

USING OF NATURALREADER: A TEXT-TO-SPEECH TECHNOLOGY USED BY EFL STUDENTS IN LEARNING PRONUNCIATION

Salamah HK^{*1}, Indra Perdana¹, Akhmad Fauzan¹

¹*Magister of English Language Education, University of Palangka Raya, Indonesia*

^{*}*salamahkasim2@gmail.com*

Abstract

Pronunciation is a fundamental yet frequently neglected component of English as a Foreign Language (EFL) education, often overshadowed by grammar and vocabulary due to instructional time constraints. This study investigates the effectiveness and user experiences of NaturalReader, an AI-powered Text-to-Speech (TTS) application, as a learning medium to improve pronunciation skills among tenth-grade students at MA Miftahul Jannah Palangka Raya (Academic Year 2025/2026). Employing a mixed-methods approach, quantitative data were gathered through pre-tests and post-tests to measure phonetic accuracy, while qualitative insights were collected via classroom observations and semi-structured interviews to capture student perceptions. The findings reveal a significant improvement in students' command over segmental features (vowels, diphthongs, and consonants) and suprasegmental features (stress and intonation). Furthermore, the application successfully mitigated pronunciation-related anxiety by establishing an autonomous, low-stakes practice environment. This research underscores the strategic value of integrating mobile-friendly TTS technology to bridge the gap between regional phonetic interference and standard English pronunciation.

Keywords: NaturalReader, Pronunciation, Text to Speech

INTRODUCTION

English plays a dominant role as the global lingua franca, making speaking competence one of the essential skills that English as a Foreign Language (EFL) learners are expected to master. According to Harmer (2007), speaking is a productive language skill that enables learners to communicate ideas, opinions, and information effectively in various contexts. Among the components of speaking, pronunciation is fundamental because it directly affects speech intelligibility and successful communication. Kelly (2000) argues that inaccurate pronunciation may lead to misunderstandings even when learners possess sufficient grammatical and lexical knowledge. Likewise, O'Connor (1980) emphasizes that intelligible pronunciation is essential for effective oral communication. Pronunciation itself consists of both segmental features, including vowels, consonants, and diphthongs, and suprasegmental features such

as stress and intonation, all of which contribute to comprehensible speech (Hasan, 2014).

Despite its importance, pronunciation instruction often receives limited attention in EFL classrooms. Levy et al. (2011) reported that pronunciation tends to receive less emphasis than other language skills in many English language teaching contexts because teachers often prioritize grammar and vocabulary to meet curriculum demands. This condition is further supported by Tergujeff (2012), who found that pronunciation instruction is frequently constrained by limited instructional time and insufficient teaching resources. Furthermore, Indonesian learners commonly experience first-language (L1) interference because the phonological systems of Indonesian and regional languages differ considerably from English. Pranoto and Suprayogi (2019) explain that differences between learners' native language and English phonology often result in pronunciation

errors, particularly in the production of unfamiliar English sounds. At MA Miftahul Jannah Palangka Raya, where students come from diverse linguistic backgrounds such as Banjarese, Dayakese, and Javanese, these phonological differences present additional challenges in learning English pronunciation.

Recent developments in educational technology have introduced Artificial Intelligence (AI)-based applications that can support pronunciation learning. According to Fitria (2023), AI technologies have increasingly been integrated into English language teaching, offering innovative learning media that enhance students' engagement and independent learning. One of these technologies is Text-to-Speech (TTS), which converts written text into natural-sounding speech. As explained by Bakken et al. (2012), TTS technology provides learners with immediate audio models that can be replayed repeatedly, enabling them to practice pronunciation more independently. Unlike conventional audio recordings, TTS applications allow learners to generate pronunciation models from any digital text, adjust playback speed, and repeatedly listen to difficult words or sentences according to their learning needs. Such flexibility encourages autonomous learning while reducing learners' anxiety during pronunciation practice. Among various TTS applications, Victor (2014) describes NaturalReader as a versatile TTS platform capable of converting multiple document formats into natural-sounding speech, making it a promising learning medium for pronunciation practice.

