

## **AUTOMATIC SPEECH RECOGNITION (ASR) BASED ON PROGRESSIVE WEB APPS TO DEVELOP PRONUNCIATION LEARNING**

**Muhammad Iqbal<sup>1\*</sup>**

<sup>1</sup>Information System, Universitas Royal

*email: \*iqbalmh@royal.ac.id*

**Abstract:** Good pronunciation plays a crucial role in enhancing students' confidence, encouraging active participation in learning, and preparing them for academic and professional opportunities, such as English-language interviews. Poor pronunciation during scholarship or job interviews can hinder the interviewer's understanding, thereby reducing the chances of acceptance. This study aims to improve students' pronunciation fluency and develop a learning medium based on Automatic Speech Recognition (ASR) technology. The method employed involves the development of Progressive Web Apps (PWA) integrated with ASR technology from the app.lumi.education platform, supported by manual labeling for pronunciation validation. The research was conducted at LKP Vijaya Learning Centre, Tanjungbalai City. The results demonstrate that ASR-based media significantly enhances students' pronunciation accuracy and confidence. Thus, the integration of ASR technology into PWA effectively supports innovative and efficient pronunciation learning.

**Keywords:** automatic speech recognition; language learning; pronunciation; web-based application.

**Abstrak:** Pengucapan yang baik berperan penting dalam meningkatkan kepercayaan diri siswa, mendorong partisipasi aktif dalam pembelajaran, dan mempersiapkan mereka menghadapi peluang akademik maupun profesional, seperti wawancara berbahasa Inggris. Saat menghadapi wawancara beasiswa atau pekerjaan berbahasa Inggris, pengucapan yang buruk dapat mengurangi pemahaman pewawancara, sehingga mengurangi peluang diterima. Penelitian ini bertujuan untuk meningkatkan kelancaran pengucapan siswa dan mengembangkan media pembelajaran berbasis teknologi Automatic Speech Recognition (ASR). Metode yang digunakan adalah pengembangan Progressive Web Apps (PWA) yang terintegrasi dengan ASR dari aplikasi app.lumi.education, didukung oleh pelabelan manual untuk validasi pengucapan. Penelitian dilakukan di LKP Vijaya Learning Centre, Kota Tanjungbalai. Hasil penelitian menunjukkan bahwa media berbasis ASR secara signifikan meningkatkan akurasi pengucapan dan kepercayaan diri siswa. Dengan demikian, integrasi teknologi ASR dalam PWA terbukti mendukung pembelajaran pengucapan secara inovatif dan efisien.

**Kata kunci:** aplikasi berbasis web; pengenalan suara otomatis; pembelajaran bahasa; pengucapan

### **INTRODUCTION**

English holds significant importance in the international context, especially in the current era of globaliza-

tion. As a language widely used in the global community, English serves as a communication tool among individuals from various countries. One of the critical aspects of learning English is pronun-

ciation[1]. Learning proper pronunciation is one of the most crucial steps an individual can take to enhance their communication skills[2].

Pronunciation in English becomes complex due to the lack of clear guidelines and contradictory practices[3]. In practice, pronunciation learning has not been given priority, as most teachers believe that teaching pronunciation is too challenging and unengaging for students[4]. Therefore, the aim of this study is to enhance students' fluency in pronunciation and develop digital-based English learning media utilizing Automatic Speech Recognition (ASR) technology. Additionally, the research seeks to optimize learning time, considering the large number of students in a class, which requires extensive time for individual assessment. Given the limited instructional time, additional sessions are often needed, leading to increased operational costs.

This research was conducted at LKP Vijaya Learning Centre, which focuses on English and Computer learning in Tanjungbalai City. Automatic Speech Recognition (ASR) is a technology that enables computers to interpret and understand human speech automatically. Meanwhile, Progressive Web Apps (PWA) represent a new architecture for mobile application development, offering advantages over their predecessors (websites). These applications can function with weak or no internet connection, run in the background, and support push notifications[5]. An integrated system is also utilized to enhance system value through new functionalities provided by connecting different system functions[6]. Currently, RESTful API has become the standard API widely adopted when developing systems that need to provide APIs for third parties[7].

Research on Automatic Speech

Recognition conducted by previous discusses the issue of unsupervised Automatic Speech Recognition (ASR), which excludes any form of manual labeling by humans. However, there remains a significant gap between supervised, semi-supervised, and unsupervised models[8]. Riefkyanov's study (2024), titled Analysis of Whisper Automatic Speech Recognition Performance on Low-Resource Languages[8], aims to improve Automatic Speech Recognition (ASR) for low-resource languages by utilizing a dataset in Javanese. The research highlights that current ASR models predominantly focus on speech recognition in high-resource languages such as English

The study Towards Inclusive Automatic Speech Recognition[9], indicates that only a small portion of bias (in pronunciation) can be attributed to differences in articulation among speaker groups. Furthermore, the study emphasizes the need to develop language- and architecture-specific solutions to mitigate such bias. Unfortunately, this research does not yield a product design that can be tested by the wider public. The development of an English Speed Reading Practice Application based on Progressive Web App (PWA) using the Prototyping method [1] demonstrates that the application is compatible with mobile devices and operates across browsers. However, the implementation of PWA is intended for practicing speed reading, not pronunciation

PWA is supported by most major browsers, particularly on Android and Chromium-based browsers, except on iOS, where only Safari provides support for PWA-related functionalities[10]. This technology enables the creation of applications that can be accessed like native apps, interact with device features, run offline, and be accessed directly through

a browser, allowing for a broader reach. On the other hand, conventional mobile application development requires significant financial investment for implementation and demands greater resources[11].

**METHOD**

The problem-solving method employed includes needs analysis through surveys and interviews to understand the characteristics and needs of students in pronunciation learning. Information is gathered about ASR technology, pronunciation teaching methods, and PWA application development. An initial prototype of the learning media based on PWA integrated with ASR technology is created. The prototype is tested involving students and teachers to evaluate its effectiveness, and feedback from the testing is used to refine the prototype to improve its quality and alignment with the needs of the partners.

The strategies implemented for

problem-solving involve forming a multidisciplinary team consisting of language experts and software developers. Students and teachers are engaged in every development stage, from needs analysis to prototype evaluation. ASR technology from lumi.education is utilized and embedded into the PWA to create responsive and interactive learning media. This study involves training teachers and students in the use of the application, followed by data collection and analysis of learning outcomes.

This research utilizes the Automatic Speech Recognition (ASR) from the lumi education application through the "Speak the Word Set" to enhance the features and development of digital pronunciation learning via an embedded system displayed using a Progressive Web App (PWA). Several measurable variables are used to assess the success of the research, including the ASR performance, which can be seen in Table 1.

Table 1. Variables to Measure the Success of the Research

No.	Variable	Description
1	Accuracy (%)	Percentage of total speech correctly recognized by the system
2	Response Time (Seconds)	The time required by the system to detect and transcribe speech
3	Error(%)	Percentage of total speech incorrectly recognized by the system.
4	Dialect (%)	The percentage of system success in recognizing and transcribing speech in dialect or language variations.

In achieving the research objectives, the – outlines the research steps as shown in Image 1.

**Problem Identification**

The problem identified in this study is the limited ability of teachers and students to provide feedback on pronunciation, making it difficult for students to correct their pronuncia-

tion errors. Subsequently, the Principal Investigator and Research Team members conducted a literature review aimed at gaining an understanding of the issues to be investigated related to Pronunciation, Automatic Speech Recognition, and the design of learning media based on Progressive Web Apps, as well as the methods used as solutions.

**Data Collection**

The data collection techniques used in this study include interviews, observations, and documentation.

