

Technology-Assisted Pronunciation: Using ELSA Speak Application To Master -Ed Endings

Septa Aryanika^{1*}, Nava Anggria Safitri², M. Ridho Kholid³, Ernando Rizki Dalimunte⁴

¹⁻⁴ Universitas Islam Negeri Raden Intan Lampung, Lampung, Indonesia

Abstract

This study aimed to investigate whether the integration of ELSA Speak could improve seventh-grade students' pronunciation of the three regular past-tense *-ed* allomorphs (/t/, /d/, and /ɪd/) at SMPN 1 Candipuro, Indonesia. A Classroom Action Research design was employed with 32 students across two cycles, each comprising planning, action, observation, and reflection. Data were collected through a pre-test, two post-tests, and classroom observations; pronunciation scores were analyzed descriptively, while observational data were used to evaluate learning processes and guide instructional refinement between cycles. The results showed a consistent increase in mean pronunciation scores from 30.00 at baseline to 59.06 after Cycle I and 80.31 after Cycle II. The proportion of students achieving the minimum score of 75 also increased from 0% to 21.88% and finally to 81.25%, exceeding the predetermined classroom success criterion of 80%. Classroom observations further indicated stronger attention, participation, independence, confidence, and more effective use of corrective feedback in Cycle II. These findings suggest that ELSA Speak can support the learning of difficult *-ed* pronunciation patterns when embedded in explicit, structured, and teacher-guided practice, consistent with prior evidence on automatic speech recognition and mobile-assisted pronunciation learning. However, the single-school sample, absence of a control group, short intervention, and lack of delayed or spontaneous-speech assessment limit causal and long-term conclusions. Future studies should employ broader samples, comparison groups, delayed post-tests, and communicative pronunciation tasks to examine retention and transfer.

ARTICLE HISTORY

Received : 2 September 2025

Revised : 18 October 2025

Accepted : 10 December 2025

KEYWORDS

-ed ending; Classroom action research; ELSA Speak; Pronunciation

PUBLISHER'S NOTE

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike (CC BY 4.0) license



CORRESPONDING AUTHOR

Septa Aryanika, Universitas Islam Negeri Raden Intan Lampung, Lampung, Indonesia. Email: septaaryanika@radenintan.ac.id

Introduction

Pronunciation is a fundamental aspect of English learning, yet it often receives less attention compared to grammar and vocabulary in classroom practices (Harmer, 2007; Roach, 1991; Gilakjani, 2017). For non-native learners, including Indonesian students, mastering pronunciation is crucial not only for comprehensibility but also for building communicative competence (Morley, 1991). However, many students still face persistent challenges in producing correct sounds, stresses, and intonation patterns, particularly when dealing with English phonological features that do not exist in their first language (Sholeh, 2015).

Indonesian students often encounter difficulties in learning English pronunciation, primarily due to the influence of their native language. When acquiring a second language, the first language frequently interferes with the target language, transferring phonological habits into the new system (Avery & Ehrlich, 1992; Celce-Murcia et al., 2010). This interference can affect several aspects of pronunciation, including the articulation of individual sounds, stress, intonation, and rhythm. For instance, the absence of certain phonemes in Indonesian, such as the interdental fricative [θ], leads to difficulties in pronouncing words like *think* or *thought* (Hakim, 2012). Similarly, Indonesian

learners struggle with English consonant clusters, particularly in final positions, as Indonesian does not allow such structures (Kurniawan, 2016).

The phonological system of Indonesian differs significantly from that of English, and this often results in systematic errors. For instance, Indonesian lacks certain sounds that are common in English, such as the interdental fricative [θ], which leads learners to substitute it with [t] or [s] when pronouncing words like *think* or *thought*. In addition, consonant clusters pose further difficulties. While English allows both initial clusters (e.g., [fl] in *fly*) and final clusters (e.g., [pt] in *stopped*), Indonesian phonology permits only a limited number of initial clusters (e.g., [pr] in *pramuka*) and does not allow final consonant clusters. Consequently, Indonesian learners often omit, simplify, or insert vowels in cluster environments, which negatively affects their intelligibility.

Given these differences, Indonesian EFL learners require explicit instruction and extensive practice to overcome first language interference. Research has shown that targeted pronunciation training and the use of technological tools, such as speech recognition applications, can help learners improve their accuracy and confidence in English pronunciation (Anggraini, 2022).

