

## Portabel Text to Speech yang Terintegrasi dengan Telepon Seluler untuk Tunawicara

Akhmad Hendriawan, Ardik Wijayanto, Paulus S.W, Muhammad Taufiq,  
Email: hendri@eepis-its.edu, ardik@eepis-its.edu, wardana@eepis-its.edu, wongmetet@student.eepis-its.edu

### Abstrak

*Penelitian terhadap implementasi metode text to speech telah banyak dilakukan. Namun demikian, penelitian yang ada mempunyai kekurangan pada kemampuan pengenalan suku kata serta kapasitas database suara yang digunakan. Oleh karena itu, penelitian ini bermaksud mengatasi kelemahan penelitian sebelumnya dalam melakukan pengenalan suku kata. Informasi suku kata suara disimpan dalam Database suara yang berjumlah 4700 buah dengan masing-masing pola V, VK, VKK, K, KV, KVK, KVKK, KVKKK, KKV, KKVK, KKVKK, KKVKKK, KKKV, KKKVK (V adalah vokal dan K adalah konsonan). Teks yang ada dinormalisasi menjadi teks baru yang berupa deretan karakter huruf kapital dan kemudian dikonversi menjadi deretan suku kata menggunakan metode Finite State Automata (FSA). Deretan suku kata tersebut kemudian diproses menggunakan syllable concatenation dengan cara mencocokkan setiap database suara suku kata yang sesuai kemudian digabungkan satu sama lain sehingga diperoleh hasil akhir berupa suara sintesis. Berdasarkan hasil pengujian, sistem telah mampu memenuhi kontribusi yang diharapkan yaitu mampu mengenali suku kata dan mengkonversinya menjadi suara dengan tingkat keberhasilan 90% dari 10 macam teks yang diujikan. Hasil pengujian sistem dalam pengonversian suku kata menjadi suara juga diperoleh tingkat keberhasilan maksimal 75% dari 20 responden.*

Kata kunci: Text to Speech, Syllable Concatenation, Finite State Automata (FSA), Database Suku Kata, Tunawicara

### 1. Pendahuluan

Banyak sekali penelitian yang mengupas tentang permasalahan yang dihadapi oleh penyandang disabilitas. Permasalahan yang sering kali dihadapi oleh penyandang disabilitas khususnya tunawicara ialah komunikasi. Padahal pada jaman modern ini, banyak alat-alat telekomunikasi vital yang bekerja menggunakan masukan suara, seperti telepon, pesawat HT, internet dan yang lainnya. Padahal peralatan ini sangat penting untuk mengantisipasi keadaan darurat misalnya untuk menghubungi pos kepolisian, pos

pemadam kebakaran, rumah sakit, dan lain sebagainya. Dapat dipastikan mereka tidak dapat menggunakan alat-alat telekomunikasi tersebut[1]. Hal ini jelas membuktikan bahwa teknologi telekomunikasi belum menjangkau sepenuhnya bagi kalangan khusus seperti mereka. Disamping itu, sebagai bagian dari insan sosial tentunya mereka ingin tetap dapat berkomunikasi dengan orang lain, keluarga utamanya, walaupun keberadaannya tidak memungkinkan untuk berkomunikasi secara langsung. Sehingga keberadaan alat bantu telekomunikasi seperti ini pasti sangat diperlukan.

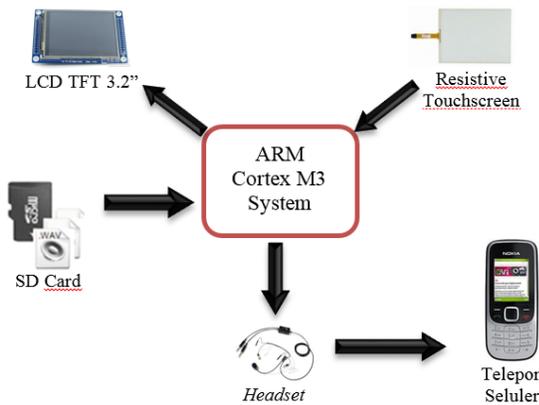
Namun demikian, sampai saat ini alat bantu telekomunikasi semacam ini masih terbatas seperti telepon TTY[2]. Telepon tersebut masih mengandalkan kabel PSTN sebagai saluran data dan belum beroperasi pada jaringan GSM ataupun CDMA sehingga hanya dapat digunakan di tempat-tempat tertentu seperti di rumah atau di kantor. Selain itu, harga dari telepon tersebut masih terbilang mahal yaitu \$339.17[3]. Terlebih, perangkat ini belum ada di Indonesia, sehingga perlu biaya tambahan untuk mendatangkannya dari luar negeri.

Melihat fenomena tersebut, maka sangat diperlukan suatu rekayasa teknologi yang mampu mewujudkan alat bantu telekomunikasi yang efektif, ekonomis, dan berdaya guna, serta mampu menjawab kebutuhan penyandang tunawicara. Banyak penelitian yang mengupas rekayasa teknologi tersebut seperti yang dilakukan oleh Dwi Lisnasari untuk mengimplementasikan alat bantu bicara portabel dalam bentuk pengubah teks ke suara[4]. Namun masih mempunyai kekurangan pada kemampuan pengenalan suku kata serta kapasitas database suara yang digunakan. Oleh karena itu, pada tugas akhir ini, penulis ingin menciptakan rekayasa teknologi berupa alat bantu untuk telekomunikasi jarak jauh dalam bentuk gadget portabel yang dilengkapi dengan papan ketik virtual sebagai media untuk menginputkan teks. Kemudian melalui sistem text to speech yang sudah terintegrasi di dalamnya, alat ini mampu mengubah teks menjadi suara yang berguna sebagai pengganti suara pengguna saat melakukan panggilan telepon. Dengan begitu, diharapkan mampu memberikan kemudahan dan kenyamanan dalam melakukan

telekomunikasi serta mampu mengatasi kesenjangan teknologi yang selama ini dialami oleh penyandang disabilitas khususnya tunawicara.

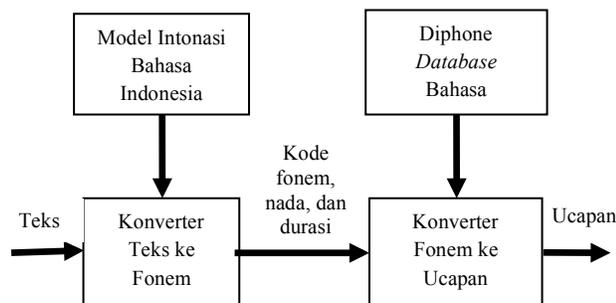
**2. Metode**

Secara umum alat ini merupakan embedded system dari integrasi perangkat LCD TFT Touchscreen yang difungsikan sebagai tampilan interface berupa keyboard virtual dengan mikrokontroler ARM Cortex M3 dan kemudian dipasang pada telepon seluler melalui headset. