REKAYASA
Journal of Science and Technology
https://journal.trunojoyo.ac.id/rekayasa

REKAYASA ,19(1): 2026 pp.80-95
ISSN: 0216-9495 (Print)
ISSN: 2502-5325 (Online)

Pengembangan Asisten Pembelajaran Bahasa Inggris Portabel Berbasis Artificial Intelligence dengan Integrasi Speech Recognition dan Speech Synthesis

Muhammad Ricky Rizaldi [1*], Ayu Novia Lisdawati[2)]
[1, 2] Universitas Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin
*email: mrickyrizaldi@gmail.com

DOI: https://doi.org/10.21107/rekayasa.v19i1.33686
Submitted March, 15th 2026, Accepted April, 27th 2026, Published May, 7th 2026

Abstrak

Pembelajaran Bahasa Inggris berbasis aplikasi telepon pintar umumnya bergantung pada koneksi internet dan rentan terhadap distraksi penggunaan perangkat. Kondisi tersebut mendorong kebutuhan akan perangkat pembelajaran khusus yang mampu beroperasi secara mandiri dan lebih terfokus. Penelitian ini bertujuan untuk merancang dan mengembangkan asisten pembelajaran Bahasa Inggris portabel berbasis kecerdasan buatan bernama POCALA yang mampu beroperasi dalam mode daring dan luring. Kontribusi penelitian ini adalah pengembangan perangkat pembelajaran Bahasa Inggris portabel yang mengintegrasikan model bahasa, pengenalan suara, dan sintesis suara dalam satu sistem embedded dual-mode untuk mendukung pembelajaran yang lebih terfokus dan aksesibel. Metode penelitian yang digunakan adalah pendekatan perancangan sistem dan prototyping yang meliputi perancangan perangkat keras dan perangkat lunak, pengembangan prototipe, pengujian fungsional, serta evaluasi pengguna. Perangkat diimplementasikan pada platform Raspberry Pi 5 dengan integrasi model bahasa berbasis kecerdasan buatan, pengenalan suara, sintesis suara, serta interaksi berbasis suara dan tampilan visual. Hasil pengujian menunjukkan akurasi pengenalan suara sebesar 96% pada mode daring dan 84% pada mode luring. Mode daring memberikan waktu respons lebih cepat, sedangkan mode luring memungkinkan operasional tanpa ketergantungan pada koneksi internet. Daya tahan baterai dan stabilitas sistem berada dalam batas operasional yang baik. Evaluasi terhadap 20 responden menunjukkan tingkat penerimaan tinggi dengan nilai rata-rata di atas 4,0 pada aspek fungsionalitas, aksesibilitas, dan kemudahan penggunaan. Hasil tersebut menunjukkan bahwa perangkat layak digunakan sebagai asisten pembelajaran Bahasa Inggris portabel yang lebih terfokus, adaptif, dan aksesibel dibandingkan aplikasi pembelajaran berbasis telepon pintar konvensional.

Kata Kunci: interaksi berbasis suara, model bahasa, perangkat edge, perangkat edukatif, prototipe, sistem tertanam.

Abstract

Smartphone-based English learning applications generally depend on internet connectivity and are vulnerable to device-related distractions. This condition highlights the need for a dedicated learning device that can operate independently and provide a more focused learning experience. This study aims to design and develop a portable artificial intelligence-based English learning assistant named POCALA that can operate in both online and offline modes. The research contribution is the development of a portable English learning device that integrates language models, speech recognition, and speech synthesis into a dual-mode embedded system to support more focused and accessible learning. The research employed a system design and prototyping approach, including hardware and software design, prototype development, functional testing, and user evaluation. The device was implemented on a Raspberry Pi 5 platform with integrated artificial intelligence-based language models, speech recognition, speech synthesis, voice-based interaction, and visual display output. The results showed speech recognition accuracy of 96% in online mode and 84% in offline mode. Online mode provided faster response time, while offline mode enabled operation without internet dependency. Battery endurance and system stability remained within acceptable operational limits. Evaluation involving 20 respondents showed high user acceptance, with mean scores above 4.0 in functionality, accessibility, and usability. These findings indicate that POCALA is feasible as a portable English learning assistant that is more focused, adaptive, and accessible than conventional smartphone-based learning applications.

Keywords*: edge device, educational device, embedded system, language model, prototype, voice-based interaction*

© 2026 Authors

PENDAHULUAN

Kemampuan berbahasa Inggris memiliki peran strategis dalam era globalisasi, khususnya dalam ranah akademik, profesional, dan sosial. Bahasa Inggris telah menjadi kompetensi esensial untuk beradaptasi dan bersaing di berbagai sektor kehidupan. Sebagai bahasa internasional yang secara resmi digunakan di lebih dari 60 negara dan dituturkan oleh lebih dari 400 juta penutur asli, kemahiran berbahasa Inggris merupakan aset penting dalam menghadapi dinamika global (Pranoto & Inharjanto, 2020). Namun, tingkat kemahiran berbahasa Inggris masyarakat Indonesia masih tergolong rendah. Berdasarkan laporan EF English Proficiency Index 2024, Indonesia menempati peringkat ke-80 dari 116 negara dan dikategorikan memiliki tingkat kemahiran rendah (EF Education First, 2024), yang mengindikasikan adanya kesenjangan signifikan dalam penguasaan bahasa Inggris.

Rendahnya tingkat kemahiran tersebut dipengaruhi oleh berbagai faktor, termasuk keterbatasan kosakata, kesulitan pelafalan, rendahnya rasa percaya diri, terbatasnya kesempatan praktik, serta metode pembelajaran yang cenderung monoton (S. Ahmad, 2023). Selain itu, keterampilan *reading*, *speaking*, *listening*, dan *writing* yang belum optimal masih menjadi tantangan utama. Salah satu pendekatan yang efektif untuk meningkatkan kemahiran berbahasa Inggris adalah menumbuhkan kepercayaan diri peserta didik dalam mempraktikkan bahasa serta mengaplikasikan materi yang telah dipelajari secara konsisten (Hidayah, 2021).

Untuk mengatasi berbagai permasalahan tersebut, sebagian besar solusi berfokus pada modifikasi metode pengajaran dan pemanfaatan teknologi melalui aplikasi pembelajaran berbasis *smartphone*. Aplikasi *mobile* berbasis AI, seperti *ELSA Speak* dan *CAKE*, memanfaatkan teknologi *speech recognition* untuk memberikan umpan balik langsung serta meningkatkan akurasi *speaking* peserta didik (Marlinda & Huda, 2024; Saragih et al., 2023). Namun, pembelajaran melalui *smartphone* kerap terganggu oleh distraksi perangkat, seperti akses ke media sosial dan permainan digital, yang berpotensi menurunkan fokus peserta didik selama sesi pembelajaran (Zulfa & Mujazi, 2022).

Permasalahan tersebut tidak hanya berkaitan dengan pendekatan pedagogis, tetapi juga merepresentasikan tantangan rekayasa dalam merancang sistem pembelajaran yang terdedikasi, portabel, minim distraksi, dan mampu mengintegrasikan perangkat keras serta perangkat lunak secara efisien untuk mendukung interaksi belajar yang lebih terfokus.

Selain itu, penerapan *artificial intelligence* (AI) memungkinkan terciptanya pengalaman pembelajaran bahasa yang lebih personal dan interaktif. Teknologi seperti *machine translation*, *speech recognition*, *chatbot*, *virtual assistant*, serta *generative AI* terbukti mendorong motivasi dan meningkatkan keterlibatan peserta didik dalam praktik bahasa secara mandiri (Patty, 2024). Fitur berbasis AI yang menyediakan umpan balik instan serta latihan adaptif sesuai kebutuhan individu juga berkontribusi terhadap peningkatan keterampilan *speaking*, *listening*, dan *writing* (Hidayatullah, 2024). Meskipun demikian, sebagian besar solusi berbasis AI masih bergantung pada konektivitas internet yang stabil, sehingga membatasi kontinuitas dan aksesibilitas pembelajaran, khususnya dalam konteks negara berkembang seperti Indonesia.

Pembelajaran berbasis audio melalui *text-to-speech* (TTS) telah banyak diterapkan untuk mendukung pembelajaran bahasa Inggris dengan memfasilitasi transfer pengetahuan, memperbaiki *pronunciation* dan *listening comprehension*, serta meningkatkan kelancaran membaca (Widyana et al., 2022). TTS juga berperan sebagai media yang menjembatani teks tertulis dengan bahasa lisan serta berkontribusi terhadap peningkatan pemahaman dan konsentrasi, khususnya bagi peserta didik dengan disabilitas (Manu & Masan, 2020). Meskipun demikian, keterbatasan pada aspek *naturalness* suara dan interaksi *real-time* masih menjadi tantangan dalam implementasinya.

Automatic speech recognition (ASR) juga semakin luas dimanfaatkan untuk mendukung praktik lisan dengan memungkinkan latihan berbicara secara berulang serta menyediakan umpan balik yang langsung dan

objektif. Integrasi ASR dalam pembelajaran bahasa terbukti memfasilitasi praktik mandiri, menurunkan kecemasan berbicara, serta mengoptimalkan performa lisan seperti *fluency* dan *lexical complexity* (Jiang et al., 2022).

