Global Synthesis in Education Journal
https://gse-journal.net/index.php/gse
Vol. 03. No. 03. (2025) 23-46
ISSN:3025-5724, DOI: https://doi.org/10.61667/dj19b884

Evaluation of Elsa Speak Application for Teaching English Pronunciation at the 10th Grade of Automotive Engineering Class of SMK Muhammadiyah Prabanan: A Cipp Model

Akmal Tanjung*[1] Bunga Satya hardika[2], Intan Suzila[3]

1.2, Graduate Program of English Language Education, Faculty of Teacher Training and Education, Ahmad Dahlan University, Yogyakarta. Indonesia.

3. Universiti Teknologi Mara Cawangan,Pahang,Malaysia

Article Info

Received: 04/0/2024
Accepted: 12/10/2025
Published: 29/11/2025

Keywords:
Learning management system, blended learning course, English language teaching and learning

Abstract

Pronunciation, a key indicator of speaking ability in the Common European Framework of Reference (CEFR), is crucial for accurate communication of technical terms; while native-like pronunciation is unnecessary, minimizing errors prevents misunderstandings. This study examined the effectiveness of the ELSA Speak mobile application—designed specifically for English pronunciation—in the English program of Prambanan Vocational High School, Yogyakarta, where it has been used for more than two years but has not been formally evaluated. A purposive sample of 76 students (out of a total of 380) was selected and the research applied the CIPP evaluation model (Context, Input, Process, Product) together with a quasi-experimental pre-test/post-test design. Four research questions guided the inquiry: (1) the needs of current students and alumni for English communication skills; (2) the resources, planning, and readiness of school management, teachers, and learners for implementing ELSA Speak; (3) the challenges encountered by students, alumni, and teachers during use; and (4) the degree of pronunciation improvement attributable to the app. Contextual analysis revealed a strong demand for English, especially Received Pronunciation, to enhance career prospects. Input evaluation identified barriers such as limited internet access, mother-tongue interference, and phonetic complexity. Process results showed statistically significant gains in pronunciation (t = 7.885, p < 0.001) and increased learner engagement (t = 4.88, p < 0.001). Product outcomes highlighted notable improvements in word-level intonation and stress patterns. The school addressed input challenges through offline materials, targeted teacher training, and clear usage guidelines, which contributed to the observed performance gains and positive student feedback. The study concludes that ELSA Speak markedly enhances pronunciation for vocational learners and recommends its continued adoption with structured support. Future research should involve larger, multi-major samples to confirm generalizability.

*corresponding author: akmal@mpbi.uad.ac.id

APA Citation: Akmal T, Bunga S. H., Intan S. (2025). Evaluation of Elsa Speak Application for Teaching English Pronunciation at the 10th Grade of Automotive Engineering Class of SMK Muhammadiyah Prambanan: A Cipp Model*: Global Synthesis in Education Journal* (3), 23-46. DOI: https://doi.org/10.61667/dj19b884

1. INTRODUCTION

English has been formally mandated as a core subject in Indonesian vocational high schools through Law No. 20/2003 on the National Education System, Government Regulation No. 13 of 2015 on national education stand-ards, and Ministry of Education and Culture Regulation No. 22 of 2016 on process standards, all of which are sustained in the 2022 Merdeka Curriculum. These regu-lations position English as a strategic competence to support students' transition to globalized workplaces, particularly in sectors that require engagement with in-ternational partners and technical documentation. In this context, vocational schools (SMK) are expected to de-velop students' technical English and oral communica-tion skills rather than focusing primarily on read-ing-oriented academic literacy. Pronunciation is a critical component of speaking proficiency and a key indicator in frameworks such as the Common European Framework of Reference (CEFR), because inaccurate production of technical terms can easily lead to misunderstanding and diminish perceived professionalism in workplace com-munication. While native-like accents are not re-quired—given the legitimacy of diverse "World Englishes" such as Singaporean, Malaysian, or Indonesian Eng-lish—students are expected to achieve intelligible pro-nunciation and minimize errors that could interfere with meaning.

Despite this policy emphasis, many Indonesian voca-tional students continue to experience serious challenges in pronouncing English accurately, especially when deal-ing with verb phrases and segmental features that are absent or rare in their first languages. For example, verb phrases containing the particle "up" (e.g., "set up," "pick up," "clean up") occur frequently in technical instructions in automotive contexts and must be pronounced clearly to ensure safe and precise task execution. Similarly, stu-dents in vocational high schools often struggle with the English /u/ and /ju:/ sounds in words such as "utilize," "upgrade," "tune," "clutch," or "flush," as these sounds and related consonant clusters have no straightforward equivalents in Indonesian phonology. These phonological mismatches lead students to substitute unfamiliar sounds with those from their first language, producing fossilized errors that reduce intelligibility and may un-dermine confidence in professional settings. Studies of Indonesian learners further highlight that regional lan-guages such as Javanese and Sundanese shape systemat-ic pronunciation difficulties (e.g., confusion between /f/ and /p/ or between /v/ and /w/), reinforcing the need for explicit, context-sensitive pronunciation instruction in vocational classrooms.

Teachers, however, face multiple constraints in providing systematic pronunciation training. Pronuncia-tion is often marginalized compared with grammar, reading, and writing, and many teachers report limited formal preparation in phonetics and phonology. Moreo-ver, conventional classroom time and resources are in-sufficient to address individual pronunciation problems, especially in large vocational classes. At the same time, the rapid growth of educational technology has opened access to a wide range of computer-based and mobile resources, including online dictionaries, audio-visual materials, and speech-recognition tools, which can offer individualized practice beyond the classroom. Yet not all teachers and students can fully exploit these resources, due to time constraints, lack of training, uneven digital literacy, and infrastructure issues such as unstable in-ternet connections. These conditions create a gap be-tween curricular expectations for communicative com-petence and the actual support available to vocational learners for improving their pronunciation.

The ELSA Speak application (English Language Speech Assistant) has emerged as a promising Mobile-Assisted Language Learning (MALL) tool to help bridge this gap. ELSA Speak is an artificial

intelligence–driven mobile app designed specifically for pronunciation practice; it uses automatic speech recognition to analyze learners' speech and provide phoneme-level feedback, focusing on segmental features (e.g., minimal pairs, consonant clus-ters, schwa, /θ/ and /ð/) and suprasegmental aspects such as stress, rhythm, and intonation. The app offers personalized learning paths based on diagnostic tests, real-time feedback, thousands of high-frequency words, and gamified activities that can be accessed anytime via smartphones. Prior descriptions emphasize that ELSA Speak allows learners to repeatedly record, listen to, and correct their pronunciation in an interactive, self-paced environment, thereby supporting autonomy and sus-tained practice beyond class hours. At the same time, the effective use of ELSA Speak in schools depends on tech-nical prerequisites (adequate internet connectivity, de-vice storage, and quiet environments) and on teacher mediation, such as guidance in navigating English-only interfaces and interpreting automated feedback.

A growing body of research has investigated ELSA Speak in various EFL contexts, but several limitations remain, particularly regarding vocational education. Fakdawer (2020) reported that the app can foster learn-er autonomy by enabling university students to set goals and monitor their pronunciation progress independently, but also noted challenges related to microphone quality, unstable internet, paywalled content, and competition with more attractive social media platforms; importantly, many learners still expressed a need for teacher support. Mahmudah and Daulay (2024) found that senior high school students perceived ELSA Speak as having strong content quality and helpful real-time feedback, yet half of the participants questioned the accuracy and accent flexibility of its assessment, and the small, self-selected sample limited generalizability. Research with Indone-sian higher-education students by Maghdalena and Wahyuningsih (2024) showed that ELSA Speak enhanced speaking confidence and pronunciation, but the study involved only one institution and relied heavily on self-report data. Anggraini's (2022) classroom action research documented significant gains in pronunciation scores across cycles when ELSA Speak was integrated into instruction, while also warning about dependence on technology, the necessity of careful pedagogy, and the importance of learners' readiness. Other studies have examined ELSA Speak in blended or flipped learning en-vironments and emphasized its potential as a supple-mentary pronunciation tool rather than a replacement for teacher-led instruction. Collectively, these studies suggest that ELSA Speak can improve learners' pronun-ciation and motivation, but evidence is fragmented across levels and contexts, and rigorous experimental evaluations in vocational high schools, particularly with-in the Merdeka Curriculum framework, remain scarce.

