

---

---

## Desain dan Implementasi Speech Recognition Sebagai Media Pembelajaran Pronunciation Bahasa Inggris

Irfan Hadiatma, Hanung Adi Nugroho, Eko Nugroho  
Departemen Teknik Elektro dan Teknologi Informasi  
Universitas Gajah Mada Yogyakarta, Indonesia

hadiatma.cio14@mail.ugm.ac.id; adinugroho@ugm.ac.id; nugroho@mail.ugm.ac.id

**Abstrak**— Kemampuan teknis dan komputasi perangkat mobile yang semakin berkembang berbanding lurus dengan teknologi pengolahan data, salah satunya machine learning. Di bidang teks atau information retrieval terdapat speech recognition, yaitu kemampuan untuk mengolah suara menjadi teks. Dalam perangkat mobile teknologi tersebut terakomodasi melalui layanan API cloud speech yang merupakan pengembangan model pre-trained learning machine dari google. Komunikasi antara individu secara verbal dalam bahasa inggris memerlukan pronunciation yang baik. Selama ini proses tersebut untuk evaluasinya masih dilakukan secara manual yaitu dengan saling memberikan feed back antara individu, sehingga dinilai kurang efisien. Penelitian ini memberikan rancangan dan konsep pemanfaatan teknologi tersebut sebagai media pembelajaran pronunciation bahasa inggris secara mandiri dan mobile dimana parameter yang dibandingkan adalah tingkat penekanan suara yang kemudian dicocokkan dengan template database yang ada. Rancangan aplikasi berbasis android mobile dengan menggunakan metode SDLC dengan UML Diagram, yaitu Use Case diagram, Activity dan Sequence Diagram sebagai bahasa pemodelannya. Hasil pengujian aplikasi menunjukkan bahwa aplikasi dapat digunakan sebagai media pembelajaran bahasa inggris.

**Kata Kunci**— *Android, Aplikasi Pembelajaran, Pronunciation, SDLC, Voice Recognition*

### I. PENDAHULUAN

Komunikasi antara individu secara verbal dalam bahasa inggris memerlukan pronunciation yang baik. Bagi pemula yang belajar bahasa inggris secara mandiri tentunya akan mengalami banyak hambatan dibandingkan belajar di lembaga pendidikan bahasa inggris yang formal maupun informal. Perkembangan teknologi mobile yang semakin pesat tentunya memberikan dampak yang baik untuk menerapkan aplikasi pembelajaran bahasa inggris. Ditambah lagi fitur multimedia yang terdapat dalam teknologi mobile dimana user dapat berinteraksi secara audio visual dengan perangkat mobile tersebut.

Android sebagai sistem operasi yang dapat berjalan di mobile device sekarang ini terus berkembang. Hal tersebut dikarenakan kita sebagai pengembang memiliki hak akses penuh terhadap framework API yang sama digunakan oleh aplikasi inti, atau biasa disebut opensource[1]. Semakin kedepan fitur - fitur yang di kembangkan juga semakin canggih dimana salah satunya voice recognition. Voice recognition merupakan sistem yang dapat mengidentifikasi suara[2].

Dalam proses belajar bahasa inggris tentunya ada beberapa aspek yang dipelajari antara lain reading, writing, listening, dan speaking[3]. Pada beberapa aspek diatas, penelitian ini akan membahas tentang speaking. Dimana media pembelajaran ini menitik beratkan bagaimana user dapat belajar mengucapkan kata dan kalimat dalam bahasa inggris dengan benar.

Perancangan sistem menggunakan tahapan yang terdapat pada model pengembangan perangkat lunak SDLC (System Development Life Cycle) [4] dan pemodelan alur sistem digambarkan menggunakan UML Diagram, sehingga dapat dijelaskan secara spesifik, tervisualisasikan dan terdokumentasikan dengan baik[5]. Rancangan aplikasi pembelajaran ini, digambarkan menggunakan 3 jenis diagram dari UML yaitu, Use case diagram, Activity diagram dan Sequence diagram. Use case diagram digunakan untuk menunjukkan fungsi-fungsi utama dari sebuah sistem dan relasi antar fungsi tersebut, pada dasarnya Use case menunjukkan bagaimana aktor (pengguna) berinteraksi dengan sistem dan bagaimana modul sebuah sistem dijalankan [6]. Activity diagram digunakan untuk menggambarkan alur aktivitas, juga untuk memodelkan respon yang akan terjadi apabila suatu operasi dalam sistem dijalankan[7]. Sequence diagram memuat pengiriman pesan antar objek [8] dan merupakan eksekusi dari diagram Use Case . Sequence diagram mengilustrasikan gambaran detail dari setiap Use Case diagram yang dibuat dan direncanakan sebelumnya.

## II. PENELITIAN TERKAIT

Balaji dkk [9] dalam penelitiannya mengajukan rancangan pengembangan perangkat lunak untuk memfasilitasi komunikasi seseorang dengan keterbatasan penglihatan. Software yang dikembangkan menggunakan inputan suara sebagai jembatan interaksi antara manusia dengan mesin. Software ini dibangun menggunakan MATLAB, C++ dimana signal yang ditangkap melalui microphone kemudian diproses untuk menjalankan perintah dengan feedback berupa audio.

Ayse dkk [10] dalam penelitiannya mengembangkan sebuah software dengan memanfaatkan speech recognition di fakultas teknik sebuah universitas di turki untuk membantu mahasiswa dengan keterbatasan fisik dalam merancang sirkuit DC. Dengan menggunakan teknologi speech API (WSA) pada web browser chrome dan Symbolic Circuit Analysis dalam program MatLab (SCAM) yang digunakan untuk analisis dan perhitungan rangkaian. Dengan antar muka handsfree memudahkan siswa menjalankan simulasi dengan berbicara dan mengamatinya.

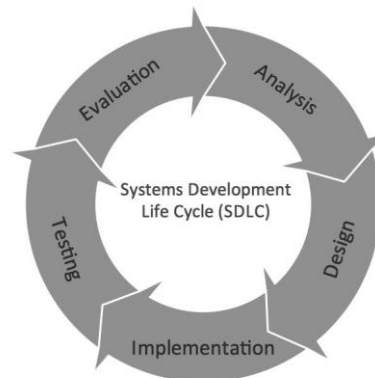
Ikhsan dkk [11] fokus pada pengembangan aplikasi pembelajaran daerah komoditas tambang. Aplikasi tersebut dikembangkan dengan metode prototyping dimana speech recognition difungsikan sebagai media untuk menginputkan perintah pencarian dalam bentuk suara. Aplikasi yang dikembangkan berbasis client server .

Supriyono [12] dalam penelitiannya mengembangkan software dengan memanfaatkan teknologi ASR (Automatic Speech Recognizer) pada Microsoft. Hal pertama yang dilakukan adalah program menangkap suara orang tersebut melalui microphone dan mengubahnya menjadi sinyal digital. Kemudian program menganalisa sinyal digital tersebut dengan membandingkannya dengan digital pattern yang ada dalam databasenya. Setelah itu akan diambil digital pattern yang paling besar prosentase kemiripannya, kemudian dari digital pattern tersebut diubah menjadi teks.

Gressia [13] dalam penelitiannya menggunakan algoritma FFT (Fast fourier transform) untuk menunjukkan proses dari speech recognition. Dimana proses awalnya ialah mengkonversi data spectrum suara kedalam bentuk digital dan mengubahnya dalam bentuk diskrit. Software yang dikembangkan berbasis desktop dengan memanfaatkan teknologi yang terdapat pada Microsoft voice dan Microsoft dictation.

## III. METODOLOGI PENELITIAN

Aplikasi pembelajaran bahasa inggris ini dirancang dengan menggunakan metode perancangan perangkat lunak SDLC (System Development Life Cycle). Dimana perancangan dilakukan secara bertahap sesuai dengan tahapan yang ada pada SDLC.



Gambar 1. Tahapan SDLC

### 1. Analisa

Fase analisa adalah sebuah proses investigasi terhadap sistem yang sedang berjalan dengan tujuan untuk mendapatkan jawaban mengenai pengguna sistem, cara kerja sistem dan waktu penggunaan sistem.

### 2. Perancangan

Merupakan proses penentuan cara kerja sistem dalam hal architecture design, interface design, database dan spesifikasi file, dan program design.

### 3. Implementasi

proses pembangunan sistem, instalasi sistem, dan rencana dukungan system.