Previous studies have demonstrated the positive contribution of TTS technology to English language learning. For example, Ekşi and Yeşilçınar (2016) found that online Text-to-Speech tools significantly improved EFL learners' pronunciation performance by providing consistent pronunciation models. Similarly, Bao (2018) reported that repeated exposure to TTS applications helped learners develop greater pronunciation automaticity during oral reading. In addition, Pratiwi and Ayu (2020) concluded that smartphone-based

pronunciation applications promoted autonomous learning by allowing students to practice English independently beyond the classroom. However, most previous studies have focused on general TTS applications or other pronunciation tools, while relatively limited research has specifically investigated NaturalReader in Indonesian EFL classrooms. Moreover, previous research has predominantly emphasized learning outcomes, with fewer studies simultaneously examining students' learning experiences through a mixed-methods approach.

Considering these gaps, this study investigates the implementation of NaturalReader as a pronunciation learning medium among tenth-grade students at MA Miftahul Jannah Palangka Raya. Using an explanatory sequential mixed-methods design, this research evaluates the effectiveness of NaturalReader in improving students' pronunciation performance while also exploring students' experiences during the learning process. The findings are expected to contribute to the growing body of literature on AI-assisted language learning and provide practical insights for integrating Text-to-Speech technology into English pronunciation instruction.

RESEARCH METHOD

Research Design

This study employs an explanatory sequential mixed-methods research design. This framework allows for a comprehensive exploration of the research problem by combining quantitative data with qualitative insights.

The quantitative phase uses a pre-experimental design featuring a single-group pre-test and post-test matrix to evaluate changes in phonetic accuracy after the technology intervention. Following the collection of quantitative data, the qualitative phase utilizes semi-structured interviews and structured classroom observations. This secondary phase provides depth to the statistical results by exploring the students' psychological experiences, difficulties, and

perceptions while using the NaturalReader software.

Participants and Setting

The research was conducted at MA Miftahul Jannah Palangka Raya during the 2025/2026 academic year. The participants consisted of all 30 students enrolled in the tenth grade. They were selected using a total sampling technique because the entire class represented the target population of the study. This cohort was chosen due to their observable difficulties in standard English pronunciation, which were influenced by the phonological characteristics of their regional languages (Banjarese, Dayakese, and Javanese), as well as their access to personal smartphones needed to use the NaturalReader application throughout the study.

Data Collection Instruments

Data were collected using three complementary instruments to ensure methodological triangulation. The first instrument was an oral pronunciation test administered before and after the implementation of NaturalReader. The test consisted of a standardized reading task containing 20 sentences designed to assess both segmental features, including short and long vowels, diphthongs, and voiced–voiceless consonant contrasts, as well as suprasegmental features such as word stress and intonation. Students' pronunciation performances were audio-recorded and assessed using a structured speaking rubric with a five-point rating scale.

The second instrument was classroom observation conducted throughout the instructional process. A non-participant observation protocol was employed to document students' engagement, interaction with the NaturalReader application, learning behaviors, and any challenges encountered during the implementation of the technology.

To complement the quantitative findings, semi-structured interviews were

conducted after the intervention with selected students representing high-, medium-, and low-achieving groups. These interviews were intended to explore students' experiences, perceptions, learning motivation, and opinions regarding the usefulness of NaturalReader in supporting English pronunciation learning.

Data Analysis Procedures

The quantitative data obtained from the pre-test and post-test were analyzed using descriptive and inferential statistics. Descriptive statistics were employed to summarize students' pronunciation performance, while a paired-samples *t*-test was conducted using statistical software to determine whether the differences between pre-test and post-test scores were statistically significant.