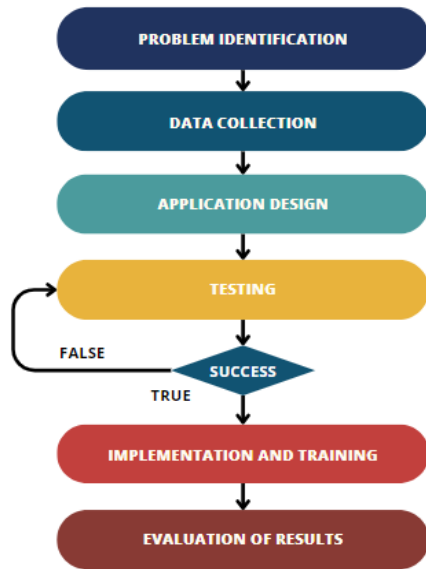


Image 1. Framework

**Application Design**

In this phase, the database design and the user interface/user experience (UI/UX) design are developed to ensure an intuitive and responsive interface, assisted by team members. The application design includes features for creating learning materials, quizzes, and discussions that align with the curriculum at Vijaya Learning Centre, which are input by the team members. Additionally, the integration of the Automatic Speech Recognition (ASR) technology from lumi.education, accessible via the Progressive Web App (PWA), is implemented. The mechanism of ASR from lumi.education is illustrated in Image 2.

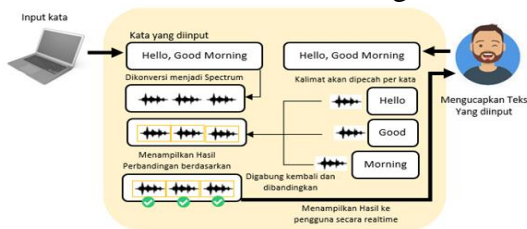


Image 2. Ilustration of ASR Application via Lumi.education

**Testing**

In this phase, the team conducts application testing to ensure that the system responds accurately to students' pronunciation and provides precise and helpful feedback to aid students in improving their pronunciation.

**Implementation and Training**

The implementation involves the project leader and team members in deploying the ASR application using PWA at Vijaya Learning Centre. Instructors and students will receive training on how to use the application. During the training sessions, users will be provided with knowledge about the application's features, how to correct pronunciation, and how to interpret the feedback provided and documented by the research team members.

**Evaluation of Results**

This evaluation is conducted by the lead to assess the effectiveness of the ASR application using PWA in improving students' pronunciation skills. Instructors collect data and observe the progress of students' pronunciation after using the application. The method used to measure success is through the System Usability Scale (SUS) questionnaire. When calculating the SUS score, there are several rules for scoring. For questions with odd numbers, the score for each question is obtained by subtracting 1 from the user's score. For even-numbered questions, the final score is derived by subtracting the user's score from 5. The SUS score is obtained by summing the scores of all questions and then multiplying the result by 2.5.

**RESULTS AND DISCUSSION**

This study obtained data presented in Table 2, which shows the results of ASR testing based on several measured variables, namely Accuracy (%), Response Time (Seconds), and Error Rate

(%).

Table 2. ASR Performance

No.	Users	Accuracy (%)	Response Time (Seconds)	Error (%)
1	A	92	1.2	8
2	B	87	1.5	13
3	C	95	1.0	5
4	D	89	1.3	11
5	E	93	1.1	7
<b>Average</b>		91.2	1.22	8.8

From Table 2, it can be observed that the average accuracy of speech recognition reached 91.2%, with an average response time of 1.22 seconds and an average error rate of 8.8%. The relatively high accuracy level indicates that ASR is capable of recognizing speech well in most cases. The accuracy ranged from 87% to 95%, demonstrating consistent performance in recognizing correctly pronounced words. The highest accuracy was achieved by user C (95%), while the lowest was by user B (87%).

The response time ranged from 1.0 to 1.5 seconds, with the fastest response at 1 second (user C) and the slowest at 1.5 seconds (user B). The average response time of 1.22 seconds is still within acceptable limits for educational applications, where speed is important but can be compromised if accuracy is maintained. The error rate ranged from 5% to 13%, with user C showing the lowest error rate (5%) and user B recording the highest (13%). These errors could be attributed to variations in dialect or the clarity of speech.