### 4. Pengujian

Merupakan proses pengujian fungsi- fungsi pada aplikasi yang dibangun.

### 5. Evaluasi

Sistem yang sudah berhasil dibangun dievaluasi terkait kelebihan dan kekurangan serta keberlangsungan aplikasi kedepannya[4].

## IV. HASIL DAN PEMBAHASAN

### A. Analisis Sistem

Aplikasi ini menyajikan modul pelatihan bahasa inggris dimana user diminta untuk mengucapkan kata yang terdapat pada modul tersebut. Inputan yang masuk berupa suara user akan diubah menjadi teks, kemudian dicocokkan dengan kata yang terdapat dalam database. Untuk dapat mengubah suara menjadi text diperlukan speech recognition dimana diperlukam koneksi internet untuk melakukan streaming di Google.

Hasil dari pencocokan akan bernilai benar dan salah, dimana nilai tersebut akan ditampilkan dalam bentuk jawaban dan skor pada akhir pelatihan. Dalam menyelesaikan modul – modul tersebut tentunya user akan mengalami beberapa

kesulitan, seperti pengucapan yang salah. Oleh karena itu disediakan fitur evaluasi berupa text to speech dimana user akan dibantu dengan cara mendengarkan lalu menirukan sebuah kata atau kalimat.

**B. Perancangan Sistem**

*Skenario Aplikasi*

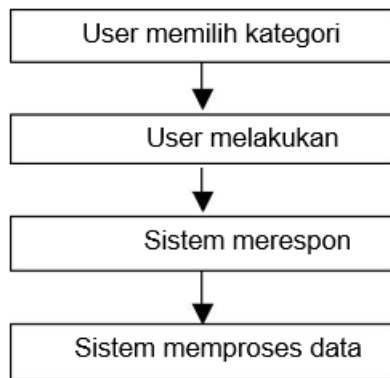
Secara garis besar aplikasi ini dibangun dengan menggunakan bahasa pemograman Java yang berbasis Android. RecognizerIntent dimanfaatkan untuk mengubah suara menjadi text. Untuk mengimplementasikan hal tersebut aplikasi perlu memverifikasi terlebih dahulu apakah perangkat target mampu mengenali speech input. Selanjutnya aplikasi menggunakan startActivityForResult() sebagai pemberitahuan bahwa aplikasi sedang meminta voice recognition. Penulisannya sebagai berikut :

```

PackageManager pm = getPackageManager();
List<resolveinfo> activities = pm.queryIntentActivities(new Intent(
    RecognizerIntent.ACTION_RECOGNIZE_SPEECH), 0);
if (activities.size() == 0) {
    mbtSpeak.setEnabled(false);
    mbtSpeak.setText("Voice recognizer not present")
    Toast.makeText(this, "Voice recognizer not present",
        Toast.LENGTH_SHORT).show();
}
    
```

Gambar 2. Pengenalan perangkat target

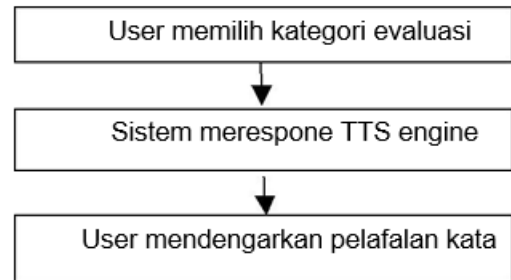
Pada aplikasi ini terdapat 2 kegiatan utama yaitu pembelajaran melalui modul dan evaluasi pembelajaran.



Gambar 3. Skema Pembelajaran Modul

Pada skema diatas menjelaskan ketika seorang user memulai melakukan pembelajaran maka hal pertama yang dilakukan adalah memilih kategori. Kategori ini terdiri dari tiga modul soal yang berbeda tingkat kesulitannya. Misalkan user memilih modul pertama, maka sistem akan merespon dengan menampilkan modul soal berupa kata yang harus diucapkan oleh user. Setelah itu user melafalkan kata tersebut

dan sistem akan memproses suara tersebut untuk diubah menjadi text yang kemudian dicocokkan dengan database. Dari proses tersebut nantinya akan diketahui benar atau salah pelafalannya.