One of the most common difficulties among Indonesian learners is pronouncing the -ed ending in regular verbs. Preliminary findings at SMPN 1 Candipuro revealed that 75% of students struggled with this feature, often mispronouncing all -ed endings as /ɪd/. A diagnostic test also showed an average accuracy of only 56.25%, with frequent errors in differentiating between the /t/, /d/, and /ɪd/ sounds. This highlights a specific problem that requires targeted pedagogical intervention. The findings from the question 'Do you have difficulty pronouncing English words ending in -ed?' (n=32) are visualized below:

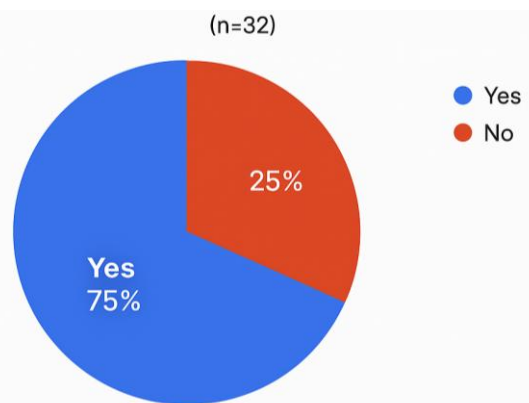


Figure 1. Students' Opinion Regarding Difficulties in Pronouncing English Word With Ending -ed

As can be seen from the statement above, the data shows that 75% of students have difficulty pronouncing English words with -ed ending, while students who do not have difficulty pronouncing English words with -ed ending are 25%. It can be concluded that students still have difficulty mastering the pronunciation of the -ed ending.

In the teaching process, students play a central role in determining learning outcomes, as they are the primary subjects of learning activities. Their achievement is largely influenced by motivation, willingness to learn, active participation, and the capacity to receive and process instructional materials. However, student-related factors are not the only determinants of success. Learning outcomes are also shaped by the quality of teaching, curriculum design, learning environment, availability of resources, and individual differences such as learning styles and preferences.

Teachers likewise hold a significant responsibility in shaping students' learning progress, particularly in pronunciation instruction. Their role extends beyond delivering material to include motivating learners, facilitating practice, and serving as pronunciation models. Effective pronunciation teaching requires teachers to design appropriate learning activities, select suitable materials, monitor progress, and provide constructive feedback (Anggraini, 2022; Dalton, 1994; Dziubalska-Kołaczyk, 2008). Yet, when teachers face excessive workloads, they may be unable to devote sufficient time and energy to the teaching process. This often results in less effective learning experiences and limited opportunities for students to receive the individualized support and feedback essential for developing accurate pronunciation mastery.

Speech recognition technology has increasingly been recognized as an effective tool to help improve and refine English pronunciation (Hongmei Yang and Xiaoli Wang, 2020). ELSA Speak differs from many other language learning applications that focus primarily on vocabulary and grammar, as it specifically incorporates speech recognition technology designed to provide immediate, individualized feedback on learners' pronunciation. With this feature, the application is reported to detect more than 95 percent of user pronunciation errors and subsequently provide corrective feedback to facilitate improvement (Van Dinh et al., 2022). This function distinguishes ELSA from other English language learning applications, as it not only identifies errors but also supplies actionable suggestions for refinement. Furthermore, ELSA Speak offers learners more than 1,200 lessons and over 60 topics that encompass a wide range of communicative contexts, enabling practice of both words and phrases relevant to learners' needs (Anggraini, 2022). Given these advantages, the present research proposes the integration of ELSA Speak as an instructional medium to support the teaching of English pronunciation, with a particular focus on the accurate production of the inflectional ending *-ed*.

Previous studies emphasize the importance of technology-assisted learning for pronunciation improvement (Rinaepi et al., 2022). While many applications focus on vocabulary and grammar, limited research has explored the role of intelligent speech recognition tools in addressing segmental pronunciation difficulties such as the *-ed* ending. ELSA Speak, an AI-powered application, offers individualized feedback and has shown effectiveness in improving learners' pronunciation in broader contexts, but its potential in addressing this particular issue has not been sufficiently examined.

Given these gaps, the present study investigates the implementation of ELSA Speak as a learning medium to overcome students' pronunciation errors of the *-ed* ending. This research is urgent because effective pronunciation instruction directly supports students' communicative competence, boosts their confidence in speaking, and contributes to their academic and professional readiness in a globalized era. Therefore this study identifies students' difficulties in pronouncing the *-ed* ending in implementing ELSA Speak as a digital pronunciation tool. This research investigated the improvement of the students' pronunciation of ending *-ed* at the seventh grade through the implement of the ELSA Speak Application, focusing on the accurate pronunciation of *-ed* endings in terms of sounds (/d/, /t/, /ɪd/). The findings are expected to contribute both theoretically and practically: theoretically, by enriching the literature on pronunciation pedagogy in EFL contexts; and practically, by providing teachers with an innovative and accessible medium to enhance students' learning outcomes.

Methods

This study used a Classroom Action Research (CAR) design following the model proposed by Kemmis and McTaggart (Creswell, 2012; Kemmis, S., & McTaggart, R., 1988), which consists of four stages: planning, action, observation, and reflection. The research was conducted in two cycles, with each cycle comprising two meetings. The second cycle was designed as a continuation and improvement of the first, addressing issues identified during the initial cycle. The main objective of applying CAR was to improve the learning process and to strengthen the professionalism of teachers in classroom practice. By adopting this cyclical approach, the study sought to create a systematic process of evaluation and refinement to enhance the overall quality of education.

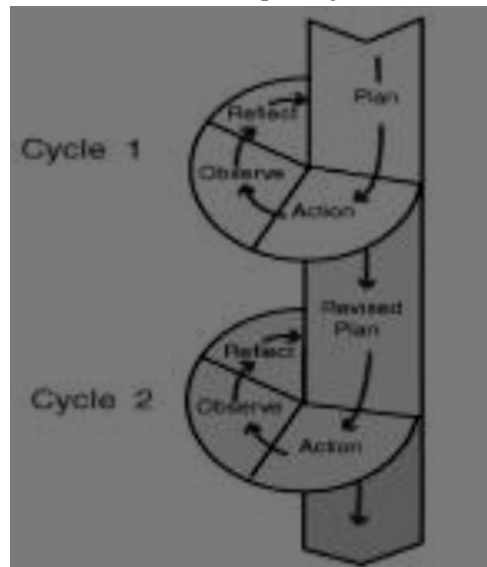


Figure 1. Action Research model according to Kemmis and Mc Taggart

Based on the description of action research by Kemmis and McTaggart, action research is a cyclical process consisting of planning, action, observation, and reflection, carried out in more than one cycle as needed. The number of cycles depends on the context and purpose of the research, and the process may be concluded once the targeted improvement in this case, students' pronunciation has been achieved (Burns, 2010). In line with the collaborative principle of Classroom Action Research, the researcher conducted the study in partnership with the English teacher of class VII at SMPN 1 Candipuro to design, implement, and evaluate each stage of the cycles.

Results and Discussion

Baseline Pronunciation Performance

The study involved 32 seventh-grade students at SMPN 1 Candipuro during the 2024/2025 academic year. Before the classroom action was implemented, a pre-test was administered on January 3, 2025, to identify the students' initial ability to pronounce regular English verbs ending in *-ed*. Each student read and recorded 20 regular verbs within approximately 25 minutes. The assessment focused on the three phonologically conditioned realizations of the *-ed* ending, namely /t/, /d/, and /ɪd/. The pre-test served as the baseline for evaluating changes in pronunciation performance throughout the two classroom-action cycles.

The baseline results indicated substantial difficulty with the target pronunciation feature. The total pre-test score was 960, with a mean of 30.00, a median of 35.00, and a standard deviation of

11.57. Scores ranged from 15 to 65, and none of the 32 students attained the predetermined minimum score of 75. This quantitative pattern corresponded with preliminary classroom data showing that 75% of students reported difficulty pronouncing English words ending in *-ed*. Typical problems included pronouncing all *-ed* endings in the same way, omitting the ending, and failing to distinguish accurately among /t/, /d/, and /ɪd/. The low and relatively widespread baseline performance confirmed that the problem was not restricted to a few learners and therefore justified targeted pronunciation instruction.

The difficulty observed at baseline can be understood in relation to the morphophonological characteristics of the English regular past-tense ending. Although the orthographic form remains *-ed*, its pronunciation changes according to the final sound of the verb stem. Learners therefore need to associate one written form with three different phonological realizations. Similar difficulties with recognizing and producing *-ed* allomorphs have been reported among Indonesian EFL learners (Giantari et al., 2020; Amalia et al., 2022). For this reason, ELSA Speak was integrated into the classroom action as a supporting tool that could provide pronunciation models, repeated oral practice, and immediate feedback on students’ production.