Selain teknologi berbasis ucapan, *machine translation* (MT) dimanfaatkan untuk membantu peserta didik memahami dan menghasilkan teks lintas bahasa. Perkembangan *neural machine translation* meningkatkan kecepatan dan akurasi terjemahan sehingga mendukung pengembangan kosakata serta pemahaman teks, terutama ketika tingkat kemahiran masih terbatas (Kirchhoff, 2024). Namun demikian, MT memiliki keterbatasan dalam memahami konteks dan nuansa budaya sehingga lebih tepat diposisikan sebagai alat bantu dalam proses pembelajaran (Naveen & Trojovský, 2024).

Perkembangan *generative AI* mendorong lahirnya *Large Language Models* (LLMs) yang menunjukkan potensi sebagai asisten pembelajaran melalui interaksi bahasa alami dan penyediaan umpan balik terpersonalisasi (Dong et al., 2024). Dalam konteks *LLM-enabled education*, integrasi dengan *speech recognition* dan *speech synthesis* memungkinkan praktik bahasa lisan yang lebih adaptif (Gan et al., 2023). Namun, implementasi LLM memerlukan pengendalian berbasis *prompt* agar respons yang dihasilkan tetap bersifat instruksional dan selaras dengan tujuan pedagogis (Pester et al., 2024).

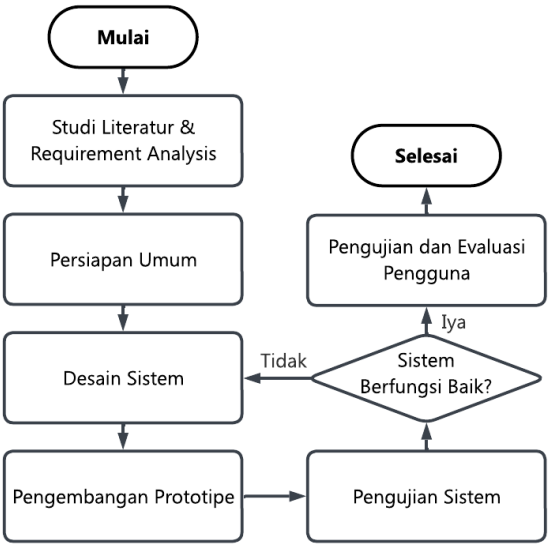
Meskipun berbagai teknologi tersebut telah berkembang secara signifikan, sebagian besar penelitian sebelumnya masih berfokus pada implementasi parsial, seperti aplikasi pembelajaran berbasis *smartphone*, layanan berbasis *cloud*, atau integrasi tunggal pada fitur tertentu seperti *speech recognition* maupun *text-to-speech*. Pendekatan tersebut menunjukkan potensi yang baik dalam mendukung pembelajaran bahasa, namun masih menyisakan keterbatasan pada aspek distraksi penggunaan, ketergantungan konektivitas, dan keterpaduan antarmuka pembelajaran. Penelitian yang secara khusus mengintegrasikan *Large Language Model*, pengenalan suara, dan sintesis suara dalam satu perangkat tertanam yang dirancang untuk pembelajaran bahasa serta mampu beroperasi secara fleksibel pada mode daring dan luring masih relatif terbatas.

Berdasarkan berbagai tantangan tersebut serta potensi transformatif dari perkembangan terbaru dalam *artificial intelligence*, diperlukan pengembangan alat pembelajaran bahasa Inggris yang lebih terspesialisasi, portabel, dan minim distraksi dibandingkan aplikasi berbasis *smartphone* konvensional. Oleh karena itu, penelitian ini bertujuan untuk merancang dan mengembangkan asisten pembelajaran bahasa Inggris portabel berbasis kecerdasan buatan yang mengintegrasikan *Large Language Model,* pengenalan suara, dan sintesis suara dalam satu sistem *embedded* yang mampu beroperasi pada mode daring maupun luring. Kontribusi penelitian ini adalah pengembangan perangkat pembelajaran bahasa Inggris yang lebih terfokus, adaptif, dan aksesibel bagi berbagai kalangan pengguna, termasuk individu dengan disabilitas, serta evaluasi fungsional dan penerimaan pengguna terhadap sistem yang dikembangkan untuk menilai kelayakan sistem dalam mendukung proses belajar bahasa Inggris secara interaktif pada berbagai kondisi penggunaan.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan perancangan sistem dan prototyping untuk mengembangkan serta mengevaluasi asisten pembelajaran bahasa Inggris portabel bernama POCALA (Pocket A Language). Pendekatan ini dipilih karena memungkinkan pengembangan sistem secara iteratif, pengujian fungsional bertahap, serta validasi kinerja terhadap integrasi perangkat keras dan perangkat lunak dalam satu sistem terpadu.

Proses prototyping dilaksanakan selama empat bulan di Laboratorium Robotika, Program Studi Teknik Elektro, Universitas Islam Kalimantan Muhammad Arsyad Al Banjari, Banjarmasin. Tahapan penelitian meliputi persiapan umum, perancangan sistem perangkat keras dan perangkat lunak, pembangunan prototipe, serta pengujian kinerja dan evaluasi pengguna. Alur pelaksanaan penelitian secara keseluruhan ditunjukkan pada diagram alir pada Gambar 1.



Gambar 1. Diagram alir tahapan penelitian

2.1 Persiapan Umum

Tahap awal penelitian diawali dengan studi literatur yang bersumber dari jurnal ilmiah, artikel, serta repositori open-source yang berkaitan dengan penerapan speech recognition, speech synthesis, machine translation, dan Large Language Models (LLMs) dalam sistem pembelajaran bahasa. Studi pendahuluan ini bertujuan untuk menyusun landasan konseptual dalam perancangan arsitektur sistem, penentuan spesifikasi komponen, serta perumusan strategi implementasi guna memastikan bahwa prototipe yang dikembangkan memenuhi aspek kelayakan teknis dan kebutuhan fungsional.

Selanjutnya dilakukan persiapan komponen perangkat keras dan perangkat pendukung. Komponen utama yang digunakan meliputi Raspberry Pi 5, ReSpeaker Lite, TFT LCD, speaker, push button, modul power bank, baterai isi ulang, konektor, serta material casing. Raspberry Pi dipilih sebagai komputer papan tunggal yang memiliki ukuran ringkas namun mampu menjalankan program sistem serta mengelola proses input dan output secara terintegrasi. ReSpeaker Lite berfungsi sebagai antarmuka input audio untuk menangkap suara pengguna dalam proses pengenalan suara, sedangkan TFT LCD digunakan sebagai media keluaran visual untuk menampilkan teks pembelajaran dan respons sistem. Keluaran audio dihasilkan melalui speaker untuk mendukung proses sintesis suara dan umpan balik interaktif.

Lingkungan perangkat lunak menggunakan bahasa pemrograman Python dengan pustaka pendukung untuk pengembangan sistem, serta Raspberry Pi OS sebagai sistem operasi utama. Perancangan rangkaian elektronik dan pemodelan casing tiga dimensi dilakukan menggunakan perangkat lunak Autodesk Fusion 360 dan Blender.

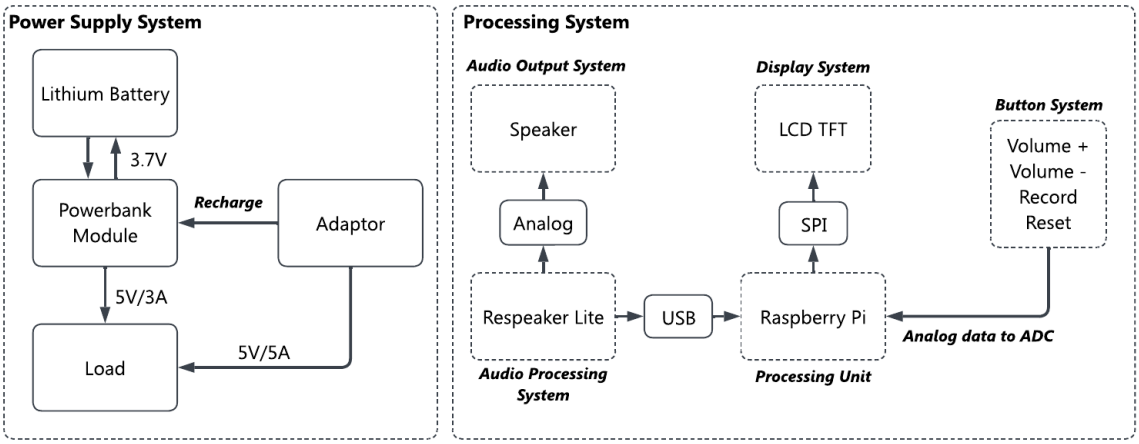
2.2 Perancangan Sistem

2.2.1 Perancangan Perangkat Keras

Perancangan perangkat keras dibagi menjadi dua subsistem utama, yaitu subsistem elektrik dan subsistem mekanik. Subsistem elektrik mencakup konfigurasi pengkabelan (*wiring*) antar komponen utama, yang terdiri atas ReSpeaker Lite sebagai modul input suara, speaker sebagai keluaran audio, LCD sebagai media tampilan visual, serta empat push button yang terhubung ke pin GPIO dan port komunikasi pada Raspberry Pi.