Gambar 4. Skema Evaluasi Pembelajaran

Pada skema diatas menjelaskan tentang evaluasi modul pembelajaran. Fitur ini dibuat sebagai evaluasi pembelajaran dan membantu user untuk menyelesaikan modul pembelajaran tersebut. Langkah pertama saat user memilih fitur evaluasi maka user akan menginputkan kata atau kalimat berupa text. Dimana text ini nantinya akan dikonversi dalam bentuk suara kemudian didengarkan oleh user. Dengan demikian user dapat menirukan pengucapan yang didengarnya. Agar aplikasi dapat mulai berbicara diperlukan metode ttsSpeak()

*Perancangan Proses*

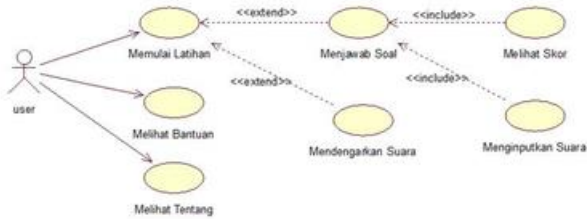
Proses yang berjalan pada aplikasi pembelajaran ini mencakup dua proses utama. Yaitu proses menginputkan suara pada modul pembelajaran dan proses evaluasi.

*Flowchart*



Gambar 5. Flowchart Proses Memulai Pembelajaran

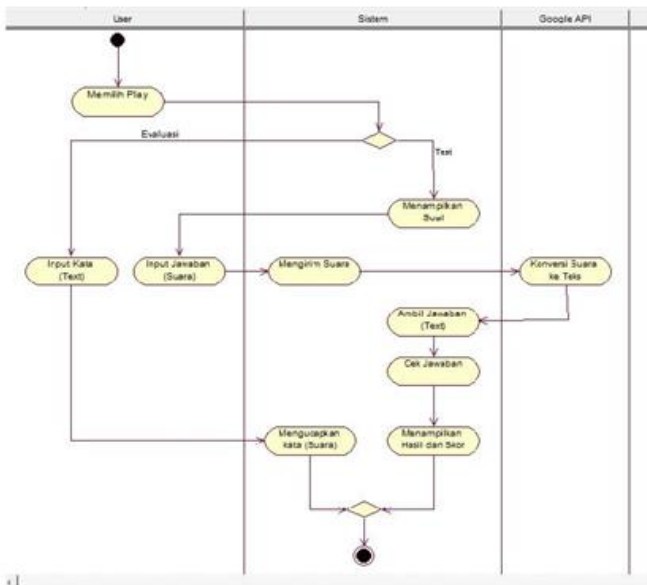
Usecase Diagram



Gambar 6. Use Case Diagram

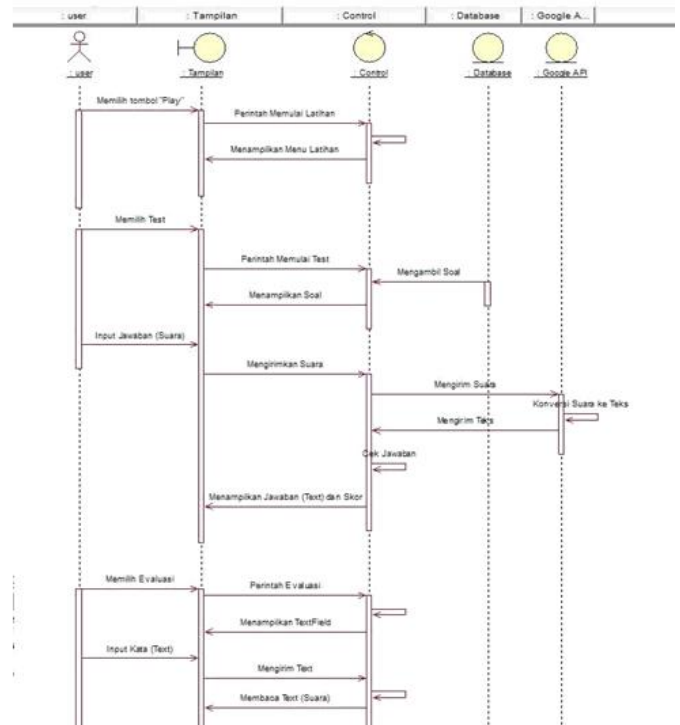
Use case diagram diatas menggambarkan perilaku (behavior) dari aplikasi ini. Dalam diagram ini user dapat memilih latihan, melihat bantuan, dan melihat tentang.