**Presentation of Data on PWA Usability**

The usability of the PWA was measured using the System Usability Scale (SUS), which provides a score based on users' perceptions of the system's ease of use.

Table 3. System Usability Scale (SUS) Scores for PWA

No.	Users	SUS Score (0-100)
1	A	82
2	B	76
3	C	85
4	D	78
5	E	80
<b>Average</b>		80.2

From the results in Table 3, the average SUS score reached 80.2, indicating that users generally perceive the PWA-based application as easy to use and intuitive. This score exceeds the common SUS benchmark, where a score of 68 is considered the threshold between acceptable and unacceptable usability. User C provided the highest score of 85, reflecting a highly positive user experience, while User B recorded the lowest score of 76.

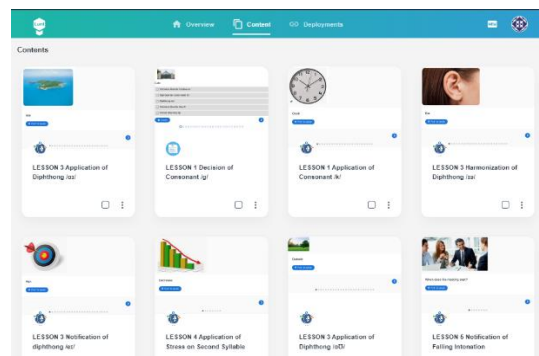


Image 3. The dashboard page contains English lesson materials.

The dashboard page organizes English lesson materials in a structured and accessible manner. The materials are categorized by topic or level, accompanied by brief descriptions to help users understand the content. With simple navigation and a search feature, users can easily locate topics such as grammar, vocabulary, or other skills. The user-friendly design ensures an efficient learning experience.



When does the meeting start?

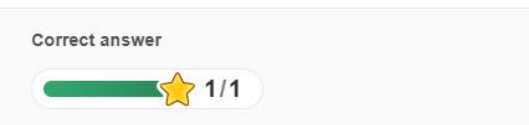


Image 4. Pronunciation testing with correct results through the application.

The pronunciation testing through the application is conducted to evaluate users' ability to articulate words or phrases accurately according to established standards. This process involves users speaking the words or sentences displayed by the application, which are then analyzed using voice recognition technology. The application provides immediate feedback, indicating whether the pronunciation is correct or requires improvement. With this feature, users can enhance their speaking skills independently and effectively. Meanwhile, the pronunciation testing with incorrect results through the application is conducted to identify errors in users' articulation.

After users pronounce the words or sentences displayed, the system analyzes the pronunciation using voice recognition technology. If the pronunciation does not meet the specified standards, the application provides feedback in the form of an error notification. Additionally, the system offers users two options: repeat the pronunciation to try again or view the captured pronunciation results. This feature is designed to help users understand their mistakes and improve their pronunciation gradually and independently.



Would you like a cup of tea?

Your answer(s) was interpreted as:  
would you like a cup ❌ would you like a cup of ❌

The correct answer(s):  
Would you like a cup of tea?

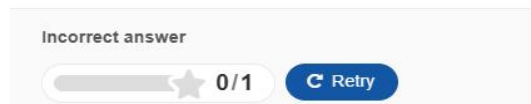


Image 5. Pronunciation testing with incorrect results through the application.

Despite this, the score remains within the range considered good, signifying that the PWA application features a user-friendly interface and strong performance.

## CONCLUSION

The performance of the ASR demonstrates exceptional capability with a high accuracy rate of 91.2%, indicating superior proficiency in recognizing users' pronunciation. Furthermore, the fast response time of 1.22 seconds enhances the ASR's ability to provide real-time feedback effectively. However, the error rate of 8.8%, while still within acceptable limits, suggests potential improvement in recognizing dialects or variations in pronunciation. In parallel, the usability of the PWA is validated by an average SUS score of 80.2, signifying positive user reception of the application as a learning medium. The implementation of the PWA as a pronunciation learning platform is deemed effective in delivering an intuitive and easily

accessible user experience. Thus, the findings of this study indicate that the combination of ASR and PWA holds significant potential to advance digital-based pronunciation learning media.