Implementation of the Classroom Action Research

The intervention consisted of two Classroom Action Research cycles, each incorporating planning, action, observation, and reflection. Cycle I was conducted on January 7 and 9, 2025, whereas Cycle II took place on January 10 and 14, 2025. The researcher collaborated with the English teacher throughout the intervention. In both cycles, students received explicit instruction on the pronunciation of *-ed* endings and then practiced the target forms using ELSA Speak. Pronunciation tests were used to assess learning outcomes, while classroom observations documented students’ attention, participation, independence, confidence, and responses to corrective feedback.

The first cycle introduced the relevant pronunciation rules and familiarized students with application-assisted practice. Reflection on the Cycle I results was subsequently used to refine the second cycle. Cycle II therefore emphasized more structured practice, repeated differentiation among /t/, /d/, and /ɪd/, and more explicit teacher guidance in interpreting feedback generated by ELSA Speak. Thus, the two cycles represented an iterative instructional process rather than simple repetition of the same activities.

Table 1. Schedule of the Classroom Action Research

Cycle	Meeting	Date
Cycle I	Meeting 1	January 7, 2025
Cycle I	Meeting 2	January 9, 2025
Cycle II	Meeting 1	January 10, 2025
Cycle II	Meeting 2	January 14, 2025

Note. The pre-test was administered on January 3, 2025. Each cycle consisted of planning, action, observation, and reflection.

Cycle I: Initial Improvement After ELSA Speak Integration

During the first meeting of Cycle I, the teacher explained the rules governing the pronunciation of regular past-tense *-ed* endings and modeled the /t/, /d/, and /ɪd/ realizations. Students subsequently practiced with ELSA Speak by listening to pronunciation models, producing target words, receiving application-generated feedback, and repeating inaccurate forms. In the second meeting, students completed Post-test I by reading 20 regular verbs aloud. The assessment followed

the same general pronunciation target as the pre-test so that changes in performance could be compared.

The results demonstrated considerable initial improvement. The total class score increased from 960 in the pre-test to 1,890 in Post-test I. The mean rose from 30.00 to 59.06, representing an average increase of 29.06 points, while the median increased from 35.00 to 60.00. Post-test I scores ranged from 15 to 100, with a standard deviation of 21.61. At the individual level, 28 students achieved higher scores than in the pre-test, while four students recorded no change. No student obtained a lower score.

Despite this improvement, Cycle I did not satisfy the predetermined classroom success criterion. Only seven students achieved a score of 75 or higher, representing 21.88% of the class, whereas 25 students (78.12%) remained below the minimum standard. Thus, although the substantial increase in the mean demonstrated early progress, most students had not yet developed the expected level of pronunciation accuracy.

Table 2. Descriptive Pronunciation Performance Across Assessment Stages

Assessment	n	M	Mdn	SD	Min	Max	≥ 75, n (%)
Pre-test	32	30.00	35.00	11.57	15	65	0 (0.00%)
Post-test I	32	59.06	60.00	21.61	15	100	7 (21.88%)
Post-test II	32	80.31	85.00	17.27	40	100	26 (81.25%)

Note. M = mean; Mdn = median; SD = standard deviation. The minimum mastery score was 75. Classroom success required at least 80% of students to attain this score.

Classroom observations clarified why a second cycle remained necessary. Students became more responsive and attentive and generally followed the instructions for using ELSA Speak. However, participation remained uneven, and several students experienced difficulty translating automated feedback into accurate production. Common errors included omission of the final *-ed*, overgeneralization of /d/ to words requiring another realization, and inaccurate production of /ɪd/. These findings indicated that the first cycle successfully increased students’ awareness of the target feature, but a single phase of instruction and application-based practice was insufficient to establish consistent control of the three allomorphs.

Reflection and Instructional Refinement

The reflection stage focused on the contrast between the substantial improvement in the class mean and the relatively small proportion of students reaching the minimum criterion. ELSA Speak was therefore retained in Cycle II, but its instructional use was refined. First, /t/, /d/, and /ɪd/ were treated more explicitly as separate pronunciation targets instead of presenting *-ed* words as one undifferentiated category. Second, practice became more structured and repetitive, with particular attention given to forms that had remained problematic during Cycle I. Third, teacher guidance was strengthened so that students were encouraged not merely to repeat an unsuccessful pronunciation but to identify the error and use the feedback to make a specific correction.