Activity Diagram



Gambar 7. Activity Diagram Memulai Latihan

Sequence Diagram



Gambar 8. Sequence Diagram Memulai Latihan

C. Perancangan Data

Pada sub-bab ini akan dijelaskan mengenai perancangan data yang berada dalam sistem, yaitu perancangan konseptual data. Terdiri dari tabel penyimpanan data dan Class yang digunakan. Class yang digunakan untuk penyimpanan data adalah DatabaseManager(). Berikut merupakan tabel penyimpanan data dan tabel class DatabaseManager() yang berjalan pada aplikasi.

Tabel 1. Tabel Penyimpanan Data

No	Tabel	Field	Type Data	Fungsi
1	Word & Sentence	Pertanyaan	varchar	Berisi pertanyaan pada aplikasi
		Jawaban	varchar	Jawaban dari soal

Tabel 2. Tabel Class DatabaseManager()

No	Nama Class	Atribut	Method
1	DatabaseManager	<i>Row_id:string</i> <i>Row_nama:string</i> <i>Nama_db:string</i> <i>Nama_table:string</i> <i>Db_version:int</i> <i>Create_table:string</i>	<i>addRow()</i> <i>close()</i> <i>Arraylist()</i> <i>Getversion()</i> <i>execSQL()</i>
2	DatabaseOpenHelper		<i>onCreate()</i> <i>onUpgrade()</i> <i>getWritableDatabase()</i>

**D. Implementasi**

*Proses Membangun Database*

Proses membangun database merupakan proses untuk memasukkan modul soal yang akan dipakai dalam aplikasi pembelajaran.

*Proses Random Soal*

Proses random soal merupakan proses dimana soal yang akan ditampilkan akan muncul secara random. Hal ini berfungsi agar soal yang muncul tidak dapat diprediksi.

*Pengecekan dukungan SpeechRecognition*

Pada proses ini terdapat proses pengecekan dukungan speech input, dimana terdapat tombol untuk memasukkan jawaban user dan hasil identifikasi jawaban user. Fungsi speech input akan dicek terlebih dahulu apakah Library tersedia. Yaitu dengan mengirim RecognizerIntent sekaligus menentukan model bahasa yang digunakan. Apabila Library tersedia maka proses dapat dilanjutkan. Namun apabila tidak tersedia proses akan berhenti.

*Pengecekan dukungan Text to Speech*

Pada proses pengecekan dukungan text to speech proses yang terjadi hampir sama dengan pengecekan speech input. Dimana diperlukan pengecekan terlebih dahulu apakah engine TTS tersedia pada perangkat.

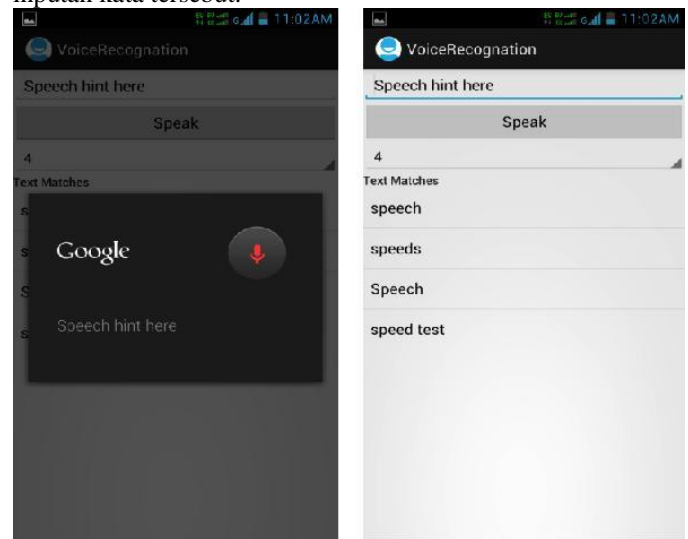
**E. Pengujian**

Pengujian akan dilakukan pada smartphone Andromax U, dengan spesifikasi sebagai berikut:

1. Sistem operasi Android 4.1 Jelly Bean
2. Prosesor Dual Core Snapdragon 1.2 GHz GPU Adreno 203.
3. Ukuran Layar 540×960 pixels.
4. Memori RAM 2 GB.
5. Memori eksternal 2 Gb.
6. Support speech input & Text to speech.