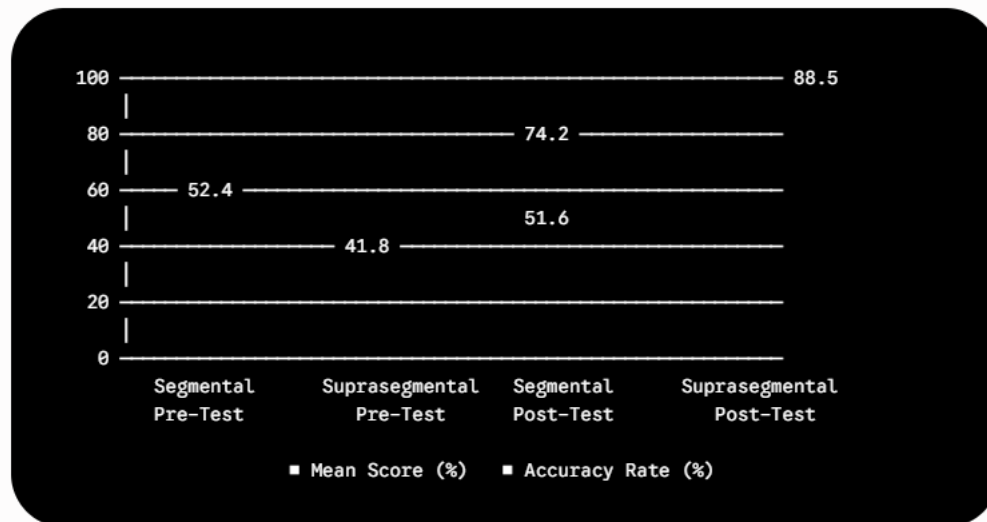
The qualitative data derived from classroom observations and semi-structured interviews were analyzed using thematic analysis. The analysis involved repeatedly reviewing the collected data, generating initial codes, organizing related codes into broader categories, and identifying recurring themes that represented students' experiences and perceptions of using NaturalReader as a Text-to-Speech learning tool. This analytical approach enabled the integration of qualitative findings with the quantitative results, thereby providing a more comprehensive understanding of the effectiveness of the application in pronunciation learning.

RESULT

Quantitative Results: Evaluating Phonetic Effectiveness

The quantitative findings indicate that students' pronunciation performance improved following the implementation of NaturalReader. Improvements were observed not only in the overall pronunciation scores but also in both segmental and suprasegmental aspects of pronunciation.

Figure 1. Comparison of Segmental and Suprasegmental Pronunciation Accuracy Before and After the Intervention



As illustrated in Figure 1, students demonstrated noticeable improvements in both pronunciation components after the intervention. The accuracy of segmental pronunciation increased from 52.4% in the pre-test to 74.2% in the post-test. Likewise, suprasegmental pronunciation improved from

41.8% to 51.6%, indicating better performance in stress placement and intonation patterns after using NaturalReader.

To provide a broader description of students' overall pronunciation achievement, descriptive statistics of the pre-test and post-test scores are presented in Table 1.

Table 1. Descriptive Statistics of Students' Pronunciation Scores

Evaluation Phase	Minimum Score	Maximum Score	Mean Score	Standard Deviation
Pre-Test	42.00	68.00	54.50	6.85
Post-Test	68.00	92.00	81.35	5.40

As shown in Table 1, students' pronunciation performance improved substantially after the implementation of NaturalReader. The mean pronunciation score increased from 54.50 in the pre-test to 81.35 in the post-test, indicating a notable improvement in students' overall pronunciation achievement. In addition, both the minimum and maximum scores increased considerably, with the minimum score rising from 42.00 to 68.00 and the maximum score increasing from 68.00 to 92.00. These results suggest that improvement was experienced not only by higher-performing students but also by those who initially demonstrated lower levels of pronunciation proficiency. Furthermore, the standard deviation decreased from 6.85 in the pre-test to 5.40 in the post-

test, indicating that students' pronunciation scores became more homogeneous after the intervention. Overall, the descriptive statistics demonstrate a clear positive trend in students' pronunciation performance following the use of NaturalReader.