At SMK Muhammadiyah Prambanan in Yogyakarta, ELSA Speak has been integrated into English instruction for more than two years, especially in the tenth-grade Automotive Engineering program, yet no systematic evaluation has been conducted to determine its effec-tiveness for vocational pronunciation needs. Preliminary observation and teacher reports indicate persistent problems in pronouncing verb phrases, especially those with the particle "up," and in producing English /u/ and related clusters in automotive procedure texts, despite regular exposure to English. In addition, students report limited opportunities to practice with native or near-native models, a lack of specialized pronunciation materials aligned with vocational content, and con-straints in using digital tools independently. These con-ditions define the core problem addressed in this study: the misalignment between the curriculum's demand for intelligible, profession-oriented spoken English and the insufficiently evaluated, technologically mediated sup-port available for pronunciation learning in a vocational setting.

To address this problem, the present study focuses specifically on the implementation of ELSA Speak for teaching English pronunciation to tenth-grade Automo-tive Engineering students at SMK Muhammadiyah Prambanan. The scope is deliberately limited to one ac-ademic year, two intact classes (X TOC as the experi-mental group and X TOB as the control group), and the pronunciation of vocabulary embedded in vocational procedure and descriptive texts aligned with phase E learning outcomes. The research adopts a qua-si-experimental non-equivalent group design with pre-test and post-test measures, combined with a CIPP (Context, Input, Process, Product) evaluation framework to capture not only learning outcomes but also needs, resources, implementation processes, and perceived impacts. This delimitation recognizes that findings will primarily generalize to similar vocational settings and proficiency levels, while more diverse samples and in-stitutions would be required for broader external validi-ty.

Guided by this design, the study seeks to answer four research questions: (1) How effective is the ELSA Speak app in improving English pronunciation among X TOC students compared with a control class? (2) What obsta-cles do teachers encounter when using ELSA Speak to teach pronunciation? (3) What solutions do teachers im-plement to overcome these obstacles? and (4) How do teachers perceive the effectiveness of these solutions in supporting the integration of ELSA Speak? These ques-tions correspond to four objectives: (1) to describe the effectiveness of ELSA Speak in teaching English pronun-ciation to X TOC students; (2) to identify the main obsta-cles faced in its classroom use; (3) to analyze the solu-tions adopted by teachers to address technical and ped-agogical challenges; and (4) to evaluate how effective these solutions are from the teachers' perspective.

This study is expected to yield several contributions. For students, it aims to demonstrate how a mobile, AI-based application can extend pronunciation practice beyond the classroom, increase awareness of useful dig-ital tools, and enhance confidence in speaking English in public and professional contexts. For teachers, the find-ings can inform diagnostic practices for assessing stu-dents' speaking abilities, guide the selection and design of methods and media that better match vocational learners' needs, and provide a concrete model for inte-grating pronunciation technology into lesson planning. For researchers, the study offers empirical evidence from a quasi-experimental and CIPP-based evaluation of ELSA Speak in a vocational high school, thereby enriching the literature on MALL and pronunciation teaching in un-derexplored contexts and suggesting avenues for further inquiry.

In this study, ELSA Speak is operationally defined as an artificial intelligence–powered mobile application de-signed to support English pronunciation learning through automatic speech recognition, phoneme-level feedback, and structured practice activities accessible via smartphones. It is treated as a form of Mo-bile-Assisted Language Learning that enables two-way interaction: learners speak into the device, the app ana-lyzes their output, and it delivers corrective feedback and recommendations for improvement. Teaching English pronunciation is operationally understood as instruc-tional practices that help learners perceive and produce segmental and suprasegmental features of English accu-rately and intelligibly, drawing on articulatory, auditory, cognitive, communicative, prosodic, multisensory, task-based, and lexical approaches as outlined by Starkey (2008) and subsequent work. Within this introduction, these operational definitions provide a shared frame for interpreting the role of ELSA Speak as both a technolog-ical tool and a pedagogical resource in addressing voca-tional students' pronunciation needs

2. Literature review

A theoretical framework provides the conceptual lens through which a study defines its key constructs, organ-izes prior knowledge, and explains the expected rela-tionships among variables (Creswell & Creswell, 2018). In quantitative and mixed-methods research, theory serves as the "rainbow" that links independent and de-pendent variables, offering a rationale for why one is ex-pected to influence the other (Labovitz & Hagedorn, 1971; Creswell, 2023). In turn, the framework informs the re-search design, the selection of instruments, and the in-terpretation of findings. Within this study, the theoretical framework integrates perspectives from pronunciation research, English for specific and vocational purposes, mobile-assisted language learning (MALL), technology integration in schools, and program evaluation (CIPP model), culminating in a conceptual model that situates the ELSA Speak application as a technology-mediated intervention for improving vocational students' pronun-ciation.

2.1 Pronunciation, Communicative Competence, and Intelligibility

Modern language teaching views pronunciation as one of the core components of communicative competence, alongside grammar, vocabulary, discourse, and pragmat-ics (Canale & Swain, 1980; Celce-Murcia et al., 2010). In frameworks such as the Common European Framework of Reference for Languages (CEFR), pronunciation is ex-plicitly included under the descriptor of "phonological control," with emphasis on intelligible segmental pro-duction and appropriate stress, rhythm, and intonation rather than on native-like accent. This shift reflects what Levis (2005, 2018) calls the "intelligibility principle": the primary pedagogical goal is speech that is understanda-ble to interlocutors, not the elimination of all traces of a foreign accent.

Empirical work with adult learners shows that com-prehensibility and intelligibility are only moderately correlated with accentedness; many speakers with no-ticeable accents are easily understood if key segmental and suprasegmental features are produced accurately (Derwing & Munro, 2015). Jenkins's (2000) proposal of a Lingua Franca Core further underscores that in interna-tional communication, especially among non-native speakers, certain phonological contrasts (e.g., vowel length, some consonant distinctions, nuclear stress placement) are more critical to intelligibility than others. In vocational contexts, where mispronounced technical items may result in costly or dangerous misunderstand-ings (e.g., "shut down" vs. "start up"), intelligible pronun-ciation of domain-specific lexis becomes particularly salient.

In the Indonesian EFL context, learners' first and re-gional languages (e.g., Indonesian, Javanese, Sundanese) shape systematic pronunciation challenges such as diffi-culty with final consonants, consonant clusters, certain vowels (/uː/, /ʊ/, /ɜː/), or fricatives (/v/, /f/, /θ/, /ð/). These difficulties are well-documented in research on Indonesian learners and are often traced to phonological transfer from L1 and L2, as well as to limited exposure to authentic spoken input. In vocational high schools (SMK), these general EFL challenges intersect with the need to pronounce field-specific terminology (e.g., automotive, hospitality, engineering) accurately in interactions with supervisors, colleagues, and clients from diverse linguis-tic backgrounds. Hence, the present study adopts the intelligibility principle as its overarching phonological stance: students are not expected to achieve native-like Received Pronunciation but should reduce error patterns likely to impede understanding in technical and ser-vice-related communication.

2.2 Principles of Pronunciation Teaching in EFL and Vocational Contexts

The field of pronunciation pedagogy has evolved from mechanical drills and minimal pairs to more communi-cative, integrated approaches that combine attention to form with meaningful use (Celce-Murcia et al., 2010). Contemporary frameworks distinguish between seg-mental features (consonants, vowels, clusters) and su-prasegmental features (word and sentence stress, rhythm, intonation), arguing that both require pedagog-ical attention but may differ in relative weight depending on learners' needs and contexts. In line with Derwing and Munro's (2015) recommendations, effective pronuncia-tion instruction typically includes (a) explicit infor-mation about how sounds are articulated, (b) focused perception training to sharpen learners' ability to dis-tinguish contrasts, (c) guided production practice with feedback, and (d) opportunities to use target features in communicative tasks.

Pronunciation teaching is also informed by broader language learning theories. The noticing hypothesis (Schmidt, 1990) posits that learners must consciously notice phonological gaps between their production and target forms for lasting change to occur. Interactionist views (Long, 1996) emphasize that negotiation of meaning and feedback in communicative interaction can draw attention to problematic forms and prompt modi-fied output. Social constructivist perspectives (Vygotsky, 1978) highlight the role of scaffolding within learners' "zone of proximal development," suggesting that tech-nology and teachers can jointly provide stepwise support from controlled practice to autonomous use.

In vocational English, pronunciation instruction is sit-uated within English for Specific Purposes (ESP) and English for Vocational Purposes (EVP). Here, the focus lies on language tasks, genres, and lexis that are directly relevant to the workplace (Hutchinson & Waters, 1987). Pronunciation work, therefore, should target high-frequency technical terms, collocations, and proce-dural language that students will encounter in manuals, safety instructions, customer interactions, and on-the-job communication. Task-based approaches, in which learners complete authentic tasks (e.g., explaining a procedure, giving safety instructions, handling a cus-tomer complaint) while receiving pre-task focus on problem sounds and post-task feedback, are particularly compatible with EVP.

The Merdeka Curriculum and prior regulations for SMK emphasize speaking and listening for workplace readi-ness, positioning pronunciation as an integral part of communicative competence. However, classroom reali-ties—large classes, exam-oriented culture, limited teacher training in phonetics, and scarcity of authentic oral models—often constrain teachers' ability to provide individual, sustained pronunciation feedback. This ten-sion between curricular expectations and practical limi-tations creates a rationale for integrating technology, particularly tools that can deliver individualized, repeat-ed feedback on pronunciation outside classroom time.

2.3 Mobile-Assisted Language Learning and AI-Based Pronunciation Tools

Mobile-Assisted Language Learning (MALL) leverages the ubiquity and portability of smartphones to support language learning anytime and anywhere (Ku-kulska-Hulme & Shield, 2008). From a theoretical per-spective, MALL aligns with constructivist and so-cio-cultural views in which learning is distributed across people, tools, and contexts, and with self-regulated learning frameworks that emphasize learners' active role in planning, monitoring, and evaluating their learning (Zimmerman, 2002). In medical education, for example, self-regulated learning is seen as a cyclical process in which learners set goals, select strategies, monitor pro-gress, and adjust behavior based on feedback. Similar processes underlie effective use of MALL: learners can set pronunciation goals,

choose practice modules, re-ceive automated feedback, and adjust their productions accordingly.

Automatic speech recognition (ASR)–driven tools, of-ten categorized under Computer-Assisted Pronunciation Training (CAPT), provide a specific instantiation of MALL for pronunciation. They record learners' speech, compare it to target models, and provide feedback at the word or phoneme level. Studies have shown that CAPT can en-hance segmental accuracy, rhythm, and overall compre-hensibility when integrated into classroom practice, par-ticularly because it allows for high-frequency, individu-alized repetition and immediate feedback—something difficult to achieve in traditional teacher-fronted instruc-tion. CAPT tools also support the noticing process by visually highlighting errors, replaying mispronounced segments, and offering articulatory instructions, thereby making phonological features more salient.

ELSA Speak belongs to this family of ASR-powered tools but is distinctive in its focus on English pronuncia-tion, AI-driven personalization, and gamified learning paths. From a theoretical standpoint, ELSA Speak can be seen as a mediational tool that (a) increases exposure to native-speaker models, (b) scaffolds learners' production via incremental difficulty and targeted feedback, (c) promotes self-regulation and learner autonomy, and (d) potentially increases motivation through interactive, goal-oriented practice. Within Vygotskyan terms, the app functions as a "more capable peer" that guides learners through their zone of proximal development for pronun-ciation, while teachers play a complementary role in or-chestrating tasks, interpreting data, and aligning app use with curricular goals.

However, technology itself does not guarantee learn-ing. Effective MALL and CAPT use depends on learners' digital literacy, access to devices and internet, the quality and appropriateness of input, the transparency and ac-curacy of feedback, and integration with teacher-led in-struction. Research on MALL in formal education under-scores that teacher mediation, institutional support, and alignment with assessment practices are crucial for sus-taining innovations (Stockwell & Hubbard, 2013). These insights connect directly to the next component of the framework: technology integration at the school system level.

2.4 Technology Integration and Whole-School Capacity for Innovation

The implementation of digital tools in education is shaped not only by individual teachers' skills and beliefs but also by broader school-level capacities and leader-ship. Drawing on an in-depth case study of an innovative primary school in Hong Kong, Tong, Trinidad, and col-leagues developed a whole-school systems model of ICT integration that highlights four interrelated dimensions: teacher ICT professional attributes, school ICT capacity, school system capacity, and school leadership. The model conceptualizes how these dimensions jointly influence the quality of the learning environment and the extent to which ICT can support "innovative pedagogical practices" such as learner-centered, knowledge-centered, assess-ment-centered, and community-centered activities.

Teacher ICT professional attributes encompass teach-ers' visions for using technology, their integration prac-tices, ICT capabilities, and attitudes. When teachers per-ceive technology as genuinely enhancing pedagogy and feel confident in using it, they are more likely to adopt innovative practices rather than merely digitizing tradi-tional methods. School ICT capacity refers to infrastruc-ture, hardware and software, digital resources, and tech-nical support. Reliable infrastructure, appropriate devic-es, well-organized digital materials, and just-in-time technical support are conceptualized as necessary or suf-ficient conditions for innovation.

School system capacity involves incentives and sup-port systems, professional development, cultural change, curriculum support, and community connections. For example, school-based professional development fo-cused not only on ICT skills but also on pedagogy is seen as a sustainable condition for innovation, while clear or-ganizational structures and defined roles support ac-countability and collaboration. School leadership, finally, includes the school's vision and policy for ICT, the prin-cipal's instructional leadership, and participatory, inno-vative decision-making processes. Tong et al. categorize conditions into necessary (without which innovations cannot occur), sufficient (which facilitate but do not guarantee innovation), and sustainable (which help maintain and deepen innovations over time).

This whole-school perspective is directly relevant to the present study, which investigates ELSA Speak im-plementation in a vocational high school. The success of such an app does not depend solely on the app's design or on individual teacher enthusiasm but on the alignment of teacher attributes, technical infrastructure (e.g., Wi-Fi, devices, quiet spaces), systemic support (e.g., policies, schedules, incentives, collaboration), and leadership (e.g., a clear vision for digital language learning). The CIPP evaluation conducted in this study examines these di-mensions in context, input, and process terms, drawing conceptually on this systems view of ICT integration.

2.5 CIPP Evaluation Model and Quasi-Experimental Approach

Program evaluation models offer another layer of the-oretical grounding for this study. The Context–Input–Process–Product (CIPP) model, originally formulated by Stufflebeam, frames evaluation as a decision-oriented activity: context evaluation identifies needs and oppor-tunities; input evaluation assesses strategies, resources, and action plans; process evaluation monitors imple-mentation; and product evaluation examines outcomes, including intended and unintended effects. The model is well-suited to educational technology innovations be-cause it emphasizes not only impact but also feasibility, relevance, and sustainability.

Within the present study, context evaluation focuses on vocational students' and alumni's needs for English pronunciation and broader communicative competence, particularly in light of global workplace demands. Input evaluation considers the readiness of school manage-ment, teachers, and students, as well as the adequacy of technical and human resources for using ELSA Speak. Process evaluation captures how the app is implemented in teaching, what challenges arise (e.g., connectivity, in-terface language, phonetic complexity), and how teach-ers mediate app use. Product evaluation examines quan-titative outcomes in pronunciation gains and engage-ment, as well as qualitative perceptions of usefulness.

From a research design perspective, the evaluation is embedded in a quasi-experimental framework. In educa-tional settings, random assignment is often impractical; therefore, quasi-experimental designs, such as non-equivalent control group designs with pre-tests and post-tests, are frequently used to estimate intervention effects while acknowledging threats to internal validity (Creswell & Creswell, 2018; Sugiyono, 2013). Creswell describes experiments as designs in which an inde-pendent variable is manipulated and its effect on a de-pendent variable is measured, ideally controlling for al-ternative explanations through randomization, compar-ison groups, and statistical control. Sugiyono similarly distinguishes experimental from non-experimental de-signs and notes that in applied educational research, quasi-experimental approaches are commonly used when intact classes must be retained for pragmatic or ethical reasons.

In this study, the independent variable is the use of ELSA Speak integrated into pronunciation instruction for the experimental class, while the control class receives conventional pronunciation teaching without the app. The primary dependent variable is pronunciation per-formance, measured through pre-test and post-test tasks aligned with communicative, vocationally relevant out-comes. Combining the CIPP framework with the qua-si-experimental design allows the study to address not only whether ELSA Speak improves pronunciation but also under what contextual and process conditions it is most effective.

2.6 Conceptual Framework of the Present Study

Synthesizing the strands outlined above, the conceptu-al framework for this study can be articulated as follows. First, the study adopts the intelligibility principle as the guiding phonological goal: ELSA Speak is expected to help vocational students produce key segmental and su-prasegmental features of English technical vocabulary more intelligibly, rather than to approximate native-like accents. Second, pronunciation is conceptualized as an integral component of vocational communicative com-petence, central to safe, efficient, and professional work-place communication. Third, ELSA Speak is framed as a MALL/CAPT tool that, when effectively integrated into instruction, can provide additional, individualized prac-tice, facilitate noticing of phonological gaps, and support self-regulated learning, thereby complementing teach-er-led pronunciation work.