#### ACKNOWLEDGMENT

We express our gratitude to LKP Vijaya Learning Centre, Kota Tanjung-balai, for their support and participation in conducting this research. Our heartfelt thanks also go to the ASR technology development team at app.lumi.education for providing an innovative platform that supported this study. The contributions of all parties have enabled the achievement of significant results in developing technology-based learning media that benefit students. We hope this research will make a positive contribution to the field of education, particularly in improving students' English pronunciation skills.

#### BIBLIOGRAPHY

[1] M. F. Alfath, L. Fanani, and A. P. Kharisma, "Pengembangan Aplikasi Berlatih Membaca Cepat Berbahasa Inggris Berbasis Progressive Web App dengan Metode Prototyping," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 7, pp. 1461–1468, 2023, doi: 10.25126/jtiik.1077982.

[2] F. N. Amalia<sup>1</sup>, S. S. Rambe<sup>2</sup>, and Y. Lubis<sup>3</sup>, "Literature Review On English Consonant Pronunciation Problems Encountered By Indonesian Students," *J. Pendidik. dan Sos. Hum.*, vol. 3, no. 3, pp. 175–184, 2023, [Online]. Available: <https://doi.org/10.55606/khatulistiwa.v3i3.1922>

[3] A. P. Gilakjani, "English Pronunciation Instruction: A Literature Review - International Journal of Research in English Education," *Ijree*, vol. 1, no. 1, pp. 1–6, 2016, [Online]. Available: <https://ijreeonline.com/article-1-21-en.html>

[4] S. Nasional and D. Mengabdi, "SENADA ;," vol. 04, no. 01, pp. 10–19, 2023.

[5] I. Malavolta, K. Chinnappan, L. Jasmontas, S. Gupta, and K. A. K. Soltany, "Evaluating the impact of caching on the energy consumption and performance of progressive web apps," *Proc. - 2020 IEEE/ACM 7th Int. Conf. Mob. Softw. Eng. Syst. MOBILESoft 2020*, pp. 109–119, 2020, doi: 10.1145/3387905.3388593.

[6] M. Iqbal, A. . Syahputra, Y. M. Nurwati, and A. Nasution, *Sistem Terintegrasi Menggunakan API: Implementasi Website dan Android*. 2022.

[7] M. Iqbal and N. Nurwati, "Penerapan Sistem Terintegrasi Menggunakan Restful Api Pada Dealer Management System Panca Niaga Sei Piring," *J. Sci. Soc. Res.*, vol. 6, no. 1, p. 219, 2023, doi: 10.54314/jssr.v6i1.1161.

[8] H. Aldarmaki, A. Ullah, S. Ram, and N. Zaki, "Unsupervised Automatic Speech Recognition: A review," *Speech Commun.*, vol. 139, pp. 76–91, 2022, doi: 10.1016/j.specom.2022.02.005.

[9] S. Feng, B. M. Halpern, O. Kudina, and O. Scharenborg, "Towards inclusive automatic speech recognition," *Comput. Speech Lang.*, vol. 84, no. September 2023, p. 101567, 2024,

- doi: 10.1016/j.csl.2023.101567.
- [10] B. T. Kane Examiner and J. Sören Pettersson, “Dirk-Jan Rensema The Current State of Progressive Web Apps,” 2020.
- [11] C. Huda, A. Ramadhan, A. Trisetyarso, E. Abdurachman, and Y. Heryadi, “Smart Tourism

Recommendation Model: A Systematic Literature Review,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 12, pp. 167–174, 2021, doi: 10.14569/IJACSA.2021.0121222.