These modifications are important because the effectiveness of pronunciation technology depends substantially on how it is incorporated into instruction. Research on automatic speech recognition and mobile-assisted pronunciation learning indicates that technology is most productive when learners receive explicit feedback and repeated opportunities to modify their production (Dai & Wu, 2023; Ngo et al., 2024). Accordingly, Cycle II combined automated feedback with continued

teacher explanation and guided practice rather than treating the application as an independent instructional agent.

Cycle II: Consolidation of Pronunciation Accuracy

Cycle II was implemented over two meetings. On January 10, 2025, the teacher reviewed Post-test I results and discussed the pronunciation errors identified during the first cycle. Students reflected on words they had mispronounced and then completed customized ELSA Speak exercises focusing on distinctions that still caused difficulty. The activities included listening to model pronunciation, repeatedly producing target items, responding to corrective feedback, and presenting selected words orally in groups. Observation indicated that students had become more familiar with the application and were increasingly willing to repeat pronunciation attempts when their initial production was inaccurate.

Post-test II was administered during the second meeting on January 14, 2025. Students again read 20 regular verbs ending in *-ed*. The total class score reached 2,570, corresponding to a mean of 80.31 and a median of 85.00. The standard deviation was 17.27, with scores ranging from 40 to 100. Compared with Post-test I, the class mean increased by a further 21.25 points. At the individual level, 28 students improved their scores from Post-test I, three obtained the same score, and one student obtained a lower score. Thus, although the magnitude of improvement varied, the overall distribution shifted clearly toward higher pronunciation performance.

More importantly, Cycle II fulfilled the predetermined success criterion. Twenty-six of the 32 students achieved scores of at least 75, equivalent to 81.25% of the class, while six students (18.75%) remained below the standard. Because the study required at least 80% of students to attain a minimum score of 75, the result exceeded the established target. Consequently, the classroom action was concluded after Cycle II and no additional cycle was required.

Table 3. Achievement of the Minimum Pronunciation Criterion

Assessment	Students ≥ 75	%	Students < 75	%
Pre-test	0	0.00	32	100.00
Post-test I	7	21.88	25	78.12
Post-test II	26	81.25	6	18.75

Note. The minimum achievement score was 75. The classroom success criterion required at least 80% of students to achieve this score.

Overall Progression and Classroom Observation Findings

Across the complete intervention, pronunciation performance showed a consistent upward progression. The mean increased from 30.00 at baseline to 59.06 after Cycle I and 80.31 after Cycle II. The overall mean gain from the pre-test to Post-test II was therefore 50.31 points. All 32 students obtained higher final scores than their respective baseline scores, with individual gains ranging from 10 to 85 points. The percentage of students reaching the minimum criterion similarly increased from 0% to 21.88% and finally to 81.25%. The largest increase in mastery occurred between Cycle I and Cycle II, after the instructional strategy had been refined on the basis of classroom reflection.

Nevertheless, individual outcomes were not uniform. Several students obtained final scores between 90 and 100, whereas six students remained below 75. This variation demonstrates that the classroom-level intervention successfully met its target without eliminating individual differences in learning progress. Students who remained below the criterion would therefore benefit from

continued individualized pronunciation support even though the formal action-research success criterion had been achieved.

The quantitative changes were accompanied by observable developments in classroom behavior. During Cycle I, students became more attentive but continued to rely considerably on teacher guidance, particularly when interpreting feedback from the application. During Cycle II, they demonstrated greater independence in following instructions and more active participation in application-based practice. Students were also increasingly willing to repeat words after receiving feedback and appeared more aware of the accuracy of their own production.

Table 4. Summary of Classroom Observation Findings Across the Two Cycles

Observed Aspect	Cycle I	Cycle II
Attention and focus	Attention improved, but consistency remained limited.	Students demonstrated stronger and more sustained focus.
Participation	Participation increased but remained uneven.	Students participated more actively in pronunciation practice.
Use of feedback	Some students struggled to interpret and apply feedback.	Students increasingly used feedback to revise pronunciation attempts.
Pronunciation accuracy	Errors remained in omission and differentiation of /t/, /d/, and /ɪd/.	Errors decreased, although difficulties persisted for some students.
Independence and confidence	Considerable teacher guidance was still required.	Students followed instructions more independently and responded more confidently.