Subsistem daya terdiri atas baterai isi ulang yang terintegrasi dengan modul power bank sebagai pengatur tegangan dan proteksi arus. Sistem distribusi daya dikendalikan melalui slide switch yang berfungsi untuk mengatur aliran daya ke seluruh komponen perangkat. Konfigurasi ini dirancang untuk mendukung operasi portabel sekaligus menjaga keamanan sistem dari fluktuasi daya.

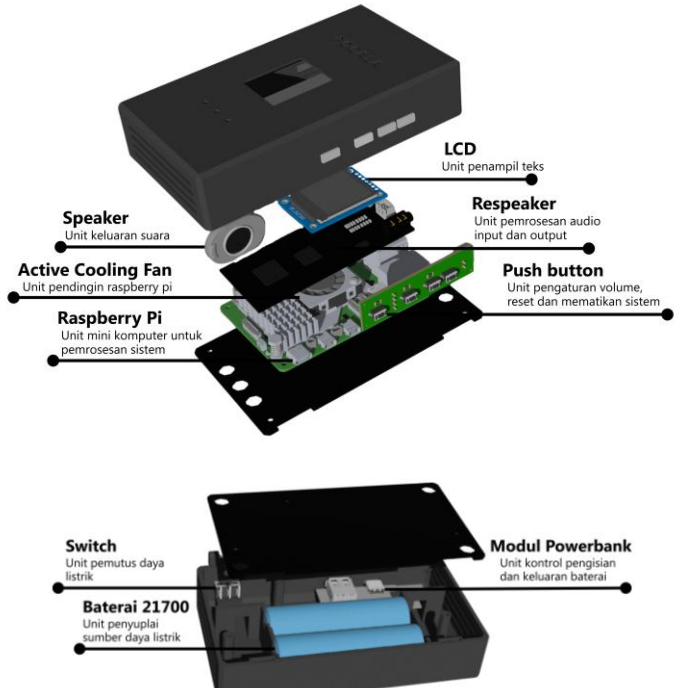
Arsitektur perangkat keras secara keseluruhan ditunjukkan dalam diagram blok pada Gambar 2, yang menggambarkan hubungan fungsional serta alur data antar komponen sistem.



Gambar 2. Diagram blok perangkat keras

Untuk menyederhanakan integrasi tombol, dirancang PCB push button kustom berukuran 60 × 15 mm menggunakan Autodesk Fusion 360. PCB tersebut mengintegrasikan empat tombol dengan konfigurasi pull-up resistor guna memastikan stabilitas logika sinyal pada saat pembacaan input melalui GPIO. Setelah proses perakitan, dilakukan pengujian integrasi dengan memprogram Raspberry Pi agar dapat berkomunikasi dengan LCD untuk menampilkan teks, ReSpeaker Lite untuk menerima input suara, speaker untuk menghasilkan keluaran audio, serta push button yang dikonfigurasi untuk fungsi kontrol volume, perekaman suara, dan restart/shutdown sistem.

Perancangan mekanik dilakukan melalui pembuatan desain casing cetak 3D menggunakan Blender. Casing terdiri atas dua bagian utama, yaitu bagian atas sebagai kompartemen komponen pemrosesan dan bagian bawah sebagai kompartemen sistem daya, masing-masing berukuran 130×75 ×25 mm, dengan tambahan penutup berukuran 110× 55×15 mm. Desain mekanik mempertimbangkan aspek sirkulasi udara untuk menjaga stabilitas suhu perangkat, lubang lanyard untuk mendukung portabilitas, serta mekanisme buka-tutup berbasis engsel yang dikunci menggunakan kawat pengunci guna memudahkan akses saat perawatan. Desain visual prototipe perangkat, termasuk tata letak komponen internal dan struktur casing eksternal, ditampilkan pada Gambar 3.



Gambar 3. Desain prototipe perangkat

2.2.2 Perancangan Perangkat Lunak

Perancangan perangkat lunak diawali dengan penyusunan diagram alir (*flowchart*) alur kerja sistem untuk memastikan kesesuaian antara kebutuhan fungsional pengguna dan implementasi teknis sistem. Diagram alir tersebut menggambarkan proses pemilihan mode, penerimaan input suara, pemrosesan teks, hingga penyajian keluaran dalam bentuk visual dan audio. Sistem dirancang dalam tiga mode operasional utama, yaitu:

1. Translator Mode, yang berfungsi menerjemahkan kalimat antara bahasa Indonesia dan bahasa Inggris secara dua arah.
2. Vocabulary Mode, yang menampilkan informasi kosakata meliputi terjemahan, representasi fonetik (*phonetic*), definisi, serta contoh penggunaan dalam kalimat.
3. Assistant Mode, yang terdiri atas empat fitur fungsional:
    - Grammar, memberikan koreksi tata bahasa disertai penjelasan
    - Asking, merespons pertanyaan terkait pembelajaran secara ringkas dan kontekstual
    - Question, menyediakan soal latihan beserta evaluasi jawaban
    - Speaking, memfasilitasi praktik percakapan interaktif dengan kesinambungan konteks dialog.

Arsitektur perangkat lunak dirancang secara modular, di mana setiap mode dan fitur diimplementasikan sebagai modul terpisah yang saling terhubung melalui manajemen alur utama sistem. Pendekatan modular ini memungkinkan pengembangan dan pemeliharaan sistem secara fleksibel, serta memastikan setiap fungsi pembelajaran dapat berjalan secara independen tanpa mengganggu stabilitas sistem secara keseluruhan.

2.3 Pengembangan Prototipe

Tahap pengembangan prototipe diawali dengan perakitan seluruh komponen perangkat keras sesuai dengan konfigurasi *wiring* yang telah dirancang pada tahap sebelumnya. Setelah integrasi perangkat keras selesai, pengembangan perangkat lunak dilakukan menggunakan struktur pemrograman modular untuk memastikan keterpisahan fungsi serta kemudahan pemeliharaan sistem.

Struktur perangkat lunak terdiri atas beberapa modul fungsional, yaitu system button module untuk pengelolaan input dari push button melalui GPIO, modul input/output yang menangani proses perekaman audio, tampilan teks pada LCD, transkripsi suara, serta sintesis suara dalam mode daring dan luring, serta utility module yang bertanggung jawab atas pembersihan hasil transkripsi (transcription cleaning), konversi format teks, ekstraksi respons model, dan pengelolaan konteks percakapan. Pendekatan modular ini diterapkan untuk menjaga konsistensi alur data mulai dari input suara hingga keluaran visual dan audio.

Selain modul inti, dikembangkan pula modul pengelolaan API client untuk integrasi dengan Ollama dan Google Cloud. Fitur pendukung seperti help menu sebagai panduan penggunaan, learning audio mode untuk memutar podcast dan musik berbahasa Inggris, idle animation, serta indikator status perangkat juga diintegrasikan ke dalam sistem utama tanpa mengganggu fungsi inti pembelajaran.

Setiap mode pada POCALA diprogram berdasarkan flowchart sistem yang telah disusun pada tahap perancangan. Pengembangan dimulai dari versi luring dengan memanfaatkan Whisper.cpp sebagai modul pengenalan suara, Piper TTS untuk sintesis suara, MarianMT untuk penerjemahan, serta Gemma 3 yang diakses melalui Ollama API sebagai Large Language Model lokal. Implementasi versi luring dilakukan terlebih dahulu untuk memastikan sistem dapat beroperasi secara mandiri tanpa ketergantungan jaringan.

Setelah seluruh fungsi pada mode luring berjalan stabil, dikembangkan versi daring menggunakan layanan Google Cloud API, meliputi Cloud Speech-to-Text, Cloud Text-to-Speech, Cloud Translation, serta Gemini sebagai Large Language Model. Integrasi versi daring dilakukan dengan mempertahankan struktur alur data yang sama, sehingga peralihan antara mode daring dan luring hanya memengaruhi sumber pemrosesan model, bukan logika sistem utama. Arsitektur dual-mode ini dirancang untuk memberikan fleksibilitas operasional sesuai kondisi konektivitas jaringan.

Pada kedua implementasi LLM, teknik prompt engineering diterapkan untuk mengarahkan keluaran model agar tetap relevan dengan konteks pembelajaran dan tidak menghasilkan respons yang bersifat terlalu umum atau tidak instruksional. Setelah seluruh modul terintegrasi dan diuji secara internal, program utama sistem dikonfigurasi menggunakan layanan systemd pada Raspberry Pi OS sehingga sistem dapat berjalan otomatis saat perangkat dinyalakan.

