timeline	Resources	Expected output
Share the results, the arrangement and the scoring sheet with the Grade 10 English teachers in a learning action cell session	Researcher, English department Head Teacher	Term 2, October 2026	Report, script of "The Bells", scoring sheet	Teachers who decide whether to try it, with the limits known
Second cycle in the same section: every learner given one line to say alone, and a one-minute unprepared talk added to the recordings	Researcher, two raters	Term 2, October to December 2026	School tablet, the same passages, a talk prompt per occasion	Evidence on whether the gains reach learners' own speaking
Compare the speech choir with the same minutes of other reading-aloud practice in two sections	Researcher and one Grade 10 English teacher	Term 3, January to March 2027	The arrangement, a matched reading-aloud plan	An estimate of the speech choir beyond added practice
Train the raters with anchor recordings for comprehensibility	Raters, Master Teacher in English	Before the second cycle	Ten anchor recordings with agreed scores	Higher agreement on comprehensibility
Prepare a second arrangement of another public-domain poem for the Term 2 texts	Researcher	Term 2	The term's texts	A second piece for the next class

References

The Aesop for children (M. Winter, Illus.). (1919). Rand McNally.

Botes, E., van der Westhuizen, L., Dewaele, J.-M., MacIntyre, P., & Greiff, S. (2022). Validating the Short-form Foreign Language Classroom Anxiety Scale. *Applied Linguistics, 43*(5), 1006–1033. https://doi.org/10.1093/applin/amac018

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Department of Education. (2012). *DepEd child protection policy* (DepEd Order No. 40, s. 2012).

Department of Education. (2016). *Adoption of the basic education research agenda* (DepEd Order No. 39, s. 2016).

Department of Education. (2017). *Research management guidelines* (DepEd Order No. 16, s. 2017).

Department of Education. (2022). *Adopting tools to improve quality management of completed research at the Department of Education* (DepEd Memorandum No. 028, s. 2022).

Department of Education. (2024a). *Policy guidelines on the implementation of the MATATAG curriculum* (DepEd Order No. 010, s. 2024).

Department of Education. (2024b). *Amendment to DepEd Order No. 010, s. 2024 (Policy guidelines on the implementation of the MATATAG curriculum)* (DepEd Order No. 012, s. 2024).

Department of Education. (2026a). *Three-term budget of work for learning competencies: English, Grade 10*.

Department of Education. (2026b). *Foundational guidelines on artificial intelligence in basic education* (DepEd Order No. 003, s. 2026).

Department of Education. (2026c). *Guidelines on the implementation of the three-term school calendar in basic education* (DepEd Order No. 009, s. 2026).

Department of Education. (2026d). *Revised guidelines on classroom assessment, grading system, and awards and recognition for the K to 12 basic education program* (DepEd Order No. 015, s. 2026).

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*(2), 149–164. https://doi.org/10.1177/001440290507100202

Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Annals of Family Medicine, 13*(6), 554–561. https://doi.org/10.1370/afm.1865

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly, 10*(2), 135–159. https://doi.org/10.1080/15434303.2013.769545

Jugo, R. R. (2020). Language anxiety in focus: The case of Filipino undergraduate teacher education learners. *Education Research International, 2020*, Article 7049837. https://doi.org/10.1155/2020/7049837

Marsden, E., & Torgerson, C. J. (2012). Single group, pre- and post-test research designs: Some methodological concerns. *Oxford Review of Education, 38*(5), 583–616. https://doi.org/10.1080/03054985.2012.731208

Mastrothanasis, K., Kladaki, M., & Andreou, A. (2023). A systematic review and meta-analysis of the Readers' Theatre impact on the development of reading skills. *International Journal of Educational Research Open, 4*, Article 100243. https://doi.org/10.1016/j.ijedro.2023.100243

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Morrison, T. G., & Wilcox, B. (2020). Assessing expressive oral reading fluency. *Education Sciences, 10*(3), Article 59. https://doi.org/10.3390/educsci10030059

