Published Date: - 25-11-2026

Page No: - 19-22



Teacher-Led, AI-Assisted Instruction and Young Learners' Listening and Speaking in English: A Controlled Experiment in Three Regions of Uzbekistan

Jumazoda Shohidai Jaloliddin
PhD Candidate, Basic Researcher in "Theory and Methodology of Education and Upbringing (Methodology of Teaching English)" of the Department of "Pedagogy and Psychology", Urgench State University, Uzbekistan

ABSTRACT

Background. Young learners pick up a foreign language mainly through hearing and speaking it, yet crowded primary classrooms leave little room for individual feedback.
Aim. To test whether lessons in which artificial intelligence (AI) assists, but never replaces, the teacher can raise Grade 1–4 pupils' English listening and speaking performance.
Methods. A controlled classroom experiment was run with 755 pupils (382 experimental, 373 control) from six state schools in Khorezm, Namangan and Tashkent. Experimental classes worked with a curriculum-based task system, a teacher–pupil–AI work cycle and a speech-recognition platform (TE-FLY). Performance was scored on a 20-indicator, 80-point rubric and sorted into proficiency levels.
Results. The groups began at the same point (19.4% vs 19.3% at the high level). At the end, 79.6% of experimental pupils but only 30.6% of control pupils were at the high level (χ^2 = 186.53, $p < 0.001$; Cramér's $V \approx 0.50$).
Conclusion. Teacher-led AI assistance is a promising way to strengthen early oral English, although the composite score and the lack of delayed testing call for careful interpretation.

KEYWORDS: Artificial intelligence in education; young learners; EFL; listening; speaking; adaptive learning; primary education; Uzbekistan

INTRODUCTION

In the early years of foreign-language study, what children can understand and say aloud largely determines how well they later read, write and interact (Cameron, 2001; Field, 2008). Primary teachers, however, work against real obstacles: pupils' attention is brief, their vocabulary small, their pronunciation unstable, and a single teacher can seldom respond to every child's utterance. Tools such as speech recognition, adaptive task selection and learning analytics promise to ease this burden (Luckin et al., 2016; Holmes, Bialik & Fadel, 2019), provided they stay human-centred and suited to children's age (UNESCO, 2023).

Most empirical work, though, examines one skill, one technique or one application at a time, and a large share of it comes from universities rather than primary schools (Zawacki-Richter et al., 2019). Far less is known about a coherent approach that develops listening and speaking together, in young children, and within an official syllabus and set of textbooks. Uzbekistan, which is currently expanding AI use and renewing general education, offers a timely setting for

NEXT SCIENTISTS CONFERENCES

such a test. We therefore ask: when AI supports a teacher-led programme, do Grade 1–4 pupils make larger gains in listening and speaking than under conventional teaching?

2. Methods

Participants and design. The experiment took place in six state schools: Nos. 21 and 35 in Urgench, Nos. 5 and 10 in Kosonsoy district, and Nos. 229 and 251 in Tashkent's Uchtepa district. Altogether 755 pupils from Grades 1 to 4 participated, 382 in the experimental and 373 in the control condition. Every class followed the national English syllabus and the Guess What! Student's Book series, so the two conditions differed in teaching method rather than content.

Experimental programme. A preliminary review of the syllabus and course books revealed three weaknesses: tasks of uniform difficulty, no adjustment to individual readiness, and no objective check of pronunciation. The programme responded with a set of eight task formats, from listening and repetition to question–answer, spoken reply and open communicative activity. These were sequenced from reproductive through semi-productive to productive work, and each followed a fixed seven-step routine. Lessons were guided by six design principles (child-centredness, age fit, communicative purpose, adaptivity, gradual progression and instant feedback). In practice, the AI platform diagnosed pupils' level, delivered tasks, analysed answers and gave feedback, while the teacher set goals, sustained motivation and made the final instructional decisions. Since recognition accuracy is lower for children's voices, automatic speaking scores were advisory only; the teacher assigned the final mark. Control classes were taught in the usual way without AI tools.

Assessment and analysis. Pupils were scored on an 80-point rubric of 20 indicators grouped into four domains: motivation and attitudes, language and cognitive knowledge, technology-related practice, and creative communicative performance. Scores were then converted into proficiency levels. Because the levels are ordinal, groups were compared with Pearson's χ^2 and the Mann–Whitney U test; effect sizes are reported as Cramér's V and r, and the gap in the high-level share carries a 95% confidence interval.

3. Results

The two conditions were statistically indistinguishable at the outset (19.4% vs 19.3% at the high level; $\chi^2 = 0.02$, $p = 0.99$). By the end, the high level had been reached by 79.6% of experimental pupils and 30.6% of controls (Table 1). This 49.0-point gap (95% CI 42.8–55.2) means that experimental pupils were roughly 2.6 times as likely to reach the top level. Both the χ^2 test and the Mann–Whitney test confirmed the difference ($z = 12.80$, $p < 0.001$), and the effect sizes were large ($V \approx 0.50$; $r \approx 0.47$). Progress occurred in both conditions, but it was much stronger among experimental pupils ($V = 0.61$ for start-to-end change) than among controls ($V = 0.25$).

Table 1. Share of pupils at the high proficiency level (experimental n = 382; control n = 373)

Pupils at the high level	Experimental	Control	Statistic
Start of experiment	19.4%	19.3%	$\chi^2 = 0.02$, $p = 0.99$

NEXT SCIENTISTS CONFERENCES

End of experiment	79.6%	30.6%	$\chi^2 = 186.53$, $p < 0.001$; $z = 12.80$
Gain (percentage points)	+60.2	+11.3	$V \approx 0.50$; $r \approx 0.47$ (end-point groups)

4. Discussion

The pattern suggests that AI is most useful in primary English when it takes over the routine, individual side of the work (checking answers, adjusting difficulty, responding at once) and leaves purpose, motivation and judgement to the teacher. Two mechanisms are plausible: every child received prompt personal feedback, which a whole-class lesson cannot supply, and tasks moved step by step from imitation to real communication in line with pupils' age. Because control pupils also improved, the contribution of the programme is better read from the difference in gains than from the experimental gain alone.

Interpretation should respect several limits. Control classes had no AI tools, so the gap on the technology-related domain cannot be credited to the teaching method alone. The main outcome is a composite score pooled across four grades, which hides skill-specific and grade-specific patterns. Recognition errors with children's speech forced reliance on teacher scoring, and without a delayed post-test we cannot say whether gains last. Follow-up studies should report listening and speaking indicators separately, compare automatic and human ratings, analyse results by grade, and add retention testing.

5. Conclusion

In a three-region sample of 755 primary pupils, teacher-led instruction supported by AI diagnostics, adaptive tasks and immediate feedback coincided with far higher listening and speaking attainment than conventional lessons. The study offers a curriculum-linked task system, a model of teacher–pupil–AI collaboration and a rubric for tracking young learners' progress, all of which can support primary English teaching and teacher education as AI resources reach schools.

References

1. Cameron, L. (2001). Teaching languages to young learners. Cambridge University Press.
2. Field, J. (2008). Listening in the language classroom. Cambridge University Press.
3. Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
4. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson.
5. UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.

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

7. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16, Article 39.