Proses fabrikasi casing dilakukan secara iteratif. Prototipe awal dicetak menggunakan filament PLA+ untuk mengoptimalkan tata letak komponen dan kesesuaian dimensi. Setelah desain mekanik dinilai sesuai, casing final diproduksi menggunakan filament ASA untuk meningkatkan ketahanan terhadap panas dan penggunaan jangka panjang. Setelah perakitan akhir serta verifikasi fungsional dilakukan, prototipe POCALA yang berfungsi penuh berhasil diselesaikan.

2.4 Pengujian dan Evaluasi

Pengujian sistem dilakukan dalam dua tahap untuk memastikan sistem berfungsi sesuai dengan tujuan perancangan, yaitu pengujian kinerja internal untuk mengukur performa teknis sistem serta evaluasi pengguna eksternal untuk menilai aspek fungsionalitas, kemudahan penggunaan, dan aksesibilitas dalam konteks penggunaan nyata.

2.4.1 Pengujian Kinerja Internal

Pengujian internal dilaksanakan oleh tim pengembang bersama dosen pembimbing untuk memverifikasi kesesuaian fungsi sistem terhadap spesifikasi desain. Pengujian ini mencakup pemeriksaan integrasi perangkat keras dan perangkat lunak, kestabilan sistem saat beroperasi pada mode daring dan luring, serta validasi fungsi setiap fitur pembelajaran.

Pada tahap ini dilakukan identifikasi kesalahan sistem (debugging), pengujian respons tombol fisik, validasi alur input-proses-output, serta pengamatan stabilitas operasional selama penggunaan berkelanjutan. Pengujian internal bertujuan memastikan bahwa seluruh modul berjalan sesuai perancangan sebelum dilakukan evaluasi kepada pengguna eksternal.

Selain verifikasi fungsional, dilakukan pengukuran kinerja teknis menggunakan beberapa parameter kuantitatif. Akurasi pengenalan suara diukur pada tingkat *utterance* berdasarkan keberhasilan transkripsi 50 kata dan kalimat yang diucapkan, yang diulang sebanyak tiga kali oleh pembicara berbeda untuk mengakomodasi variasi suara.

Kecepatan transkripsi diukur sebagai selang waktu sejak pelepasan tombol rekam hingga hasil transkripsi ditampilkan pada layar. Waktu respons Large Language Model (LLM) diukur sejak input pengguna selesai hingga respons teks ditampilkan pada sistem.

Daya tahan baterai dihitung berdasarkan rata-rata waktu yang dibutuhkan untuk setiap penurunan 1% kapasitas baterai dari total pengurangan 20%, dimulai dari kondisi penuh (100%) dengan seluruh fungsi sistem digunakan secara berkelanjutan. Suhu perangkat diukur selama proses intensif pengenalan suara dan pemrosesan LLM menggunakan *thermal gun*.

Kualitas keluaran suara dinilai secara kualitatif oleh tiga evaluator berdasarkan kejernihan dan *naturalness* audio pada kedua bahasa. Seluruh pengujian dilakukan di dalam ruangan pada suhu lingkungan normal dan koneksi internet yang stabil.

2.4.2 Evaluasi Pengguna Eksternal

Evaluasi eksternal melibatkan 20 responden dari tiga lokasi, yaitu SLB Negeri 3 Banjarmasin (termasuk pengguna tunanetra, pengguna dengan disabilitas fisik, serta guru), MAN 2 Banjarmasin (siswa dan guru), serta Universitas Islam Kalimantan Muhammad Arsyad Al Banjari (mahasiswa dan tenaga akademik). Evaluasi ini bertujuan untuk menilai fungsionalitas sistem, kemudahan penggunaan (usability), serta aspek aksesibilitas pada kelompok pengguna dengan latar belakang yang beragam.

Instrumen evaluasi menggunakan kuesioner skala Likert lima poin (Pranatawijaya et al., 2019). Kuesioner dirancang untuk mengukur enam aspek utama yang merepresentasikan dimensi teknis dan pengalaman pengguna terhadap perangkat, yaitu fungsionalitas umum, fungsionalitas mode, aksesibilitas perangkat keras, aksesibilitas perangkat lunak, kebermanfaatan umum dan kebermanfaatan tiap mode.

Aspek fungsionalitas umum terdiri dari 7 pertanyaan yang mengukur reliabilitas dan performa teknis sistem secara keseluruhan. Indikator pada aspek ini meliputi kecepatan respons sistem pada mode daring dan luring, kejernihan serta naturalness keluaran *text-to-speech*, akurasi *speech recognition*, daya tahan baterai, stabilitas sistem selama penggunaan, serta kejelasan umpan balik ketika input suara tidak berhasil dikenali.

Aspek fungsionalitas mode terdiri dari 10 pertanyaan untuk mengevaluasi performa masing-masing fitur pembelajaran yang tersedia pada sistem. Penilaian mencakup akurasi terjemahan pada Translator Mode,

dukungan kosakata dan penyajian informasi pada Vocabulary Mode, kejelasan dan ketepatan koreksi pada fitur Grammar, relevansi respons pada fitur Ask, kesesuaian serta akurasi evaluasi pada Question Mode, dan koherensi percakapan pada Speaking Mode.

Aspek aksesibilitas perangkat keras terdiri dari 6 pertanyaan yang berfokus pada dimensi fisik perangkat, termasuk ergonomi saat digenggam, portabilitas ukuran dan berat, kemudahan pengoperasian tombol, kejelasan tampilan layar, serta persepsi ketahanan dan keamanan perangkat dalam penggunaan sehari-hari.

Aspek aksesibilitas perangkat lunak terdiri dari 7 pertanyaan untuk mengevaluasi kemudahan interaksi pengguna dengan sistem. Indikator yang dinilai meliputi kejelasan instruksi suara, kemudahan navigasi antar menu dan mode, kemudahan penggunaan sub-mode pada Assistant Mode, serta tingkat kepercayaan diri pengguna dalam mengoperasikan sistem secara mandiri tanpa bantuan tambahan.

Aspek kebermanfaatan umum terdiri dari 6 pertanyaan yang mengukur persepsi kebermanfaatan sistem secara menyeluruh dalam mendukung pembelajaran bahasa Inggris. Penilaian mencakup peningkatan motivasi belajar, manfaat operasional mode luring, pengurangan distraksi dibandingkan aplikasi berbasis smartphone, serta nilai kebaruan pendekatan pembelajaran yang ditawarkan.

Aspek kebermanfaatan tiap mode terdiri dari 6 pertanyaan yang secara khusus menilai kebermanfaatan masing-masing mode pembelajaran dalam mendukung praktik mandiri, pemahaman materi, serta peningkatan kepercayaan diri pengguna dalam menggunakan bahasa Inggris.

Data kuantitatif dianalisis menggunakan statistik deskriptif berupa nilai rata-rata (mean) dan standar deviasi, yang umum digunakan untuk memberikan gambaran awal mengenai persepsi responden dalam penelitian kuantitatif (Field, 2018).

Interpretasi nilai rata-rata dilakukan dengan membagi rentang skala Likert ke dalam lima kategori proporsional, yaitu sangat rendah, rendah, sedang, tinggi, dan sangat tinggi. Lebar interval setiap kategori dihitung menggunakan pendekatan pembagian rentang proporsional sebagaimana disarankan oleh Sugiyono (2016). Perhitungan interval dinyatakan dalam Persamaan (1):

$$I = \frac{(S_{max} - S_{min})}{k} \quad (1)$$

Keterangan:

- **I** = lebar interval
- $\mathbf{S_{max}}$ = skor maksimum
- $\mathbf{S_{min}}$ = skor minimum
- **k** = jumlah kategori

Berdasarkan perhitungan tersebut, skala interpretasi nilai mean yang digunakan dalam penelitian ini disajikan pada Tabel 1.

Tabel 1. Interpretasi Skala Likert

Nilai Mean	Kategori Interpretasi
1.00 – 1.80	Sangat Buruk
1.81 – 2.60	Buruk
2.61 – 3.40	Netral / Sedang
3.41 – 4.20	Baik
4.21 – 5.00	Sangat Baik

Selain analisis kuantitatif, data kualitatif yang diperoleh melalui pertanyaan terbuka dianalisis secara deskriptif untuk mengidentifikasi persepsi pengguna, pengalaman penggunaan, serta masukan terhadap pengembangan sistem pada tahap selanjutnya.