*Pengujian Pengaksesan Speech Input*

Proses pengujian pengaksesan speech input dilakukan dengan cara menginputkan suara pada device. Tujuannya adalah untuk mengetahui apakah smartphone yang digunakan support speech input dan RecognizerIntent mampu merespon inputan kata tersebut.



Gambar 9. Pengaksesan speech input

*Pengujian Tingkat Keakurasian Speech Input*

Keakurasian pengenalan suara dipengaruhi faktor faktor seperti user dan tingkat kesulitan kata untuk diucapkan. Untuk mengetahui keakurasian speech input maka diperlukan perbandingan data yang dihasilkan. Pengujian tingkat keakurasian speech input dilakukan dengan cara menginputkan 5 kata yang telah ditentukan, untuk diucapkan sejumlah user yang berbeda. Pengujian dilakukan pada ruangan dengan tingkat kebisingan rendah agar keakurasian lebih terjaga.

Tabel 3. Detail Uji coba akurasi

Nama	Kata	Benar	Dikenali sebagai
Inggar	Know	tidak	no
	No	ya	
	Write	tidak	right
	Psychology	ya	
Nizar	Know	ya	
	No	ya	
	Write	ya	
	Psychology	ya	
Dinda	Know	ya	
	No	ya	
	Write	ya	
	Psychology	ya	
Satria	Know	tidak	Now
	No	ya	
	Write	tidak	Right
	Psychology	ya	
Nisa	Know	ya	
	No	ya	
	Write	ya	
	Psychology	ya	
	Science	ya	
	Science	ya	
	Science	ya	
	Science	ya	

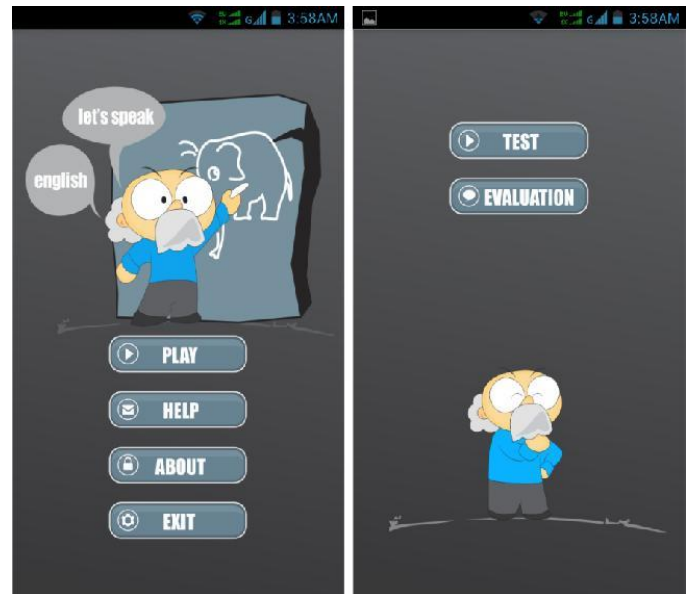
Berdasar data diatas, dapat disimpulkan bahwa tingkat akurasi speech input cukup baik. Hal ini dapat dilihat dari hasil pengujian bahwa jawaban user dinyatakan benar sebanyak 94 % , dikenali sebagai kata lain sebanyak 6 %.