Moser, G. P., Sudweeks, R. R., Morrison, T. G., & Wilcox, B. (2014). Reliability of ratings of children's expressive reading. *Reading*

Psychology, 35(1), 58–79. https://doi.org/10.1080/02702711.2012.675417

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Paige, D. D. (2011). 16 minutes of "eyes-on-text" can make a difference: Whole-class choral reading as an adolescent fluency strategy. *Reading Horizons, 51*(1), 1–20. https://scholarworks.wmich.edu/reading_horizons/vol51/iss1/3

Poe, E. A. (1849, November). The bells. *Sartain's Union Magazine of Literature and Art, 5*(5), 304.

Rasinski, T. V. (2004). *Assessing reading fluency*. Pacific Resources for Education and Learning. https://eric.ed.gov/?id=ED483166

Republic Act No. 10173. (2012). *Data Privacy Act of 2012*.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Smith, G. S., & Paige, D. D. (2019). A study of reliability across multiple raters when using the NAEP and MDFS rubrics to measure oral reading fluency. *Reading Psychology, 40*(1), 34–69. https://doi.org/10.1080/02702711.2018.1555361

Stickney, J. H. (1915). *Æsop's fables: A version for young readers*. Ginn and Company.

Thomson, R. I. (2018). Measurement of accentedness, intelligibility, and comprehensibility. In O. Kang & A. Ginther (Eds.), *Assessment in second language pronunciation* (pp. 11–29). Routledge. https://doi.org/10.4324/9781315170756-2

Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice, 30*(3), 211–217. https://doi.org/10.1080/00405849109543502

Financial Report

| Date | Item | Receipt | Amount (PhP) | |:--|:--|:--|--:| | June 10, 2026 | Printing: parents' consent letters and learners' assent forms (84 copies) | OR 2216 | 252.00 | | June 15, 2026 | Printing: S-FLCAS questionnaires for three occasions (126 copies) | OR 2231 | 189.00 | | July 23, 2026 | Printing: speech choir scripts of "The Bells" (42 copies, 5 pages each) | OR 2398 | 630.00 | | August 28, 2026 | Printing: scoring sheets and the resolution record | OR 2477 | 96.00 | | September 2, 2026 | Snacks for the scoring session (three raters and the coordinator) | OR 0915 | 480.00 | | | **Total** | | **1,647.00** |

NOTE: Fund source: the researcher's own funds. No school funds or grants were used, and no learner paid for anything. The recordings were made on a school-owned tablet already on hand. Receipt numbers are invented for this sample.

Appendices

Appendix A. Data Set and Notebook

The Excel data set (the learners, every rating by both raters and the third, the questionnaire answers, the rehearsal attendance, every learner's scores, drift, gain and contrast, and the *t* tests, Holm's correction, the agreement statistics, Cronbach's alpha and the attendance as formulas) and the Jupyter notebook that recomputes the report's tests from the data.

Appendix B. Parents' Consent Letter and Learners' Assent Form

The blank letter and form go with the supporting documents the evaluators check (DepEd, 2022, Item 13); the signed forms stay with the school and are left out of the archived report.

Appendix C. Scoring Sheet and Resolution Record

The rater's sheet for the four fluency dimensions and comprehensibility, with the third-rater rule (the scale's level descriptions are in Rasinski, 2004, p. 19).

Appendix D. The Speech Choir Arrangement of "The Bells"

Poe's 113 lines with each line's part and loudness mark and the notes beside them.

Appendix E. Interview Guide and Exit Slip

The four interview questions with their follow-ups, and the exit-slip prompt.

Appendix F. Declaration of AI Use

I, the researcher, declare that I used [name of the AI tool] only to check the grammar of the English text of this report, for the purpose of finalizing the

research study. The prompt used was: "Check the grammar of this paragraph. Do not change its meaning or add content." I reviewed every suggestion and take full responsibility for the report. No learner data were entered into the tool. The chat history is at [web address of the chat history]. (Adapted from Annex A of DepEd Order No. 003, s. 2026.)