HASIL DAN PEMBAHASAN

3.1 Gambaran Umum Sistem

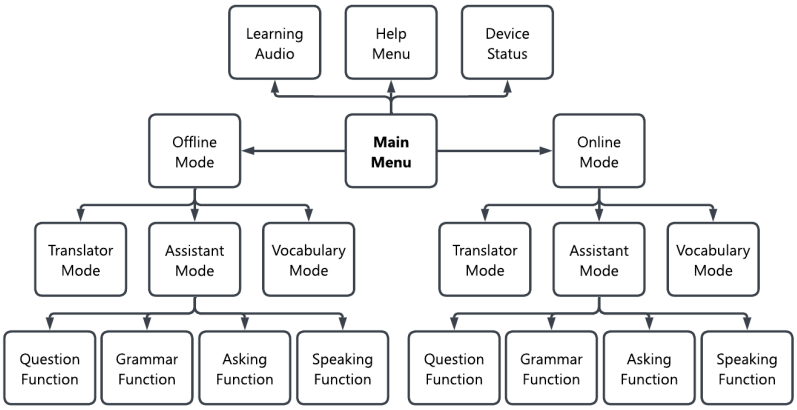
Penelitian ini menghasilkan sebuah prototipe fungsional asisten pembelajaran bahasa Inggris portabel berbasis kecerdasan buatan bernama POCALA (*Pocket a Language*). Prototipe ini dikembangkan melalui

serangkaian tahapan yang meliputi perancangan sistem, implementasi perangkat keras dan perangkat lunak, pengujian fungsional, serta evaluasi pengguna untuk menilai kinerja dan kebermanfaatan sistem yang dihasilkan.

POCALA dirancang untuk beroperasi dengan menerima input pengguna berupa perintah atau respons lisan melalui tombol perekaman khusus. Input suara tersebut kemudian diproses oleh sistem dan menghasilkan keluaran dalam dua bentuk, yaitu informasi teks yang ditampilkan pada layar serta respons audio yang disalurkan melalui *speaker*. Ketika perangkat dinyalakan, pengguna akan diarahkan ke menu utama yang terdiri atas beberapa pilihan mode, yaitu *online* mode, *offline* mode, *learning audio*, *help menu*, dan status.

Baik mode daring maupun luring menyediakan tiga fungsi pembelajaran utama dengan alur kerja yang serupa, yaitu *Translator*, *Vocabulary*, dan *Assistant*. Mode *Assistant* selanjutnya terdiri atas empat sub-fungsi pembelajaran, yaitu *Grammar*, *Asking*, *Question*, dan *Speaking*. Perbedaan utama antara mode daring dan luring terletak pada mekanisme pemrosesan data. Mode daring memanfaatkan layanan berbasis cloud yang memungkinkan waktu respons lebih cepat, kualitas keluaran suara yang lebih natural, serta cakupan pengetahuan yang lebih luas. Sebaliknya, mode luring menggunakan model yang dijalankan secara lokal pada perangkat, sehingga memiliki kecepatan transcription yang sedikit lebih lambat dan kualitas keluaran suara yang relatif lebih rendah. Meskipun demikian, mode luring tetap mampu menghasilkan keluaran fungsional yang sebanding tanpa ketergantungan pada koneksi internet.

Selain fungsi pembelajaran utama, sistem juga dilengkapi dengan beberapa fitur tambahan. Fitur *Learning Audio* memungkinkan pengguna memutar materi pembelajaran bahasa Inggris secara luring, seperti *podcast* atau konten audio edukatif. *Help Menu* menyediakan panduan penggunaan untuk setiap mode pembelajaran, sedangkan fitur Status menampilkan informasi perangkat seperti tingkat volume, konektivitas Wi-Fi, serta perangkat Bluetooth (headset) yang sedang terhubung. Struktur menu sistem secara keseluruhan ditunjukkan pada Gambar 4.

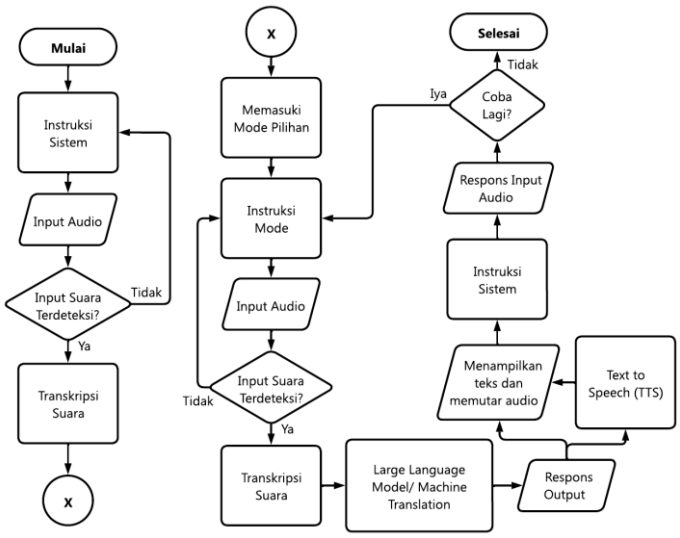


Gambar 4. Diagram struktur menu

Navigasi sistem sebagian besar dilakukan melalui interaksi suara. Pengguna dapat memberikan respons dalam bahasa Indonesia maupun bahasa Inggris sesuai dengan instruksi yang diberikan oleh sistem. Apabila input suara tidak dapat dikenali dengan jelas, sistem akan memberikan umpan balik yang meminta pengguna untuk mengulangi proses perekaman. Setelah suatu tugas atau aktivitas pembelajaran selesai, pengguna diberikan pilihan untuk melanjutkan atau keluar dari mode yang sedang digunakan. Selain itu, tersedia perintah suara tambahan "*exit*" yang memungkinkan pengguna menghentikan mode secara langsung. Desain navigasi ini menekankan kesederhanaan serta penggunaan pola instruksi yang konsisten dan berulang untuk memudahkan adaptasi pengguna, terutama bagi pengguna yang baru pertama kali menggunakan perangkat.

Dari perspektif teknis, POCALA mengintegrasikan komponen input, pemrosesan, dan output dalam satu pipeline sistem yang terpadu. Input suara pengguna ditangkap melalui mikrofon dan dikonversi menjadi teks menggunakan modul *speech recognition*, yaitu Whisper.cpp pada mode luring dan Google Speech-to-Text pada mode daring. Teks hasil transcription kemudian diproses oleh *Large Language Model* (LLM), yaitu Gemma 3 melalui Ollama untuk operasi lokal atau Gemini AI untuk pemrosesan berbasis cloud, guna menghasilkan respons yang sesuai dengan konteks percakapan atau pembelajaran.

Respons teks yang dihasilkan kemudian ditampilkan pada layar perangkat dan dikonversi menjadi suara menggunakan modul *text-to-speech*, yaitu Piper TTS pada mode luring dan Google Text-to-Speech pada mode daring. Seluruh komponen sistem dijalankan pada sistem operasi Raspberry Pi OS dengan arsitektur pemrograman modular berbasis Python. Alur kerja sistem secara umum, mulai dari pemilihan mode hingga sistem menjalankan fungsi pembelajaran, ditunjukkan pada Gambar 5.



Gambar 5. Alur kerja sistem secara umum

3.2 Spesifikasi Fisik dan Desain Perangkat

Secara fisik, prototipe POCALA terdiri atas dua unit utama, yaitu unit perangkat (*device unit*) dan unit catu daya (*power supply unit*). Unit perangkat berfungsi sebagai modul pemrosesan utama yang mengintegrasikan komponen komputasi dan antarmuka pengguna. Modul ini ditenagai oleh sebuah Raspberry Pi yang dilengkapi dengan layar LCD berukuran 1,3 inci, mikrofon sebagai perangkat input suara, speaker sebagai perangkat keluaran audio, serta empat tombol fungsional yang digunakan untuk mengontrol operasi sistem. Daya untuk unit perangkat dapat disuplai secara langsung melalui adaptor Type-C atau melalui unit catu daya menggunakan koneksi Type-C.

Unit catu daya dirancang untuk mendukung operasional perangkat secara portabel dengan mengintegrasikan baterai *rechargeable* berkapasitas 10.000 mAh serta rangkaian manajemen daya tambahan. Selain berfungsi sebagai sumber daya utama bagi unit perangkat, unit catu daya juga dapat digunakan sebagai *power bank* tambahan untuk mengisi daya perangkat eksternal apabila diperlukan.

Dari segi dimensi, unit perangkat memiliki ukuran 130×75×25 mm dengan berat sekitar 179 gram, sedangkan unit catu daya memiliki dimensi yang sama dengan berat sekitar 211 gram. Ketika kedua unit tersebut digabungkan, dimensi keseluruhan perangkat menjadi 130×75×50 mm dengan berat total sekitar 390 gram. Kedua unit dihubungkan menggunakan mekanisme *magnetic locking hinge* yang memungkinkan proses pemasangan dan pelepasan dilakukan secara mudah sekaligus menjaga kestabilan struktur perangkat saat digunakan.

Casing perangkat dibuat menggunakan material ASA (*Acrylonitrile Styrene Acrylate*), yang dipilih karena memiliki ketahanan terhadap suhu tinggi, daya tahan terhadap benturan, serta perlindungan terhadap paparan sinar ultraviolet. Material ini juga mendukung penggunaan perangkat dalam berbagai kondisi lingkungan tanpa mengalami deformasi yang signifikan. Selain itu, desain perangkat dilengkapi dengan *lanyard attachment* yang memungkinkan perangkat digantung atau dibawa dengan lebih mudah sehingga meningkatkan aspek portabilitas dalam penggunaan sehari-hari. Tampilan prototipe perangkat yang telah selesai dirakit ditunjukkan pada Gambar 6.



Gambar 6. Prototipe fungsional perangkat

3.3 Hasil Pengujian Fungsional

Pengujian fungsional dilakukan untuk mengevaluasi kinerja sistem POCALA pada dua konfigurasi operasional, yaitu mode online dan mode offline. Evaluasi ini mencakup beberapa parameter utama yang berkaitan dengan performa pemrosesan suara, respons sistem, efisiensi energi, serta stabilitas operasional perangkat.

Parameter yang diamati meliputi akurasi *speech recognition*, kecepatan transkripsi suara, waktu respons *Large Language Model* (LLM), daya tahan baterai, kualitas *voice output*, keberhasilan fungsi tombol, serta suhu perangkat selama operasi. Hasil pengujian untuk kedua mode operasional disajikan pada Tabel 2.

Tabel 2. Hasil pengujian fungsional

Parameter	Mode Online	Mode Offline
Akurasi pengenalan suara	96%	84%
Kecepatan transkripsi suara	1-3 detik	2-5 detik
Kecepatan respons *large language model*	1-5 detik	5-15 detik
Daya tahan baterai	±12 jam, ±7 menit/1%	±5 jam, ±3 menit/1%
Kualitas *voice output*	Natural dan jelas pada kedua bahasa	Sedikit kurang jelas pada bahasa Indonesia
Keberhasilan fungsi tombol	100%	100%
Suhu perangkat	30-40°C (normal)	40-55°C (terasa lebih hangat saat penggunaan lama)

Berdasarkan hasil pada Tabel 2, mode online secara konsisten menunjukkan performa yang lebih tinggi pada parameter pemrosesan suara dan respons sistem. Akurasi *speech recognition* pada mode online mencapai 96%, sedangkan pada mode offline sebesar 84%. Perbedaan ini dapat dikaitkan dengan penggunaan layanan pemrosesan berbasis *cloud* pada mode online yang memiliki kapasitas komputasi dan model yang lebih besar dibandingkan dengan model yang dijalankan secara lokal.

Perbedaan performa juga terlihat pada kecepatan transkripsi suara dan waktu respons LLM. Mode online mampu menghasilkan transkripsi suara dalam waktu sekitar 1-3 detik dan respons LLM dalam 1-5 detik. Sebaliknya, mode offline memerlukan waktu yang lebih lama, yaitu sekitar 2-5 detik untuk transkripsi dan 5-15 detik untuk respons LLM. Hal ini mencerminkan keterbatasan sumber daya komputasi pada perangkat tertanam yang menjalankan model secara lokal.

Dari sisi efisiensi energi, daya tahan baterai pada mode online tercatat lebih tinggi dibandingkan mode offline. Kondisi ini terjadi karena sebagian besar proses komputasi pada mode online dialihkan ke layanan *cloud*,

sehingga beban pemrosesan lokal menjadi lebih ringan. Sebaliknya, pada mode offline seluruh proses *speech recognition* dan pemrosesan LLM dilakukan langsung pada perangkat (*on-device computation*), yang meningkatkan konsumsi daya.

Pengujian terhadap fungsi tombol menunjukkan tingkat keberhasilan 100% pada kedua mode operasional, yang mengindikasikan bahwa mekanisme interaksi fisik perangkat berjalan secara stabil dan konsisten selama pengujian. Kualitas *voice output* pada mode online dinilai natural dan jelas pada kedua bahasa, sedangkan pada mode offline kualitas keluaran suara untuk bahasa Indonesia dinilai sedikit kurang jelas dibandingkan dengan mode online.

Dari sisi stabilitas perangkat, suhu operasional tetap berada dalam rentang yang dapat diterima. Mode online menunjukkan suhu operasi sekitar 30-40°C, sedangkan mode offline dapat mencapai 40-55°C selama penggunaan intensif. Peningkatan suhu pada mode offline disebabkan oleh aktivitas komputasi lokal yang lebih tinggi, khususnya selama proses *speech recognition* dan inferensi LLM.

Secara keseluruhan, hasil pengujian menunjukkan adanya *trade-off* antara performa sistem dan independensi operasional. Mode online menawarkan kecepatan pemrosesan dan kualitas keluaran yang lebih baik, sementara mode offline memberikan fleksibilitas penggunaan tanpa ketergantungan pada koneksi internet.

3.4 Hasil Evaluasi Pengguna

Evaluasi pengguna eksternal melibatkan 20 responden yang terdiri atas 1 guru dan 5 siswa MAN 2 Banjarmasin, 9 mahasiswa Universitas Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin, serta 1 guru pendamping dan 4 individu dengan disabilitas dari SLB Negeri 3 Banjarmasin. Evaluasi ini dilakukan untuk menilai persepsi pengguna terhadap fungsionalitas sistem, aksesibilitas, serta kebermanfaatannya dalam konteks pembelajaran bahasa Inggris.

Hasil evaluasi kuantitatif dirangkum pada Tabel 3, yang menampilkan nilai rata-rata untuk setiap aspek penilaian. Evaluasi mencakup enam aspek utama, yaitu fungsionalitas umum, fungsionalitas mode, aksesibilitas perangkat keras, aksesibilitas perangkat lunak, kebermanfaatan umum, dan kebermanfaatan mode. Penjelasan mengenai indikator dan butir pertanyaan pada masing-masing aspek telah dijelaskan pada Bagian 2.4.

Tabel 3. Skor evaluasi pengguna

Aspek Penilaian	Rata-rata Skor	Kategori
Fungsionalitas Umum	4.51	Sangat Baik
Fungsionalitas Mode	4.73	Sangat Baik
Aksesibilitas Perangkat Keras	4.18	Baik
Aksesibilitas Perangkat Lunak	4.60	Sangat Baik
Kebermanfaatan Umum	4.73	Sangat Baik
Kebermanfaatan Mode	4.84	Sangat Baik

Berdasarkan nilai rata-rata yang diperoleh, seluruh aspek evaluasi menunjukkan skor di atas 4,0, yang berada pada kategori baik hingga sangat baik menurut interpretasi skala Likert pada Tabel 1 di Bagian 2.4. Hal ini menunjukkan bahwa secara umum pengguna menilai sistem POCALA memiliki tingkat fungsionalitas, kemudahan penggunaan, serta kebermanfaatan yang tinggi dalam mendukung proses pembelajaran bahasa Inggris.

Nilai rata-rata tertinggi diperoleh pada aspek Kebermanfaatan Mode (4,84) dan Fungsionalitas Mode (4,73). Hasil ini menunjukkan bahwa fitur-fitur pembelajaran yang disediakan oleh sistem, seperti *Vocabulary*, *Translator*, *Grammar*, *Ask*, *Question*, dan *Speaking*, dinilai mampu mendukung proses belajar secara interaktif dan membantu pengguna dalam memahami serta mempraktikkan bahasa Inggris. Tingginya skor pada aspek ini mengindikasikan bahwa desain fitur pembelajaran yang terintegrasi dalam sistem mampu memberikan pengalaman belajar yang efektif dan mudah digunakan.

Aspek Aksesibilitas Perangkat Lunak juga memperoleh skor tinggi (4,60), yang menunjukkan bahwa pengguna menilai interaksi dengan sistem relatif mudah dipahami dan dioperasikan. Hal ini mencerminkan

bahwa instruksi suara, navigasi menu, serta alur penggunaan sistem telah dirancang secara cukup intuitif sehingga dapat diikuti oleh berbagai kelompok pengguna.

Sementara itu, aspek Aksesibilitas Perangkat Keras memperoleh nilai rata-rata 4,18, yang masih berada pada kategori baik namun sedikit lebih rendah dibandingkan aspek lainnya. Hal ini menunjukkan bahwa meskipun perangkat secara umum dinilai mudah digunakan, masih terdapat beberapa aspek fisik perangkat yang berpotensi untuk ditingkatkan pada pengembangan berikutnya, khususnya terkait dimensi perangkat dan kenyamanan interaksi dengan tombol fisik.

Selain evaluasi kuantitatif, umpan balik kualitatif juga dikumpulkan melalui pertanyaan terbuka dalam kuesioner untuk memperoleh gambaran yang lebih mendalam mengenai pengalaman penggunaan sistem. Ringkasan respons tersebut disajikan pada Tabel 4.

Tabel 4. Ringkasan respons pertanyaan terbuka

Pertanyaan	Kata Kunci Utama
Fitur Favorit	Semua fitur, Speaking, Translator, Question
Fitur Kurang Disukai	Tidak ada, Ukuran
Situasi Penggunaan Paling Bermanfaat	Belajar, Waktu santai
Fitur Tambahan	Lebih banyak musik, Headset
Perbaikan Hardware	Ukuran, Kekerasan tombol rekam, Volume speaker
Perbaikan Sistem	Tidak ada, Instruksi dalam Bahasa Indonesia

Berdasarkan tanggapan terbuka yang diberikan responden, fitur *speaking* sering disebut sebagai salah satu fitur yang paling menarik dan bermanfaat. Hal ini menunjukkan bahwa interaksi berbasis suara yang disediakan oleh sistem mampu mendorong pengguna untuk lebih aktif mempraktikkan kemampuan berbicara dalam bahasa Inggris.