Fourth, the adoption and sustainability of ELSA Speak in a vocational high school are understood through a whole-school ICT integration lens: teacher ICT attributes, school ICT capacity, school system capacity, and leader-ship interact to create necessary, sufficient, and sustain-able conditions for innovative pedagogical practices with technology. Within this lens, the CIPP model structures the evaluation of needs (context), resources and plans (input), implementation processes (process), and learn-ing outcomes (product). Fifth, the quasi-experimental design provides a methodological framework for esti-mating the effect of the app on pronunciation perfor-mance, recognizing the constraints of intact classes and real-world school settings.

Put in schematic terms, the framework posits that:

1. Contextual factors (students' vocational English needs, existing pronunciation difficulties, curricular demands) create a need for targeted pronunciation support.

2. Given adequate inputs (devices, connec-tivity, teacher training, guidelines, institutional support), ELSA Speak can be implemented as a pronunciation intervention.

3. During implementation, teacher media-tion, learner engagement, and the resolution of technical and pedagogical challenges shape the quality and intensity of exposure to the app.

4. Effective implementation leads to im-proved pronunciation outcomes (greater intelli-gibility of target technical lexis, better control of word stress and intonation) and increased learner engagement with pronunciation prac-tice.

5. Feedback loops from product back to context and input (e.g., positive outcomes and perceptions informing policy and resource allo-cation) determine whether ELSA Speak becomes a sustained component of pronunciation peda-gogy or remains a short-lived innovation.

This theoretical framework thus provides the concep-tual justification for examining ELSA Speak as a technol-ogy-mediated intervention for vocational pronunciation learning and organizes the study's variables and hypoth-eses.

It accommodates both micro-level processes (how individual learners interact with the app and adjust their pronunciation) and meso-level school conditions that enable or constrain such interactions, aligning with con-temporary perspectives on language learning in digitally rich educational ecosystems.

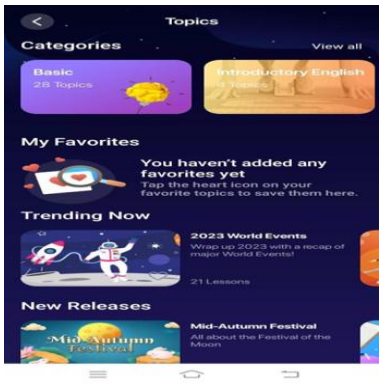

Figure 2.1 ELSA Speak App main screen view.

3. METHOD

Research Design

a. Research Design

The study belongs to quasi-experimental. A quasi-experimental, according to Safir (2022), consists of multiple categories for the control and experimental groups, but the groups weren't chosen at random. The group categories were put together organically before the study. Pre- and post-tests were provided to both the experimental and control groups. According to Isnawan (2020), non-equivalent group designs for the pre- and post-tests are frequently used in classroom experiments where the experimental and control groups are spontaneously built as complete classes that may be comparable. Because of this, the researchers had two classes: the experimental and control groups.

Group	Pre-test	Treatment	Post-test
Experimental Class X TOC	O_1	X	O_2
Control Class X TOB	O_3	-	O_4

Adopted from Hatch and Farhady (1982) cited on Basopi (2021)

b. The Population and Sample

1. Population

Pandey (2021) stated that population is a generalization area consisting of objects or subjects by the researcher to be studied and then to get conclusions. Thus, the population is not only people but also objects and other natural objects. The population is also not just the amount that exists on the object or subject studied but includes all the characteristics possessed by the subject or object. Moreover, Syafrida (2022) mentioned that in educational research, the population is usually a group of people (pupils, teachers, or other individuals) who possess certain characteristics and, in some cases, it can be defined as a group of classrooms, schools, or even facilities. Considering that reason, the population of this study was the entire tenth-grade students of the Vocational High School Muhammadiyah Prambanan. They enrolled in the academic year 2024/2025 (total population 10 classes x 38 of students).

The first is the researcher, the second is a teacher who has been teaching for 10 years, is used to teaching using technology, the third is a senior teacher who has been teaching for almost 30 years, and the last is a mutation teacher who has only been teaching at SMK Muhammadiyah Prambanan for a few months.

2. Sample

From 10 classes available, the researcher decided two classes that were taught by one teacher who was selected using purposive sampling (1.the teacher who teaches class X, 2. He/she uses technology in teaching, and 3. She is a certified teacher). Class X Automotive

Engineering C (TOC) consists of 38 students and X Automotive Engineering B (TOB) also consists of 38 students. The X TOB was assigned as the control group, while X TOC was assigned as the experimental group. Both groups were given pre- and post-test. During the experimental process, the class received five treatments (each treatment lasted for 90 minutes). The topics of the lesson given to experimental and control groups are the same by using verb phrases in vocational procedure texts.

The results of the pretest and posttest are calculated by using SPSS 26, and their statistical results such as means, median, mode, and t-test are given. T-test is a statistical method used to test whether there is a significant difference between the means of two groups, either independent or paired (Larson, 2022). The most common types of t-test include independent t-test, which is used to compare the means of two unrelated groups, and paired t-test, which is used to compare the means of two related data sets, such as pre- and post-treatment data. In addition, there is also a one-sample t-test that compares the sample mean to a specific value. This method is usually used when the sample size is relatively small and the data is normally distributed, and when the variance between groups is considered equal (Field, 2022). T-test of control class and t-test of experimental class then are compared with the t table results. If the t-test experiment is higher than t-table, it means there is a significant difference between control and experimental classes or the Hypothesis 1 (H1) is accepted or Ho is rejected.

The use of N-gain serves several purposes. First, it helps in measuring the improvement in pronunciation ability, allowing researchers to objectively assess whether ELSA Speak contributes to better pronunciation skills. Second, it aids in evaluating the effectiveness of the app by comparing how much students' skills have improved relative to their initial levels. Finally, N-gain helps to determine the impact of the application on students' learning outcomes compared to traditional pronunciation methods (Haspen & Syafriani, 2022).

To answer what obstacles that teacher faces when teaching English pronunciation using ELSA Speak App, as well as solutions made by teachers and as these solutions are considered effective, the researcher used an interview instrument. Interview guide based on the CIPP model (Context, Input, Process, Product) (Magdalena, 2024). The interview started with a context evaluation, asking about the main objectives of the program, meeting the needs of the targets, stakeholder involvement, and external factors affecting the program. At the input stage, questions focus on the availability of resources, the planning process, decision-making, and the readiness of staff or participants. For the process, interviews review the implementation of the program, challenges faced, and the flexibility of the program in responding to obstacles (Rima, 2023). At the product stage, the evaluation emphasizes the outcomes of the program, the resulting impact, sustainability, as well as areas that require improvement based on the results achieved.

C. Data Collection

1. Research Instrument

Research instrument is something that is used together with information for a study. According to Creswell (1994) cited in Khoirunnisa (2023) an instrument is a measuring tool used to reveal specific details. In this research, the instruments are pre-test and post-test are adopted from ELSA Speak App.

The test will be conducted with a student voice test using the ELSA Speak App. The instrument used in the pretest post-test was a verb phrase taken from the automotive vocational procedural text. The following illustrates an example of pre-test and post-test material:

D. Research Procedure

The research procedure of this study is as follows:

1. Preparing the lesson plan

Lesson plan was created to be implemented during the treatment sessions. The materials included in designing the lesson plan (teaching module) are vocabularies that are in the learning outcomes of phase E, which is descriptive text.

The treatment will be conducted seven times. The first and last meeting were allocated to conduct pre-test & post-test, while the other five meetings were used to learn pronunciation using ELSA Speak App.

2. Conducting pre-test

The pre-test was given during the first meeting to both the experimental and control groups. The purpose of this exam is to collect baseline data on students' core listening skills and see whether listening skills were comparable in the two groups prior to treatment. Strunk (2023) states that a pre-test provides a measure of a trait or quality that researchers evaluate for study participants prior to the therapy being administered.

3. Conducting treatment

After conducting the pre-test, treatment was given to the experimental group. Researchers physically alter conditions in studies by interfering in one or more of them, causing participants to perceive something differently in the experimental condition than in the control condition (Creswell, 2012, p. 301).

Distinctive between the experimental and control conditions (Creswell, 2012, p. 301). The treatment in question was the application of the ELSA Speak App. The time allocation for each meeting was two lesson hours where one lesson hour is forty minutes. The time schedule of this study is summarized in Table 3.2.