To determine whether the observed improvement was statistically significant, a paired-samples t-test was conducted to compare students' pronunciation scores before and after the implementation of NaturalReader. The analysis yielded a t-value of -14.82 with a significance level of $p < 0.001$, indicating that the difference between the pre-test and post-test scores was statistically significant. Since the obtained p-value was far below the conventional significance threshold of 0.05, the null

hypothesis was rejected, confirming that the improvement in students' pronunciation performance was unlikely to have occurred by chance. These statistical findings provide empirical evidence that the implementation of NaturalReader had a positive effect on students' English pronunciation performance.

Qualitative Results: Documenting Students' Experiences

The qualitative findings were derived from classroom observations and semi-structured interviews conducted after the implementation of NaturalReader. The thematic analysis identified three major themes that describe students' experiences in using the application for learning English pronunciation: reduced speaking anxiety, increased learner autonomy, and improved awareness of regional phonological interference.

Theme 1. Reduced Foreign Language Speaking Anxiety

One of the most prominent findings was the reduction of students' speaking anxiety during pronunciation practice. Many participants explained that practicing English pronunciation in front of teachers or classmates often made them feel embarrassed or afraid of making mistakes. In contrast, NaturalReader provided a private and supportive learning environment that enabled students to practice repeatedly without fear of being judged.

As one participant stated:

"I usually feel embarrassed to try reading English because my tongue feels thick with my Banjarese accent. But the app doesn't laugh at me. I can listen to a word ten times until I get it right."

This finding suggests that the application helped lower students' affective barriers and increased their confidence in practicing English pronunciation independently.

Theme 2. Increased Learner Autonomy Through Mobile Learning

Classroom observations indicated that students quickly became familiar with the features of NaturalReader on their personal smartphones. The application enabled them to upload various text formats and practice pronunciation both inside and outside the classroom. Students also adjusted the playback speed according to their learning needs, slowing the audio to approximately **0.75×** when practicing difficult words before returning to the normal speed once they felt more confident.

These observations demonstrate that the flexibility of the application encouraged students to take greater responsibility for their own learning and supported autonomous pronunciation practice beyond classroom instruction.

Theme 3. Improved Awareness of Regional Phonological Differences

The third theme concerned students' awareness of the differences between their native regional accents and standard English pronunciation. Interview responses indicated that the audio models provided by NaturalReader helped students recognize pronunciation features that had previously gone unnoticed.

Several students from Javanese and Dayakese linguistic backgrounds reported that repeatedly listening to the application's pronunciation model enabled them to distinguish more clearly between voiced and voiceless consonants, thereby improving their pronunciation of English words containing final consonant clusters. This suggests that NaturalReader served not only as a pronunciation model but also as a tool for helping learners recognize and gradually overcome phonological interference from their first language.

DISCUSSION

The findings of this study demonstrate that the implementation of NaturalReader significantly improved students' English

pronunciation, as reflected in both the quantitative and qualitative results. The increase in students' pronunciation scores indicates that repeated exposure to accurate speech models through Text-to-Speech (TTS) technology facilitates more effective pronunciation practice. This finding supports the view of Kelly (2000) that regular exposure to accurate pronunciation models is essential for developing learners' intelligible speech. Likewise, Hasan (2014) emphasizes that mastery of both segmental and suprasegmental features contributes substantially to overall pronunciation proficiency. The improvement observed in this study suggests that NaturalReader effectively assisted students in developing these two fundamental aspects of pronunciation.