Beberapa responden juga memberikan masukan terkait ukuran perangkat serta tingkat kekerasan tombol rekam. Masukan ini menunjukkan bahwa meskipun sistem telah diterima dengan baik secara fungsional, penyempurnaan pada aspek ergonomi perangkat masih dapat dilakukan untuk meningkatkan kenyamanan penggunaan dalam jangka panjang. Secara keseluruhan, hasil evaluasi pengguna menunjukkan bahwa sistem POCALA memiliki tingkat penerimaan yang tinggi serta berpotensi menjadi alat pendukung pembelajaran bahasa Inggris yang praktis dan mudah diakses oleh berbagai kelompok pengguna.

3.5 Pembahasan

Hasil pengujian fungsional dan evaluasi pengguna menunjukkan bahwa sistem POCALA mampu menyediakan lingkungan pembelajaran bahasa Inggris yang portabel, interaktif, dan inklusif melalui integrasi berbagai teknologi berbasis kecerdasan buatan. Integrasi *speech recognition*, *speech synthesis*, dan *Large Language Model* (LLM) memungkinkan sistem memproses input lisan pengguna serta menghasilkan respons yang kontekstual dalam bentuk teks dan audio, sehingga mendukung praktik bahasa secara lebih natural dan responsif.

Performa yang lebih tinggi pada mode online mencerminkan keunggulan layanan AI berbasis *cloud*, yang umumnya menyediakan kapasitas komputasi lebih besar serta model yang lebih optimal untuk pemrosesan bahasa alami. Kecepatan pemrosesan yang lebih tinggi dan kualitas keluaran suara yang lebih natural pada mode online menunjukkan bagaimana pemanfaatan layanan berbasis *cloud* dapat meningkatkan kualitas interaksi pengguna dengan sistem pembelajaran berbasis AI. Sebaliknya, mode offline menawarkan keuntungan praktis berupa independensi dari koneksi internet serta kemampuan operasional pada lingkungan dengan keterbatasan infrastruktur jaringan. Kemampuan dual-mode ini menunjukkan bahwa sistem dapat beradaptasi dengan berbagai kondisi penggunaan, baik pada lingkungan dengan konektivitas stabil maupun pada situasi dengan keterbatasan akses internet.

Hasil evaluasi pengguna juga memperlihatkan tingkat penerimaan yang tinggi terhadap sistem yang dikembangkan. Nilai rata-rata di atas 4,0 pada seluruh aspek evaluasi menunjukkan bahwa pengguna menilai sistem memiliki tingkat fungsionalitas, kemudahan penggunaan, serta kebermanfaatan yang baik dalam

mendukung proses pembelajaran bahasa Inggris. Skor tinggi pada aspek Kebermanfaatan Mode dan Fungsionalitas Mode menunjukkan bahwa fitur pembelajaran seperti *Translator, Vocabulary, Grammar, Ask, Question,* dan *Speaking* mampu memberikan pengalaman belajar yang interaktif serta membantu pengguna dalam memahami dan mempraktikkan bahasa Inggris secara mandiri.

Penerimaan positif ini dapat dikaitkan dengan desain sistem yang mengintegrasikan berbagai komponen AI yang disesuaikan dengan fungsi pembelajaran tertentu, alih-alih mengandalkan satu model interaksi generik. Pendekatan ini memungkinkan setiap fitur pembelajaran memberikan respons yang lebih terarah sesuai konteks penggunaan. Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa sistem pembelajaran bahasa berbasis AI dapat meningkatkan keterlibatan dan motivasi belajar melalui umpan balik adaptif serta interaksi dialogis dengan pengguna (Hidayatullah, 2024; Patty, 2024). Selain itu, penelitian terdahulu juga melaporkan bahwa penggunaan *machine translation* dapat membantu pengembangan kosakata dan pemahaman teks ketika digunakan sebagai alat bantu pembelajaran tambahan, meskipun tetap memiliki keterbatasan dalam memahami konteks dan nuansa semantik bahasa (Abimbola, 2023; Naveen & Trojovský, 2024).

Integrasi *Large Language Model* pada *Assistant Mode* juga memungkinkan interaksi yang lebih natural selama proses pembelajaran. LLM memungkinkan sistem memberikan respons berbasis percakapan yang tidak hanya memberikan jawaban, tetapi juga dapat membimbing pengguna melalui penjelasan, pertanyaan lanjutan, serta dialog yang mendorong eksplorasi bahasa. Temuan ini konsisten dengan penelitian sebelumnya yang menunjukkan bahwa LLM memiliki potensi sebagai asisten pembelajaran percakapan yang dapat mendukung interaksi terpersonalisasi, bimbingan kontekstual, serta umpan balik adaptif dalam lingkungan pendidikan (Dong et al., 2024; Gan et al., 2023).

Namun demikian, literatur juga menekankan bahwa efektivitas LLM dalam konteks pendidikan sangat bergantung pada kontrol pedagogis yang tepat. Penggunaan model generatif tanpa pembatasan dapat menurunkan kualitas pembelajaran dengan mendorong konsumsi jawaban secara pasif serta meningkatkan risiko keluaran yang tidak akurat atau bias. Oleh karena itu, kontrol berbasis *prompt* dan desain interaksi yang terarah menjadi penting untuk memastikan bahwa respons LLM tetap bersifat instruksional dan selaras dengan tujuan pembelajaran (Chen et al., 2023; Pester et al., 2024). Dalam penelitian ini, strategi *prompt engineering* diterapkan untuk mengarahkan keluaran LLM agar tetap berorientasi pada penjelasan, latihan, dan dukungan pembelajaran, sehingga interaksi yang dihasilkan tetap memiliki nilai pedagogis.

Selain aspek interaksi AI, implementasi sistem sebagai perangkat keras khusus juga memberikan kontribusi penting terhadap pengalaman belajar. Berbeda dengan aplikasi pembelajaran berbasis *smartphone* yang rentan terhadap distraksi digital, perangkat khusus memungkinkan pengguna berinteraksi dengan sistem dalam lingkungan yang lebih terfokus. Selain itu, kemampuan sistem untuk beroperasi secara offline membantu mengatasi keterbatasan konektivitas yang masih menjadi tantangan di banyak lingkungan pembelajaran. Tantangan serupa juga dilaporkan pada penelitian sebelumnya yang mengembangkan prototipe pembelajaran berbasis GPT pada platform Raspberry Pi, di mana sistem percakapan AI mampu meningkatkan keterlibatan siswa tetapi masih sangat bergantung pada layanan API berbasis internet sehingga membatasi kemampuan operasi mandiri perangkat (Yuly et al., 2023). Tantangan tersebut juga diidentifikasi sebagai hambatan utama dalam berbagai konteks pembelajaran, khususnya pada lingkungan dengan ketersediaan jaringan yang terbatas serta paparan tinggi terhadap gangguan digital non-pembelajaran (Y. Ahmad et al., 2025).

Dari perspektif aksesibilitas, penggunaan interaksi berbasis suara dan navigasi yang dipandu oleh audio memungkinkan sistem digunakan oleh pengguna dengan keterbatasan penglihatan maupun mobilitas. Pendekatan ini memperkuat orientasi desain inklusif pada sistem POCALA. Penelitian sebelumnya mengenai perangkat asistif berbasis suara juga menunjukkan bahwa integrasi teknologi *speech recognition* dan *speech-to-text* dapat meningkatkan aksesibilitas bagi pengguna dengan disabilitas dengan mengurangi ketergantungan pada antarmuka visual atau manual (Tharwat et al., 2024).

Meskipun demikian, beberapa keterbatasan masih ditemukan dalam implementasi sistem. Kualitas *speech synthesis* pada mode offline masih kurang natural dibandingkan dengan mode online, waktu respons sistem sedikit lebih lama, dan bobot perangkat masih dapat dioptimalkan untuk meningkatkan portabilitas. Keterbatasan ini terutama disebabkan oleh penggunaan model yang dijalankan secara lokal serta karakteristik perangkat keras prototipe tahap awal. Penurunan kecepatan dan akurasi *transcription* pada mode offline juga

sejalan dengan penelitian sebelumnya mengenai implementasi model *speech recognition* berbasis Whisper pada sistem *embedded*. Studi tersebut menunjukkan bahwa keterbatasan sumber daya komputasi pada perangkat seperti Raspberry Pi dapat menyebabkan peningkatan *word error rate* serta waktu pemrosesan yang lebih lama, khususnya pada variasi aksen dan intonasi, meskipun masih mempertahankan tingkat responsivitas yang dapat diterima untuk aplikasi *embedded* (Wafiy & Prasetio, 2024).

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa POCALA menawarkan pendekatan pembelajaran bahasa yang lebih terfokus, bebas distraksi, dan aksesibel dibandingkan aplikasi pembelajaran berbasis smartphone konvensional. Dengan mengombinasikan perangkat keras khusus dan interaksi AI multimodal, sistem ini mampu menyediakan pengalaman belajar yang lebih imersif dan berpusat pada pengguna tanpa ketergantungan pada koneksi internet secara terus-menerus. Pengembangan di masa depan dapat mencakup sinkronisasi *cloud* untuk pelacakan progres pembelajaran, dukungan multi-bahasa, serta optimasi lebih lanjut terhadap model AI lokal guna meningkatkan performa dan efisiensi sistem.