4. Conducting post-test

The next step is to conduct a post-test which aims to measure students' English pronunciation skills after receiving treatment using the ELSA Speak App. This post-test was given to the experimental group and control group.

This post-test aims to measure the difference in student scores between the experimental and control groups. Mackiewic, (2018) reiterates that after treatment, "researchers can take another reading on an attribute or characteristic and a post-test is a measurement on some attribute or characteristic assessed for participants in an experiment after treatment" (p. 297).

4. RESULTS

The data collection period was started from 5 to 15 November 2024, and the sample consisted of students in class X Automotive Engineering C as the experimental class and X Automotive Engineering B as the control class. The collected data were analyzed to determine 1) evaluating the impact of ELSA Speak App on English Pronunciation Skills 2) the teacher obstacles when using ELSA Speak App in teaching English pronunciation, 3) teacher's solution to solve the problem when teaching English pronunciation using ELSA Speak App, and 4) the effectiveness of teacher-made solutions. The results of the data analysis are presented in the following sections.

Research question 1. How effective is ELSA Speak App in teaching English Pronunciation at X TOC Class of SMK Muhammadiyah Prambanan?

a. Normality Test

Tests of Normality

	Kolmogorov-Smirnova			Shapiro Wilk		
	Statistics	df	Sig.(p)	Statistics	df	S
Pre Test Experiment	0.096	37	,200*	0.969	37	
Post Test Experiment	0.119	37	0.189	0.936	37	
Pre Test Control	0.159	37	0.016	0.905	37	
Post Test Control	0.127	37	0.129	0.933	37	

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure 1. Normality Test

The results of the normality test using Kolmogorov-Smirnov and Shapiro-Wilk showed that most of the data were normally distributed. In the experimental group, both the pre-test (p = 0.200) and post-test (p = 0.189) showed significant values greater than 0.05, indicating normality. However, in the experimental post-test and control pre-test, the p values were 0.032 and 0.004, respectively, indicating that the data were not normally distributed. However, since the majority of the data showed a normal distribution, it can be concluded that further analysis using parametric tests can be applied to these data.

b. Homogeneity Test

Test of Homogeneity of Variances

		Levene Statistics	df1	df2	Sig.
English pronunciation	Based on Mean	0.042	1	74	0.
	Based on Median	0.042	1	74	0.
	Based on Median and with adjusted df	0.042	1	73,820	0.
	Based on trimmed mean	0.040	1	74	0.

Figure 3. Homogeneity Test

The results of the homogeneity of variance test indicate that the variance between the experimental and control groups is homogeneous. All significance values, both based on the mean (0.839), median (0.838), median with adjusted degrees of freedom (0.838), and trimmed mean (0.843), are greater than 0.05. Therefore, it can be concluded that the variance between groups is homogeneous, which supports the use of parametric tests in further analysis.

c. Paired Sample Test

1) Experimental Class

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. E Mea
Pair 1	Pre Test Experiment	342.61	38	70,071	11
	Post Test Experiment	502.76	38	104,575	16

Figure 4. Paired Sample Test

In the Paired Sample Test for the experimental class, the data showed a significant difference between the pre-test and post-test scores. Before the treatment, the average pre-test score was 342.61 with a standard deviation of 70.071, while after the treatment (post-test), the average score increased to 502.76 with a standard deviation of 104.575. This shows that there is a significant increase in the post-test score compared to the pre-test score in the experimental class, with an average increase of 160.158.

The results of the Paired Sample Test in the experimental class revealed that the treatment provided significantly improved students' abilities, as seen from the striking difference between the pre- and post-treatment scores. This improvement indicates that the applied learning method is not only relevant but also effective in helping students achieve better learning outcomes. The higher standard deviation in the post-test indicates individual variation in response to the treatment, which may be influenced by factors such as initial ability level, learning speed, or level of student participation. Overall, the findings provide empirical evidence that the treatment could be a viable approach to be applied more widely in the learning process.

Paired Samples Test

Paired Samples Test

		Paired Differences					t	df	Si ta
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pre Test Experiment - Post Test Experiment	-160,158	125,212	20,312	-201,314	-119,002	-7,885	37	

Figure 5. Paired Sample Test 2

The results of the Paired Samples Test further confirmed a significant difference between the pre-test and post-test, with a p value of 0.000 (less than 0.05). The experimental t value was 7.885 and the t table was 1.68595, which means that the experimental t or t test was > than the t table. Which means that the results of the t-test show that the experimental t value (t test) is higher than the t table, meaning that the test results show statistically significant evidence. Thus, supporting the hypothesis that the treatment in the experimental class has a positive effect on improving student scores. Thus, based on the results of the t test and the confidence interval obtained, the alternative hypothesis (H_1) is accepted, which means that the use of the ELSA Speak App has been proven to improve students' English pronunciation. In other words, H0 is rejected.

2) Control Class

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre Test Control	383.68	38	78,917	12,802
	Post Test Control	403.89	38	106,475	17,273

Figure 6. Control Class

In the Paired Sample Test for the control class, the results showed a small change between the pre-test and post-test scores. The average pre-test score was 383.68 with a standard deviation of 78.917, while the post-test score increased slightly to 403.89 with a standard deviation of 106.475. Despite the increase in scores, the average difference was only 20.211, indicating a relatively small change compared to the experimental class.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pre Test Control - Post Test Control	-20,211	124,714	20,231	-61,203	20,782	-0.999	37	0.324

Figure 7. Paired Samples Test 3

The results of the Paired Samples Test showed that this difference was not statistically significant, with a p-value of 0.324 (greater than 0.05). The t-value of -0.999 with 37 degrees of freedom indicates that the difference between the pre-test and post-test was not strong enough to be considered a significant change. The 95% confidence interval for the difference in pre-test and post-test scores ranges from -61.203 to 20.782, which includes the value of zero, indicating that the change could have occurred by chance and is not sufficient to support a meaningful difference of 1.684. Means in the control class there is no significant difference.

d. independent t Test

Group Statistics

Group		N	Mean	Std. Deviation	Std. Error Mean
English pronunciation	Experiment	38	160.16	125,212	20,312
	Control	38	20,21	124,714	20,231

Figure 8. independent t Test

In the Independent t-test between the experimental and control groups for English pronunciation testing, the results showed a significant difference. The experimental group had means of 160.16 with a standard deviation of 125.212, while the control group only recorded means of 20.21 with a standard deviation of 124.714. With these results, there is a clear difference between the two groups, where the experimental group showed higher results.

Independent Samples Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
English pronunciation	Equal variances assumed	0.042	0.839	4,882	74	0,000	139,947	28,669	82,824	197,071
	Equal variances not assumed			4,882	73,999	0,000	139,947	28,669	82,824	197,071

Figure 9. independent t Test

The t-test results showed that the difference between the experimental and control groups was highly statistically significant, with a p-value of 0.000, which is smaller than 0.05. The t-value obtained was 4.882 with 74 degrees of freedom (or 73.999 if the variances are not assumed to be the same) confirming a significant difference between the two groups. In addition, the 95% confidence interval for the mean score difference between the experimental and control groups ranged from 82.824 to 197.071. This supports the finding that the use of a particular method in the experimental group had a greater impact on improving English pronunciation compared to the control group. Thus, based on the results of the t-test and the confidence interval obtained, the alternative hypothesis (H_1) is accepted, which means that the use of the ELSA Speak App is proven to improve students' English pronunciation.