The present findings are also consistent with previous studies investigating the use of Text-to-Speech technology in English language learning. Ekşi and Yeşilçınar (2016) reported that online TTS applications significantly improved EFL learners' pronunciation by providing accurate pronunciation models that learners could access repeatedly. Similarly, Bao (2018) found that repeated interaction with TTS applications helped learners develop greater pronunciation automaticity during oral reading. In the present study, students were likewise able to replay words and sentences multiple times until they felt confident with their pronunciation, which may explain the significant improvement identified in the post-test scores. These findings reinforce the argument of Bakken et al. (2012) that TTS technology enables individualized pronunciation practice beyond the limitations of conventional classroom instruction.

An interesting finding of this study is that the improvement was evident not only in students' overall pronunciation scores but also in specific pronunciation components. Segmental accuracy increased considerably, particularly in the production of English vowels, diphthongs, and voiced–voiceless consonants, while suprasegmental features

such as stress and intonation also showed positive development. According to Avery and Ehrlich (1992), learners often struggle with English sounds that do not exist in their first language, making consistent pronunciation modeling an important aspect of instruction. This explanation appears relevant to the present study because many participants came from Banjarese, Dayakese, and Javanese language backgrounds, whose phonological systems differ from English. As explained by Pranoto and Suprayogi (2019), regional language interference frequently influences Indonesian learners' pronunciation, particularly when producing unfamiliar English phonemes. Therefore, the availability of consistent pronunciation models through NaturalReader may have helped students gradually modify their pronunciation habits despite their native-language influence.

Beyond the quantitative improvement, the qualitative findings reveal that NaturalReader also influenced students' affective and behavioral aspects of learning. Many participants reported feeling more confident practicing pronunciation independently because they no longer feared making mistakes in front of teachers or classmates. This finding is in line with Önder (2025), who argues that AI-powered pronunciation tools can reduce learners' speaking anxiety by providing a supportive environment for repeated, self-paced practice. Furthermore, students appreciated the opportunity to control playback speed, repeat difficult words, and practice whenever they wished. Such flexibility reflects the characteristics of autonomous learning described by Pratiwi and Ayu (2020), who found that smartphone-based pronunciation applications encourage learners to practice independently outside formal classroom settings.

The findings also support broader perspectives regarding the integration of educational technology into language learning. Chapelle and Sauro (2019) argue that technology-enhanced language learning provides opportunities for learners to receive

richer linguistic input while promoting greater learner autonomy. Similarly, Fitria (2023) notes that AI-based applications have become increasingly valuable in English language teaching because they provide personalized learning experiences that are difficult to achieve through traditional instruction alone. In the present study, NaturalReader functioned not merely as a pronunciation model but also as an accessible learning companion that allowed students to practice repeatedly according to their own pace and learning needs. This flexibility helped overcome one of the common challenges identified by Levy et al. (2011) and Tergujeff (2012), namely the limited classroom time available for systematic pronunciation instruction.

Although the findings demonstrate the effectiveness of NaturalReader, several considerations should be acknowledged. This study was conducted with a relatively small sample drawn from a single school, which may limit the generalizability of the findings to other educational contexts. In addition, the intervention focused on short-term pronunciation improvement; therefore, the long-term retention of students' pronunciation development remains uncertain. Future studies could involve larger participant groups, compare different AI-based pronunciation applications, and investigate the long-term impact of Text-to-Speech technology on pronunciation accuracy and speaking fluency.

Overall, the findings suggest that NaturalReader is not only effective in improving students' pronunciation accuracy but also successful in fostering learner confidence and autonomy. By combining accessible AI technology with opportunities for repeated and individualized practice, the application offers a practical solution for addressing pronunciation difficulties commonly experienced by EFL learners, particularly those whose first-language phonological systems differ substantially from English.

CONCLUSION

This study demonstrates that the integration of NaturalReader as a text-to-speech learning medium significantly improves the English pronunciation skills of tenth-grade students at MA Miftahul Jannah Palangka Raya. Quantitative data show significant improvements in both individual sound production (segmental features) and overall speech rhythm (suprasegmental features). Qualitatively, the software provides a supportive environment that reduces language anxiety, increases learner autonomy, and helps students systematically address regional phonetic interference. In conclusion, NaturalReader is an effective, flexible, and scalable tool that enhances the quality of modern EFL oral communication instruction.