KESIMPULAN DAN SARAN

Penelitian ini berhasil merancang dan mengembangkan asisten pembelajaran bahasa Inggris portabel berbasis kecerdasan buatan bernama POCALA *(Pocket a Language*) yang mengintegrasikan teknologi *Large Language Model*, *speech recognition*, dan *speech synthesis* melalui pendekatan perancangan sistem dan *prototyping*. Sistem yang dikembangkan mampu mendukung proses pembelajaran bahasa Inggris yang interaktif, inklusif, dan adaptif melalui interaksi audio-visual multimodal yang dapat beroperasi baik pada mode online maupun offline.

Hasil pengujian fungsional menunjukkan bahwa seluruh komponen sistem dapat beroperasi secara andal sesuai dengan rancangan yang telah ditetapkan. Selain itu, hasil evaluasi pengguna menunjukkan tingkat penerimaan yang tinggi pada berbagai kelompok pengguna, termasuk individu dengan disabilitas. Penerapan arsitektur dual-mode memungkinkan sistem tetap dapat digunakan secara fleksibel pada lingkungan dengan keterbatasan konektivitas internet sekaligus membantu mengurangi distraksi teknologi yang umum ditemukan pada aplikasi pembelajaran berbasis smartphone. Temuan ini menunjukkan bahwa integrasi multimodal kecerdasan buatan dalam perangkat embedded portabel merupakan pendekatan yang layak untuk mendukung pembelajaran bahasa yang lebih terfokus, adaptif, dan aksesibel.

Meskipun demikian, beberapa keterbatasan masih ditemukan, antara lain kualitas *speech synthesis* pada mode offline yang relatif lebih rendah serta waktu respons yang sedikit lebih lambat dibandingkan mode online, serta bobot prototipe yang masih cukup besar. Oleh karena itu, pengembangan selanjutnya dapat difokuskan pada optimasi model AI lokal, peningkatan ergonomi perangkat keras, serta implementasi fitur tambahan seperti sinkronisasi progres pembelajaran dan dukungan multi-bahasa guna meningkatkan performa sistem dan kemudahan penggunaan dalam konteks pembelajaran bahasa yang lebih luas.

DAFTAR PUSTAKA

Abimbola, L. I. (2023). The Impact of Machine Translation on Vocabulary Acquisition and Reading Comprehension in ESL Learners. *Pulchra Lingua: A Journal of Language Study, Literature & Linguistics*, *2*(2), 80–93. https://doi.org/10.58989/plj.v2i2.27

Ahmad, S. (2023). Analisis Kesulitan Siswa dalam Berbicara Bahasa Inggris di MAS Mulia Sei Balai. *JIP: Jurnal Ilmu Pendidikan*, *1*(6), Article 6.

Ahmad, Y., Hussain, S., Mumtaz, U., & Riaz, S. (2025). Investigating the Influence of Duolingo, Elsa Speak, and Hello English Mobile Applications on Students Attitudes and Speaking Skills. *Review Journal of Social Psychology & Social Works*, *3*(1), Article 1.

Chen, Y., Chen, H., & Su, S. (2023). Fine-Tuning Large Language Models in Education. *2023 13th International Conference on Information Technology in Medicine and Education (ITME)*, 718–723. https://doi.org/10.1109/ITME60234.2023.00148

Dong, B., Bai, J., Xu, T., & Zhou, Y. (2024). Large Language Models in Education: A Systematic Review. *2024 6th International Conference on Computer Science and Technologies in Education (CSTE)*, 131–134. https://doi.org/10.1109/CSTE62025.2024.00031

EF Education First. (2024). *EF English Proficiency Index: A Ranking of 116 Countries and Regions by English Skills*. https://www.ef.com/wwen/epi/

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics (5th ed.)* (5th ed.). SAGE Publications.

Gan, W., Qi, Z., Wu, J., & Lin, J. C.-W. (2023). *Large Language Models in Education: Vision and Opportunities* (arXiv:2311.13160). arXiv. https://doi.org/10.48550/arXiv.2311.13160

Hidayah, W. N. (2021). Konsep Solusi Terhadap Problem Keterampilan (Skills) Berbahasa Inggris di Lembaga Pendidikan Indonesia. *Jurnal Pendidikan Indonesia*, *2*(10), 1824–1834. https://doi.org/10.59141/japendi.v2i10.314

Hidayatullah, R. (2024). Implementasi AI dalam Proses Pembelajaran pada Mahasiswa Semster Awal Pendidikan Bahasa Inggris. *Assyfa Journal of Multidisciplinary Education*, *2*(1), 13–18. https://doi.org/10.61650/ajme.v2i2.631

Jiang, M. Y.-C., Jong, M. S.-Y., Wu, N., Shen, B., Chai, C.-S., Lau, W. W.-F., & Huang, B. (2022). Integrating Automatic Speech Recognition Technology Into Vocabulary Learning in a Flipped English Class for Chinese College Students. *Frontiers in Psychology*, *13*, 902429. https://doi.org/10.3389/fpsyg.2022.902429

Kirchhoff, P. (2024). Machine Translation in English Language Teaching. *ELT Journal*, *78*(4), 393–400. https://doi.org/10.1093/elt/ccae034

Manu, G. A., & Masan, P. L. (2020). Aplikasi Text to Speech untuk Meningkatkan Pembelajaran Bahasa Inggris bagi Siswa Disabilitas. *Jurnal Pendidikan Teknologi Informasi (JUKANTI)*, *3*(2), Article 2. https://doi.org/10.37792/jukanti.v3i2.217

Marlinda, S., & Huda, N. N. (2024). Improving Pronunciation Skills with ELSA Speak: An AI-Based Learning. *Jurnal Intelek Insan Cendikia*, *1*(6), Article 6.

Naveen, P., & Trojovský, P. (2024). Overview and Challenges of Machine Translation for Contextually Appropriate Translations. *iScience*, *27*(10), 110878. https://doi.org/10.1016/j.isci.2024.110878

Patty, J. (2024). The Use of AI in Language Learning: What You Need to Know. *Jurnal Review Pendidikan Dan Pengajaran (JRPP)*, *7*(1), 642–654. https://doi.org/10.31004/jrpp.v7i1.24609

Pester, A., Tammaa, A., Gütl, C., Steinmaurer, A., & El-Seoud, S. A. (2024). Conversational Agents, Virtual Worlds, and Beyond: A Review of Large Language Models Enabling Immersive Learning. *2024 IEEE Global Engineering Education Conference (EDUCON)*, 1–6. https://doi.org/10.1109/EDUCON60312.2024.10578895

Pranatawijaya, V. H., Widiatry, W., Priskila, R., & Putra, P. B. A. A. (2019). Penerapan Skala Likert dan Skala Dikotomi Pada Kuesioner Online. *Jurnal Sains Dan Informatika*, *5*(2), 128–137. https://doi.org/10.34128/jsi.v5i2.185

Pranoto, Y. H., & Inharjanto, A. (2020). Merevitalisasi Penggunaan Bahasa Inggris pada Satuan Pendidikan. *Jurnal Humanipreneur*, *1*(1), Article 1. https://doi.org/10.53091/jhup.v1i1.7

Saragih, M., Sipayung, K., & Manurung, L. W. (2023). The Effect of Using Cake and ELSA Speak Application on Students' Pronunciation Skills at the Eighth Grade of SMP Negeri 25 Medan. *Journey: Journal of English Language and Pedagogy*, *6*(3), Article 3. https://doi.org/10.33503/journey.v6i3.696

Sugiyono. (2016). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta.

Tharwat, M., Wardak, Y., Balbaid, S., & Radin, E. (2024). Wearable Device With Speech and Voice Recognition for Hearing-Impaired People. *2024 21st Learning and Technology Conference (L&T)*, 221–226. https://doi.org/10.1109/LT60077.2024.10469567

Wafiy, A. D., & Prasetio, B. H. (2024). Penerapan Model Whisper Pada Embedded System untuk Speech To Text. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *8*(8), Article 8. https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/14008

Widyana, A., Jerusalem, M. I., & Yumechas, B. (2022). *The Application of Text-to-Speech Technology in Language Learning*. 85–92. https://doi.org/10.2991/978-2-494069-91-6_14

Yuly, A. R., Pradana, H., & Nurcahyo, W. H. (2023). Designing Talk to GPT Prototype to Support Learning in Vocational Schools. *Southeast Asian Journal on Open and Distance Learning*, *1*(02), Article 02.

Zulfa, N. A., & Mujazi, M. (2022). Pengaruh Penggunaan Smartphone terhadap Konsentrasi Belajar Siswa. *JRTI (Jurnal Riset Tindakan Indonesia)*, *7*(3), Article 3. https://doi.org/10.29210/30032126000