e. N gain scores

Table 1. N gain scores

Pre Ex	Criteria	Post Ex	Criteria	Pre-Control	Criteria	Post Control	Criteria
27.8	Ineffective	41.9	Not enough	35.8	Ineffective	33	Ineffective
28.8	Ineffective	36.1	Ineffective	53	Not enough	37.7	Ineffective
27.8	Ineffective	48.1	Not enough	31.5	Ineffective	43.9	Not enough
40.5	Not enough	45.7	Not enough	46.2	Not enough	47.3	Not enough
34.6	Ineffective	36.1	Ineffective	30.2	Ineffective	46.7	Not enough
24.4	Ineffective	44.1	Not enough	44.3	Not enough	39.9	Ineffective
40.8	Not enough	55.7	Not enough	44.6	Not enough	33.1	Ineffective
33.7	Ineffective	73.1	Enough	37.3	Ineffective	32.6	Ineffective
37.8	Ineffective	64.3	Enough	32.5	Ineffective	36.5	Ineffective
37.5	Ineffective	52.8	Not enough	29.1	Ineffective	24.5	Ineffective
36.3	Ineffective	51	Not enough	30.2	Ineffective	41.6	Not enough
40.8	Not enough	58	Enough	31.7	Ineffective	44.9	Not enough
33.7	Ineffective	40.4	Not enough	31.7	Ineffective	61.3	Enough
34.8	Ineffective	60.7	Enough	54.5	Not enough	57.4	Enough
44.9	Not enough	48.3	Not enough	32.8	Ineffective	34.4	Ineffective
36	Ineffective	40	Not enough	39	Ineffective	54.5	Not enough
34.9	Ineffective	53.4	Not enough	56.9	Enough	32.7	Ineffective
36	Ineffective	42.9	Not enough	35.8	Ineffective	60.4	Enough
35.9	Ineffective	38.8	Ineffective	53	Not enough	35.1	Ineffective
33	Ineffective	76	Effective	31.5	Ineffective	22.3	Ineffective
47.1	Not enough	55.2	Not enough	46.2	Not enough	46.2	Not enough
52.8	Not enough	49.1	Not enough	28.6	Ineffective	25.3	Ineffective
35.7	Ineffective	42.5	Not enough	49.3	Not enough	45.6	Not enough
26.6	Ineffective	52.3	Not enough	34.4	Ineffective	34.2	Ineffective
22	Ineffective	60.9	Enough	39.8	Ineffective	42.5	Not enough
30.8	Ineffective	52	Not enough	32.9	Ineffective	36.4	Ineffective
30.5	Ineffective	36.1	Ineffective	28.3	Ineffective	40.3	Not enough
27.3	Ineffective	77	Effective	36.2	Ineffective	38.7	Ineffective
27.1	Ineffective	48.4	Not enough	37	Ineffective	29.7	Ineffective
47.9	Not enough	50.2	Not enough	44.7	Not enough	41.8	Not enough
23.1	Ineffective	58.2	Enough	48.1	Not enough	35.9	Ineffective
36.8	Ineffective	49.4	Not enough	33.9	Ineffective	41.4	Not enough
42.7	Not enough	51.1	Not enough	33.4	Ineffective	34.7	Ineffective
33.1	Ineffective	47.2	Not enough	36.5	Ineffective	73.8	Enough
29.5	Ineffective	33.6	Ineffective	40.1	Not enough	39.3	Ineffective
31.3	Ineffective	48.5	Not enough	32.9	Ineffective	28.4	Ineffective
30.1	Ineffective	46.7	Not enough	38.3	Ineffective	34.4	Ineffective
27.5	Ineffective	44.7	Not enough	35.8	Ineffective	46.4	Not enough

Experimental Class - Pre-Test, in the pre-test of the experimental class, most participants showed results in the "Ineffective" category with scores varying from 24.4 to 47.9. This indicates that many students do not have adequate mastery of English pronunciation before the treatment was carried out.

Experimental Class - Post Test, after the treatment, in the post-test of the experimental class, there were a number of participants who showed significant improvements in pronunciation, with higher scores, such as 73.1 and 77 which were in the "Enough" or "Effective" category. However, some participants were still in the "Poor" category, with scores ranging from 36.1 to 60.7, indicating that despite the improvements, students' English pronunciation was not yet fully optimal.

Control Class - Pre-Test, in the pre -test of the control class, most students were also in the "Ineffective" category, with scores ranging from 22 to 56.9. This shows that before the treatment, the level of English pronunciation ability of students in the control group was still relatively low, almost similar to the results in the experimental class.

Control Class - Post Test, the post-test results of the control class showed a slight increase, with scores in the "Poor" and "Fair" categories, such as 42.5 and 60.4. However, although there were some participants who achieved higher scores, most were still at low scores, between 22.3 to 45.6, indicating that the treatment received by the control group did not provide significant changes in English pronunciation ability.

Research Question 2. What are the obstacles encountered by the teacher when using ELSA Speak app in teaching English pronunciation at X TOC of SMK Muhammadiyah Prambanan?

Based on the interview using CIPP (Context, Input, Process, Product) Models, the obstacles encountered by the teacher when using ELSA Speak app in teaching English pronunciation at X

TOC of SMK Muhammadiyah Prambanan identifies several key aspects:

a. Context Evaluation

In the context evaluation, the program aims to improve students' English pronunciation skills by integrating the ELSA Speak app into classroom activities. This initiative addresses students' pronunciation challenges by leveraging technology, with teachers actively supporting the process. However, external factors, such as limited internet connectivity, insufficient smartphone storage, and classroom noise, significantly impact the program's effectiveness.

b. Input Evaluation

The input evaluation reveals that while the ELSA Speak app is available as a resource, students face technical barriers, including internet and storage limitations. Teachers have taken steps to address these issues by assisting students during the installation and login process and providing clear guidance to navigate the app, which uses English instructions. Teachers are ready to support students, but the latter require additional adaptation and resources to use the app effectively.

c. Process Evaluation

In the process evaluation, students began using the ELSA Speak app for pronunciation practice but encountered significant challenges. Technical issues, such as difficulties in installation and usage due to storage and network constraints, affected students' engagement. Additionally, the app's feedback system often provided inaccurate assessments, with pronunciation scores falling below expectations, leading to confusion. The noisy classroom environment further disrupted learning by affecting the app's feedback accuracy and creating distractions.

d. Product Evaluation

The product evaluation highlights that the integration of the ELSA Speak app has potential but is hindered by technical limitations, feedback inconsistencies, and classroom atmosphere challenges. These obstacles underline the need for targeted solutions, such as better infrastructure, supplemental teacher guidance, and strategies to create a more conducive learning environment, to maximize the program's effectiveness in improving students' pronunciation skills.

Research question 3. What are the solutions made by the teacher to solve the problems related to the use of ELSA Speak App in teaching English Pronunciation at X TOC Class of SMK Muhammadiyah Prambanan?

From an interview conducted by two teachers from 4th to 12th November 2024, based on the CIPP (Context, Input, Process, Product) Models, the solutions implemented by teachers to address the challenges of using the ELSA Speak app for teaching English pronunciation at X TOC Class of SMK Muhammadiyah Prambanan are evaluated as follows:

a. Context Evaluation

In the context evaluation, the program's main objective is to enhance students' English pronunciation skills through the ELSA Speak app. The identified challenges include limited internet connectivity, insufficient phone storage, and classroom noise, which hinder the effective use of the app. To meet these needs, teachers focus on providing both technical and instructional support to ensure that the program's objectives align with students' learning requirements.

b. Input Evaluation

The input evaluation highlights the availability of resources and the strategic planning involved in overcoming technical barriers. Teachers direct students to use Wi-Fi facilities, such as those in the school library, to address internet connectivity issues. To resolve storage limitations, students are encouraged to free up space by deleting unnecessary files or using external storage devices. Teachers also actively assist students during the installation and login processes, ensuring smooth access to the app.

c. Process Evaluation

In the process evaluation, the implementation of solutions involves ensuring students practice English pronunciation regularly using the app.

Teachers focus on enhancing students' engagement and performance by encouraging consistent use of the app despite its built-in feedback limitations. To address the noisy classroom environment, teachers employ strategies such as dividing students into smaller groups, rotating practice sessions, and promoting the use of headphones. These measures aim to minimize distractions, maintain focus, and improve the app's feedback accuracy.

d. Product Evaluation

The product evaluation reflects the effectiveness of these solutions in creating a more conducive learning environment. By addressing technical challenges and fostering consistent practice, teachers successfully support students in overcoming obstacles and improving their pronunciation skills. The division of practice sessions and use of headphones also contribute to a quieter classroom atmosphere, ensuring better engagement and learning outcomes. These solutions demonstrate the teachers' commitment to optimizing the integration of ELSA Speak into their teaching practices.

Research question 4. How do the teachers perceive the effectiveness of the solutions made by the teacher?

Based on the CIPP (Context, Input, Process, Product) Models, the teachers' perception of the effectiveness of the solutions implemented to address challenges in using the ELSA Speak app highlights significant progress in achieving the program's goals.

a. Context Evaluation

In the context evaluation, the program's main objective is to improve students' English pronunciation skills by leveraging the ELSA Speak app. The solutions were designed to address the technical and pedagogical challenges faced by students, ensuring that their learning needs were met. The teachers' active involvement and flexibility played a crucial role in adapting the program to external factors, such as connectivity issues and classroom dynamics.

b. Input Evaluation

The input evaluation focuses on the resources and strategies employed by the teachers to overcome technical obstacles. Teachers provided direct guidance during installation and troubleshooting processes, helping students navigate the app's English-based instructions. To address storage and connectivity issues, they adapted learning strategies, such as recommending Wi-Fi-equipped locations or alternative methods, ensuring uninterrupted access to learning materials.

c. Process Evaluation

The process evaluation, the teachers' hands-on approach and supplementary interventions proved effective in addressing the app's feedback limitations. Teachers bridged the gap between automated and real-time feedback by providing personalized corrections and additional practice opportunities. Their efforts ensured that students remained consistent in their practice, leading to notable improvements in pronunciation skills. Strategies to manage classroom noise, such as dividing practice sessions, promoting the use of headphones, and facilitating quieter settings, created a conducive environment for focused learning.

d. Product Evaluation

Finally, the product evaluation reflects the success of these solutions in achieving the program's objectives. Students demonstrated measurable progress in pronunciation skills, as evidenced by improved post-test scores. The tailored support provided by teachers not only enhanced the effectiveness of the app but also fostered a positive learning atmosphere, sustaining students' motivation and engagement. Overall, the teachers perceived their solutions as highly effective in addressing challenges and reinforcing the program's outcomes.