Note. The findings were derived from classroom observation records. Motivation and confidence were not measured using standardized psychological instruments.

Discussion

This study examined the use of ELSA Speak within a two-cycle Classroom Action Research (CAR) framework to improve seventh-grade students’ pronunciation of the three regular past-tense *-ed* allomorphs, namely /t/, /d/, and /ɪd/. The results showed a clear progression across the intervention stages. The mean score increased from 30.00 in the pre-test to 59.06 after Cycle I and 80.31 after Cycle II. In parallel, the proportion of students achieving the minimum score of 75 rose from 0% at baseline to 21.88% after Cycle I and 81.25% after Cycle II, exceeding the predetermined classroom success criterion of 80%. These results indicate substantial improvement during the intervention. However, because the study used CAR rather than a controlled experimental design, the gains should be interpreted as improvement occurring during a teacher-guided ELSA Speak intervention rather than conclusive evidence that the application alone caused the changes.

The students’ initial difficulties can be explained by the morphophonological characteristics of English regular past-tense endings. Although *-ed* has a single written form, it is realized as /t/, /d/, or /ɪd/ depending on the final sound of the verb stem. Learners must therefore connect orthographic knowledge with phonological conditioning. Baseline errors, including omission and overgeneralization, are consistent with Giantari et al. (2020), who reported difficulties among Indonesian EFL learners in producing *-ed* endings. The results also correspond with Kruk and Pawlak (2023), who showed that targeted internet-supported practice can facilitate learning of English past-tense *-ed* pronunciation. These similarities suggest that the students’ errors reflected a broader

challenge in learning a morphophonological feature that requires explicit differentiation and repeated oral practice.

The improvement across the two cycles is also consistent with research on Automatic Speech Recognition (ASR) and Mobile-Assisted Language Learning (MALL). Ngo et al. (2024) reported a positive overall effect of ASR-supported pronunciation instruction, particularly when learners received explicit corrective feedback and practiced clearly defined segmental targets. Tseng et al. (2022) likewise found that mobile-assisted pronunciation learning generally improves L2 pronunciation outcomes. These findings are relevant because /t/, /d/, and /ɪd/ are narrow segmental targets suitable for repeated practice and immediate feedback. The progression from Cycle I to Cycle II therefore supports the view that the value of ELSA Speak lies not simply in the use of artificial intelligence, but in its capacity to provide repeated practice and immediate information about pronunciation performance.

The difference between Cycle I and Cycle II further highlights the importance of instructional design. Cycle I produced a substantial increase in the class mean, yet most students still failed to achieve the minimum score. Reflection on these results led to more structured practice in Cycle II, with greater emphasis on differentiating the three *-ed* realizations and helping students interpret application-generated feedback. The stronger outcomes in the second cycle indicate that technology alone was insufficient; its use became more productive when embedded in teacher-guided explanation, repetition, and corrective practice. This interpretation is supported by Dai and Wu (2023), who found that pronunciation development can be strengthened when ASR feedback is combined with peer or instructional support. Sun (2023) similarly reported positive effects when automatic speech recognition was integrated with corrective interaction. Thus, the present study reinforces the teacher's role in helping students translate automated feedback into meaningful pronunciation improvement.

Classroom observations indicated that students became more attentive, more independent in following instructions, and more willing to repeat pronunciation attempts during Cycle II. These changes are compatible with previous ELSA Speak studies in Indonesia. Karim et al. (2023) reported improvements in speaking performance and positive learner perceptions, while Permatasari and Lubis (2024) found stronger pronunciation improvement among students using ELSA Speak. The present study extends this evidence by focusing specifically on *-ed* allomorphs, demonstrating how AI-supported practice can address a clearly defined grammatical-phonological problem in a junior-high-school context.

The observed increase in confidence and participation should nevertheless be interpreted cautiously because these outcomes were documented through classroom observation rather than validated psychological instruments. The findings therefore support greater classroom engagement, but not long-term changes in motivation or anxiety. Metruk (2024) emphasizes that the effectiveness of mobile-assisted pronunciation instruction varies according to instructional design, learner characteristics, treatment duration, and the type of feedback provided. The present findings align with this position: ELSA Speak appeared useful when integrated into structured classroom practice, but it should not be treated as a self-sufficient replacement for teacher-led pronunciation instruction.