REFERENCES

- Avery, P., & Ehrlich, S. (1992). *Teaching American English Pronunciation*. Oxford University Press.
- Bakken, S., et al. (2012). Text-to-Speech Technologies in Computer-Assisted Language Learning. *Journal of Educational Technology Systems*, 40(3), 321-334.
- Bao, X. (2018). Utilizing TTS Applications to Build Automaticity in Second Language Oral Reading. *TESOL Quarterly*, 52(2), 411-425.
- Chapelle, C. A., & Sauro, S. (Eds.). (2019). *The Handbook of Technology and Second Language Teaching and Learning*. Wiley-Blackwell.
- Ekşi, G., & Yeşilçınar, S. (2016). The Effectiveness of Online Text-to-Speech Tools on the Pronunciation Skills of EFL Teacher Trainees. *State University Journal of Education*, 43(1), 112-127.
- Endahati, N. (2014). Improving English Pronunciation through Phonetic Drills.

- Journal of English Language Teaching*, 3(2), 45-56.
- Fitria, T. N. (2023). Artificial Intelligence (AI) Technology in English Language Teaching: Focus on Text-to-Speech (TTS) Tools. *Journal of English as a Foreign Language*, 11(1), 89-104.
- Harmer, J. (2007). *The Practice of English Language Teaching* (4th ed.). Pearson Longman.
- Hasan, M. A. (2014). Segmental and Suprasegmental Aspects of English Pronunciation. *International Journal of Linguistics*, 6(4), 201-215.
- Ilmi, Z. (2014). Digital Media in Listening and Speaking Classrooms. *Indonesian Journal of Applied Linguistics*, 4(2), 134-145.
- Kelly, G. (2000). *How to Teach Pronunciation*. Pearson Education Limited.
- Levy, M., et al. (2011). Current Trends in Foreign Language Pronunciation Instruction. *Language Learning & Technology*, 15(3), 80-98.
- Mandasari, B., & Aminatun, D. (2019). Improving Students' Speaking Skill through Digital Media. *Journal of English Language Teaching and Linguistics*, 4(2), 211-222.
- O'Connor, J. D. (1980). *Better English Pronunciation* (2nd ed.). Cambridge University Press.
- Önder, A. (2025). AI-Powered Pronunciation Tools: Minimizing Anxiety and Maximizing Accuracy in EFL Contexts. *Computer Assisted Language Learning*, 38(1), 45-67.
- Oxford University Press. (2005). *Oxford Advanced Learner's Dictionary* (7th ed.). Oxford University Press.
- Pranoto, B. E., & Suprayogi, S. (2019). Regional Phonological Interference in EFL Spoken Output. *Teknosastik*, 17(1), 24-35.
- Pratiwi, Z., & Ayu, M. (2020). The Utilization of Smartphone Applications in Practicing English Pronunciation Autonomous Learning. *Journal of English Language Teaching*, 7(2), 154-167.
- Ramelan. (1985). *English Phonetics*. IKIP Semarang Press.
- Rotatori, V., et al. (2014). Computer-Based Text-to-Speech Applications for Diverse Learners. *Procedia - Social and Behavioral Sciences*, 141, 1012-1017.
- Sharp, C. (2008). Web-Based Text-to-Speech Tools for Auditory Learners. *Multimedia in Education*, 25(4), 56-62.
- Tergujeff, E. (2012). English Pronunciation Teaching in Europe Survey. *Asiaty Journal of Applied Linguistics*, 5(1), 29-45.
- Victor, S. (2014). NaturalReader as an Assistive Tool for PDFs and Web Browsing. *Journal of Special Education Technology*, 29(2), 77-83.
- Wahyukti, C. R. (2008). *Introduction to English Phonology*. Universitas Terbuka Press.