From the interview finding the researcher concluded as given in the table below:

Aspect	Challenges Identified	Solutions Implemented	Effectiveness Perceiv
Context Evaluation	- Limited internet connectivity and phone storage - Classroom noise disrupting learning- Students' unfamiliarity with the app's English-based instructions	- Encouraged students to use Wi-Fi facilities, such as school libraries - Provided direct guidance for app installation and login- Focused on balancing technical and instructional support	- Teachers actively adapted to external challenges- Program goals aligned with students' learning nee - Flexibility in addressing issues improved student engagement
Input Evaluation	- Technical barriers: installation, storage, and network constraints - Limited readiness of students to navigate app features	- Directed students to free up storage space or use external devices - Assisted students during installation and setup - Adapted teaching strategies to address technical limitations	- Teachers' guidance ensured smoother app access - Strategies like recommending Wi-Fi and freeing storage helped overcome connectivity and stora issues - Enhanced resource utilization supported consistent learning
Process Evaluation	- Challenges in app usage due to feedback inconsistencies - Classroom noise affecting focus and app feedback accuracy	- Promoted regular use of the app despite feedback limitations - Divided students into smaller groups or rotated practice sessions - Encouraged headphone use for noise reduction and focused practice	- Supplementary teacl feedback improved learning outcomes - Noise reduction strategies enhanced classroom focus - Students showed consistent practice an engagement, overcoming app feedback limitations
Product Evaluation	- Potential of *ELSA Speak* hindered by technical and classroom challenges - Need for a conducive environment to optimize learning	- Created structured and quiet practice environments - Supplemented app feedback with real-time teacher interventions- Focused on maintaining student motivation through consistent support	- Improved pronunciation skills a evidenced by post-tes scores - Quieter and more focused classroom atmosphere supported learning - Teachers perceived solutions as highly effective in reinforcin app integration and

Table 2. Summary: Challenges, Solutions, and Effectiveness of Using the ELSA Speak App in X TOC Class at SMK Muhammadiyah Prambanan

5. Discussion

The results answered the first research question that ELSA Speak App is effective in teaching English pronunciation in class X Engineering Technology C (TOC) at SMK Muhammadiyah Prambanan. This was proven by the increase in post test scores. These results support several previous studies, such as:

Rineapi (2022) asserted that the ELSA Speak app is perfect for learning pronunciation, offering high-quality content covering vowels, consonants, diphthongs, syllables and word stress. This makes the app a useful resource for pronunciation learning.

Kholis (2021) supports this statement by mentioning that this app effectively improves students' pronunciation skills. The t-test analysis showed a significant difference between the pre-test and post-test results, which proved a substantial improvement in pronunciation after the use of ELSA Speak. The recording and instant feedback features help students correct mistakes quickly.

Anggraini (2022) research also highlighted the efficiency of this app in facilitating pronunciation learning. Students not only understand the correct pronunciation faster, but also become more confident and motivated in speaking English.

Hasma (2022) concluded that despite some drawbacks, the ELSA Speak app has a positive impact on pronunciation learning. The app not only improves students' pronunciation scores, but also provides features that allow teachers to easily integrate it into their teaching strategies.

However, in addition to the effectiveness of using the ELSA Speak App, there are also shortcomings and obstacles that can be obtained from this research. As has been revealed above from the results of interviews with teachers who apply English pronunciation teaching such as: limited technology and internet quota, difficulty in the installation and login process, inaccurate application feedback, difficulty for students to imitate pronunciation, and class noise interferes with practice. These findings are supported by research from Aswaty and Indari (2022) noted several weaknesses of this application. One of them is the dependence on technological devices such as smartphones, which can be a barrier for students who do not have access. In addition, the app relies on a stable internet connection, and most of its content is available in a paid version. They also found that the effectiveness of ELSA Speak is more limited to certain accents, so some students may not benefit to the maximum extent.

.

5.CONCLUSION

The findings of this study reveal that the ELSA Speak app is effective in enhancing English pronunciation skills among students of Class X Automotive Engineering C at SMK Muhammadiyah Prambanan. This conclusion is supported by statistical evidence from pre-test and post-test analyses in the experimental and control groups.

This study finds that the use of the ELSA Speak application is significantly effective in improving the English pronunciation skills of students in Class X Automotive Engineering C (TOC) at SMK Muhammadiyah Prambanan. This is evidenced by the t-test results, which show a t-value of 4.882 with a p-value of 0.000 ($p < 0.05$), indicating a statistically significant difference between the experimental group using ELSA Speak and the control group. The average improvement in scores for the experimental group was 160.16, which is significantly higher than the control group's average improvement of 20.21.

However, the study also identified several challenges during the implementation, such as limited internet connectivity, insufficient device storage, difficulties in the installation and login process, and inaccuracies in the application's feedback. Furthermore, classroom noise emerged as an external factor affecting the accuracy of the application's feedback and the overall effectiveness of the learning process.

To address these challenges, teachers implemented various strategies, including providing access to school Wi-Fi, assisting students with installation and navigation, dividing students into smaller groups to minimize noise, and encouraging the use of headphones to improve focus during practice. Teachers also provided direct guidance to complement the application's automated feedback, enabling students to better understand their mistakes and improve their pronunciation effectively.

Program evaluation using the CIPP (Context, Input, Process, Product) model revealed that integrating the ELSA Speak app into learning activities significantly enhanced students' motivation and led to substantial improvements in their pronunciation skills. However, some students required more time and intensive guidance to achieve satisfactory results.

In conclusion, the ELSA Speak application has proven to be an effective tool for teaching English pronunciation when accompanied by comprehensive teaching strategies and adequate technical support. This study highlights the importance of addressing technical and environmental factors to optimize the integration of technology into the learning process and achieve better outcomes.

Acknowledgments

The researcher would like to deliver their great appreciations to the school principal for allowing us to conduct the research at the SMK Prambanan. Our sincere thanks also go to the English teachers for their technical assistances before and during the research.

Author Contributions

Author's full name1: Akmal as the Conceptor and supervisor

Author's full name2: Bunga Setya as Data curator,

Author's full name3: Intan Suzila as the Formal Analysis,

Funding

The study was conducted without any supporting grants or it was conducted based on the researchers' funds or This work is not supported by any external funding.

Data Availability Statement

The data is available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

REFERENCES

Akhmad, N. W., & Munawir, A. (2022). Improving the Students' Pronunciation Ability by Using ELSA Speak App. IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, 10(1), 846–857. https://doi.org/10.24256/ideas.v10i1.2868

Akramov, X., Davlyatov, S., Umarov, S., & Abobakirova, Z. (2023). Method of experimental research of concrete beams with fiberglass reinforcement for

bending. E3S Web of Conferences, 365, 1–5. https://doi.org/10.1051/e3sconf/202336502021

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

Arifin, A., Norain, S., & Ridwan, R. (2022). Esp Course Design: the Need Analysis on English for Tourism Book for Travel Business Department of Eleventh Grade Students At Smkn 1 Tarakan. Technium Education and Humanities, 2(2), 1–7. https://doi.org/10.47577/teh.v2i2.6968

Aswaty, P., & Indari, A. (2022). the Effect of Using Elsa (English Language Speech Assistant) Speak Application on Students' Speaking Ability for the Eleventh Grade of Mas Darul Al Muhajirin in the Academic Year 2021/2022. Serunai : Jurnal Ilmiah Ilmu Pendidikan, 8(1), 18–23. https://doi.org/10.37755/sjip.v8i1.616

Basopi, P., Aziz, R. A., & Bahasa, A. P. (2021). EFFECTIVENESS OF USING LISTENING MEDIA SOFTWARE (LMS) IN TEACHING LISTENING IN EFL IN. 9(1), 115–128.