The two-cycle CAR design also offers a methodological contribution. Improvement was not complete after the first cycle; instead, reflection on students' remaining difficulties informed instructional modification in Cycle II. This process illustrates the logic of action research described

by Burns (2010), in which teaching practices are refined through cycles of action, observation, and reflection. The novelty of the study therefore lies not in claiming that ELSA Speak is superior to other pronunciation tools, but in demonstrating a contextually grounded implementation pathway that aligns automated feedback, teacher reflection, repeated learner practice, and a specific pronunciation target.

Several limitations should be considered. First, the study involved only 32 students from one school and did not include a control group or random assignment. The improvement therefore cannot be attributed exclusively to ELSA Speak because repeated testing, teacher attention, classroom instruction, and familiarity with the target words may also have contributed. Second, the intervention covered only two cycles and did not include a delayed post-test, so the durability of the gains remains unknown. Rahimi et al. (2020) showed that immediate improvement in regular past-tense learning does not necessarily guarantee long-term retention. Third, the assessment focused on targeted word-level pronunciation and therefore does not demonstrate transfer to spontaneous speech. Future research should employ delayed assessment, unfamiliar items, spontaneous speaking tasks, and comparison groups to examine retention and communicative transfer.

Overall, the findings indicate that ELSA Speak can function as a useful supporting tool for teaching regular past-tense *-ed* pronunciation when integrated into structured and reflective classroom instruction. The most important result is not merely the increase in scores, but the pattern showing that stronger outcomes emerged after the instructional approach was refined in Cycle II. The study therefore suggests that AI-supported pronunciation technology is most valuable when combined with explicit phonological explanation, repeated practice, corrective feedback, and continued teacher mediation.

Conclusion

This study investigated the integration of ELSA Speak within a two-cycle Classroom Action Research framework to improve seventh-grade students' pronunciation of the three regular past-tense *-ed* allomorphs, namely /t/, /d/, and /ɪd/. The findings demonstrate a consistent improvement in students' pronunciation performance across the intervention, with the mean score increasing from 30.00 at baseline to 59.06 after Cycle I and 80.31 after Cycle II, while the proportion of students achieving the minimum score of 75 increased from 0% to 21.88% and ultimately to 81.25%, thereby exceeding the predetermined classroom success criterion. Classroom observations further indicated improvements in students' attention, participation, independence, confidence, and ability to respond to corrective feedback. These findings address the research objective by showing that ELSA Speak can serve as an effective pedagogical support for addressing the specific morphophonological difficulty of distinguishing /t/, /d/, and /ɪd/ when it is embedded within explicit, structured, and reflective pronunciation instruction. Theoretically, the study contributes to technology-assisted pronunciation learning by demonstrating how automatic speech recognition, repeated production, corrective feedback, and teacher mediation can operate synergistically within an iterative classroom-action process; methodologically, it illustrates the value of reflection between cycles for refining technology-enhanced instruction in response to learners' emerging difficulties. Practically, the findings suggest that English teachers may incorporate ELSA Speak as a supplementary pronunciation tool while maintaining active pedagogical guidance to help learners interpret automated feedback and develop more accurate production. Nevertheless, the findings should be

interpreted within the study's limitations, including the relatively small sample from a single school, the absence of a comparison group and random assignment, the short two-cycle intervention, and the lack of delayed assessment or spontaneous-speech measures. Future research should therefore examine larger and more diverse populations, employ quasi-experimental or experimental comparisons, incorporate delayed post-tests and unfamiliar pronunciation items, and investigate whether improvements transfer from controlled word-level practice to intelligibility and accuracy in spontaneous communication. Overall, this study demonstrates that AI-supported pronunciation technology is most educationally valuable not as a substitute for teachers, but as a complementary tool within purposeful, feedback-rich, and continuously refined instruction.