Gambar 10. Tingkat akurasi

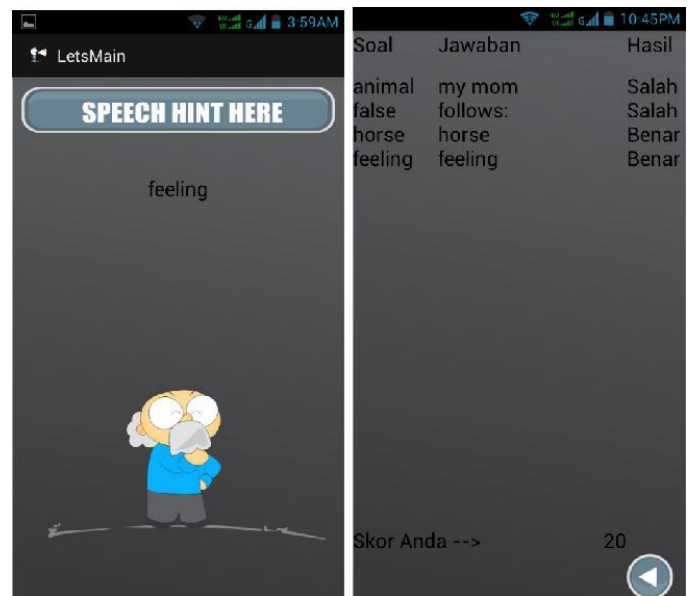
Pengujian Hasil

1. Menu Aplikasi



Gambar 11. Tampilan Awal Aplikasi

2. Tampilan memulai latihan dan skor



Gambar 12. Tampilan Menjawab Soal dan Skor

Setelah user memilih test, maka user akan memulai menjawab pertanyaan yang disiapkan dengan menekan tombol speech hint here dan next untuk melanjutkan ke soal

selanjutnya. Pada saat soal berakhir akan muncul tampilan jawaban dan skor sebagai bahan evaluasi.

## V. KESIMPULAN

Dari hasil implementasi dan pengujian pengaksesan speech input, uji tingkat akurasi dan pengujian hasil yang meliputi uji ketepatan jawaban dan ketepatan skor, sistem dapat berjalan dengan benar tanpa terjadi masalah. Aplikasi pembelajaran bahasa Inggris sudah dapat melakukan pencocokan kata dan kalimat dari modul soal yang telah ditentukan, serta memberikan nilai dan evaluasi. Hasil rancangan aplikasi pembelajaran pronunciation bahasa Inggris ini masih perlu diuji tingkat usability-nya. Juga masih perlu dilakukan analisa lebih dalam tentang seberapa efisienkah aplikasi ini dalam membantu proses pembelajaran, kendala yang mungkin terjadi dan solusi untuk mengatasinya.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada tim editorial Jurnal Teknologi Elektro atas dipublikasikannya penelitian ini.

## DAFTAR PUSTAKA

- [1] Ellis, H. J. C., & New, W., "Humanitarian open source software in computing education" (2007). *Education*, 0-3.
- [2] Gaikwad Santosh, 2010, A Review on Speech Recognition Technique, *International Journal of Computer Applications*.
- [3] Sagala, Syaiful. 2006. *Konsep dan makna pembelajaran*. Bandung : Alfabeta.
- [4] R. S. Pressman, *Rekayasa Perangkat Lunak: Pendekatan Praktisi* (Buku I). Yogyakarta: Penerbit ANDI, 2001.
- [5] A. R. Mulyanto, *Rekayasa Perangkat Lunak*. 2008
- [6] Z. Chen, "The research of software development supporting environment based on UML," in *2011 International Conference on Electronic & Mechanical Engineering and Information Technology*, 2011.
- [7] T. Shi-xiang and W. Sheng-ze, "The Conceptual Design and Simulation of Mechatronic System Base on UML," 2010, pp. 188-192.
- [8] P. Clements, F. Bachmann, L. Bass, D. Garlan, J. Ivers, R. Little, P. Merson, R. Nord, and J. Stafford, *Documenting Software Architectures: Views and Beyond* (Second Edition), Second. Boston: Pearson Education Inc., 2011.
- [9] Balaji, V., & Sadashivappa, G. (2016). Speech disabilities in adults and the suitable speech recognition software tools - A review. *2015 International Conference on Computing and Network Communications*, 2015, 559-564.
- [10] Yayla, A., & Buldu, A. (2017). An Interactive Web-Based Circuit Design and Analysis Interface for Disabled Students by Using Speech Recognition Technology. (April), 1676-1682.
- [11] Ikhsan, J & Yanto A. (2012). Implementasi speech recognition google API pada aplikasi pembelajaran daerah penghasil tambang Indonesia.
- [12] Supriyono B. (2014). Perancangan Aplikasi Voice User Interface Dengan Menggunakan Microsoft Speech API.
- [13] Melissa, G. (2008). Pencocokan Pola Suara (Speech Recognition) dengan Algoritma FFT dan Divide and Conquer. *Makalah If2251 Strategi Algoritmik*.