Dehghanpour, B., Beheshti, Z., & Jalali, M. (2022). Role of Flipped, Blended, and Conventional Learning Environments on Developing Pronunciation of EFL Learners Using ELSA Apps. Culture, and Translation (LCT), 5(1), 150–161. https://doi.org/10.30495/LCT.2023.1972987.1073

Elguanda, A. suci, Siska, S. S, & Rika AFRIYANTI. (2023). Students' Behavioral Engagement in Learning English in Vocational High School. Journal of Digital Learning and Distance Education, 2(3), 525–533. https://doi.org/10.56778/jdlde.v2i3.85

Fakdawer Maria Rosa, Y. (2020). ELSA Apps to Encourage Learner Autonomy.

Field, A. 2018. Discovering Statistics Using IBM SPSS Statistics 5th Edition. 5 ed. California: SAGE Publications Inc.

Ghozali, I. 2021. Aplikasi Analisis Multivariate dengan Program IBM SPSS. 26. 10 ed. Semarang: Badan Peneribit Universitas Diponegoro.

Hasma, N. A., Latief, H., & Mahmudah, D. R. M. (2022). Influence of the Optimal Parents' Involvement To Students Achievement in English Learning. Journal of Language Testing and Assessment, 2(2), 88–103. https://doi.org/10.56983/jlta.v2i2.153

Haspen, C. D., & Syafriani, S. (2022). Praktikalitas dan Efektifitas Emodul Fisika Berbasis inkuiri Terbimbing Terintegrasi Etnosains Untuk Meningkatkan Kemampuan Berpikir Kreatif Peserta Didik. Jurnal Penelitian Pembelajaran Fisika, 8(1), 10. https://doi.org/10.24036/jppf.v8i1.115684

Ina Magdalena, Deis Ayu Nur Hidayah, D. K. A. (2024). Cendikia pendidikan. Cendekia Pendidikan, 3(1), 1–13. https://ejournal.warunayama.org/index.php/sindorocendikiapendidikan/article/view/769

Indari, A. (2023). Detection Of Pronunciation Errors In English Speaking Skills Based On Artificial Intelligence (AI). Jurnal Serunai Bahasa Inggris, 15(2).

Kenworthy, J. (1987). Teaching English Pronunciaton (p. 163).

Khoirunnisa, K., Candraloka, O. R., & Shofiyuddin, M. (2023). Effectiveness of Using Educandy in ELT to Improve Students' Vocabulary Mastery. Darussalam English Journal (DEJ), 3(1), 67–92. https://doi.org/10.30739/dej.v3i1.2304

Kholis, A. (2021). ELSA Speak App: Automatic Speech Recognition (ASR) for Supplementing English Pronunciation Skills. Pedagogy : Journal of English Language Teaching, 9(1), 01. https://doi.org/10.32332/joelt.v9i1.2723

Marlinda, S., & Huda, N. N. (2024). IMPROVING PRONOUNCIATION SKILLS WITH ELSA SPEAK AN AI- BASED LEARNING. 1866–1871.

Meliawati, M., & Amied, A. F. (2020). Analisis Konten pada Buku Teks Bahasa Inggris untuk Kelas X Berjudul Bahasa Inggris untuk SMA/MA/SMK/MAK Kelas X Content Analysis on English Textbooks

for Class X Title English for High School. Jurnal Penelitian Pendidikan, 20(1), 83–90.

Mackiewicz, J. (2018). A Mixed-Method Approach. In Writing Center Talk over Time. https://doi.org/10.4324/9780429469237-3

Mahmudah, R., & Daulay, S. H. (2024). ELSA Speak Application as an Innovation to Improve English Pronunciations of Speaking Skills: Student's Perception Beyond Boundaries: Innovations in Arabic and English Language Teaching for the Future ELSA Speak Application as an Innovation to Improve Eng. ANCOLT International Proseeding on Language Teaching, 1(1), 1–12.

Mekkah, U. S., & Aceh, B. (n.d.). The Use of ELSA Application in Teaching Pronunciation 1 1*,2 Nyak Mila*; 2 Nora Fitria. 29–36.

Pangastuti, D. (2021). The Effect of 'ELSA Speak'Application on Students'Pronunciation in English. Prosiding Pekan Ilmiah Mahasiswa Unis, 1(1), 127–133. http://ejournal.unis.ac.id/index.php/PKIM/article/view/1976%0Ahttp://ejournal.unis.ac.id/index.php/PKIM/article/download/1976/1244

Permatasari, S. (2024). ENHANCING PRONUNCIATION SKILLS IN EFL STUDENTS THROUGH THE ELSA SPEAK APPLICATION. 10(2), 205–212.

Pramudita, M. (2023). Vowel Mispronunciations in English: Case Studies of Indonesian Learners. Indonesian Journal of English Language Teaching, 8(1), 20-34.

Pratiwi, A. (2020). Phonological Interference and Its Impact on EFL Pronunciation: A Case Study of Indonesian Learners. Journal of English Education and Linguistics, 5(2), 115-128.

Rahmawati, S. (2021). Phonetic Challenges in English Pronunciation for Indonesian EFL Learners. Journal of Language and Linguistics, 10(3), 45-58.

Rima, R., & Anriani, N. (2023). Implementasi Model Evaluasi Context, Input, Process, and Product (CIPP) pada Program Pengenalan Lapangan Persekolahan (PLP) bagi Calon Guru Bahasa Inggris. Afeksi: Jurnal Penelitian Dan Evaluasi Pendidikan, 4(6), 622–631. https://doi.org/10.35672/afeksi.v4i6.175

Rineapi, Triwardani Henni, & Nur Raysal. (2022). The effectiveness of ELSA Speak application to improve pronunciation ability. Jurnal Fakultas Keguruan & Ilmu Pendidikan, 3(1), 28–33.

Rismawati, D., Suryana, Y., & Agustiana, V. (2022). The Effectiveness of ELSA Speaking Application in Improving English Pronunciation. The Proceedings of English Language Teaching, Literature, and Translation (ELTLT), 10(1), 178. https://proceeding.unnes.ac.id/index.php/eltlt/article/view/1329

Salman, S. A. (2021). Research Methodology. In Implementing Takaful in India. https://doi.org/10.1007/978-981-16-6281-2_5

Senowarsito, S., Susanto, A. S., & Ardini, S. N. (2024). Learning Pronunciation Using Record, Listen, Revise (RLR) Method in Dictionary Speech Assistant–ELSA Speak Application: How the Flow of Thinking Goes. Journal of Languages and Language Teaching, 12(2), 883. https://doi.org/10.33394/jollt.v12i2.11048

Sholekhah, M. F., & Fakhrurriana, R. (2023). The Use of ELSA Speak as a Mobile-Assisted Language Learning (MALL) towards EFL Students Pronunciation. JELITA: Journal of Education, Language Innovation, and Applied Linguistics, 2(2), 93–100. https://doi.org/10.37058/jelita.v2i2.7596

Sonia Febri, Y., & Hubbulwathan Duri, S. (2023). The Effect Of Using ELSA Speak Application To Improve Students' Pronunciation. Jolly Journal of English Education, 1(2), 54–63. http://ejournal.staihwduri.ac.id/index.php/jjee

Starkey-Perret, R. (2008). Gerald Kelly, How to teach pronunciation. In Les Cahiers de l'APLIUT (Issue Vol. XXVII N° 3, pp. 112–114). https://doi.org/10.4000/apliut.1350

Sugiyono 2022. Metode Penelitian dan Pengembangan (Research and Development). 5 ed. Bandung: Alfabeta.

Suffa Maghdalena, F., & Wahyuningsih, S. (2024). ELSA Speak App: An Alternative Media to Improve EFL Students' Speaking Skills in Indonesian Higher Education. 327–339. https://doi.org/10.24090/celti.2024.1037

Sukmawati, L. (2021). Common Pronunciation Errors of Indonesian EFL Learners. International Journal of Language and Linguistics, 7(1), 63-75.

Widagdi, A. dan O. C. R. (2022). Joyful Learning Journal. Unnes.Ac.Id, 11(4), 195–200. https://journal.unnes.ac.id/sju/index.php/jlj/article/view/23230

Yosintha, R., & Rekha, A. (2022). "ELSA Speak" in an Online Pronunciation Class: Students' Voices. IJEE (Indonesian Journal of English Education), 9(1), 160–175. https://doi.org/10.15408/ijee.v9i1.23033

Zhang, Y. (n.d.). The Research on English Verb Phrase Learning Study Skills about College Students. 2(3), 69–75. https://doi.org/10.25236/FER.033044