References

- Amalia, I. S., Setiawan, S., Anam, S., & Purwati, O. (2022). Challenges and problems faced by EFL teachers in teaching allomorph through narrative text. *Vivid: Journal of Language and Literature*, 11(1), 80–85. <https://doi.org/10.25077/vj.11.1.80-85.2022>
- Anggraini, A. (2022). Improving students' pronunciation skill using ELSA Speak application. *Journey: Journal of English Language and Pedagogy*, 5(1), 135–141. <https://doi.org/10.33503/journey.v5i1.1840>
- Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford University Press.
- Burns, A. (2010). *Doing action research in English language teaching*. Taylor & Francis e-Library.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.
- Creswell, J. W. (2012). *Educational research*. Pearson Education.
- Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed-methods study. *Computer Assisted Language Learning*, 36(5–6), 861–884. <https://doi.org/10.1080/09588221.2021.1952272>
- Dalton, C., & Seidlhofer, B. (1994). *Pronunciation*. Oxford University Press.
- Dziubalska-Kořaczyk, K., & Przedlacka, J. (2008). *English pronunciation models: A changing scene*. Peter Lang Publishing.
- Giantari, K., Sabarudin, S., & Zahrida, Z. (2020). Pronunciation recognition of –ed ending words by the students of English Education Study Program of the University of Bengkulu. *Journal of English Education and Teaching*, 4(2), 278–293. <https://doi.org/10.33369/jeet.4.2.278-293>
- Gilakjani, A. P. (2017). English pronunciation instruction: Views and recommendations. *Journal of Language Teaching and Research*, 8(6), 1249–1255.
- Hakim, M. A. R. (2012). An analysis of phonetics b, d, g, j, z, 3, ð into English pronunciation for Java students. *International Journal of Social Science and Education*, 2(4), 706–716.
- Harmer, J. (2007). *How to teach English*. Longman.
- Karim, S. A., Hamzah, A. Q. S., Anjani, N. M., Prianti, J., & Sihole, I. G. (2023). Promoting EFL students' speaking performance through ELSA Speak: An artificial intelligence in English language learning. *Journal of Languages and Language Teaching*, 11(4), 655–668. <https://doi.org/10.33394/jollt.v11i4.8958>
- Kemmis, S., & McTaggart, R. (1988). *The action research planner*. Deakin University Press.
- Kruk, M., & Pawlak, M. (2023). Using internet resources in the development of English pronunciation: The case of the past tense -ed ending. *Computer Assisted Language Learning*, 36(1–2), 205–237. <https://doi.org/10.1080/09588221.2021.1907416>
- Kurniawan, D. (2016). The influence of first language toward students' pronunciation. *Journal of English and Education*, 4(1), 45–53.
- Metruk, R. (2024). Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, 29(13), 16255–16282. <https://doi.org/10.1007/s10639-024-12453-0>
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481–520.
- 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>

- Permatasari, S., & Lubis, Y. (2024). Enhancing pronunciation skills in EFL students through the ELSA Speak application. *Indonesian EFL Journal*, 10(2), 205–212. <https://doi.org/10.25134/ieflj.v10i2.10137>
- Rahimi, S., Ahmadian, M., Amerian, M., & Dowlatabadi, H. R. (2020). Comparing accuracy and durability effects of Jigsaw versus input flood tasks on the recognition of regular past tense /-ed/. *SAGE Open*, 10(2). <https://doi.org/10.1177/2158244020919505>
- Rinaepi, R., et al. (2022). The effectiveness of ELSA Speak application to improve pronunciation ability. *Jurnal Fakultas Keguruan dan Ilmu Pendidikan*, 3(1), 28–33.
- Roach, P. (1991). *English phonetics and phonology* (2nd ed.). Cambridge University Press.
- Sholeh, A., & Muhaji. (2015). Pronunciation difficulties encountered by EFL students in Indonesia: Sebuah studi kasus pada mahasiswa kelas Integrated Course semester 1 FKIP Bahasa Inggris Universitas Kanjuruhan Malang. *Jurnal Inspirasi Pendidikan*, 5(2), 698–707.
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, 14, Article 1210187. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Tseng, W.-T., Chen, S., Wang, S.-P., Cheng, H.-F., Yang, P.-S., & Gao, X. A. (2022). The effects of MALL on L2 pronunciation learning: A meta-analysis. *Journal of Educational Computing Research*, 60(5), 1220–1252. <https://doi.org/10.1177/07356331211058662>
- Van Dinh, N., & Nguyen, T. (2022). Mobile-assisted language learning: The case of ELSA Speak in improving English pronunciation. *International Journal of Emerging Technologies in Learning*, 17(8), 203–213. <https://doi.org/10.3991/ijet.v17i08.30215>
- Yang, H., & Wang, X. (2020). Application of intelligent speech recognition in college English pronunciation teaching. *Journal of Physics: Conference Series*, 1533(2), Article 022091. <https://doi.org/10.1088/1742-6596/1533/2/022091>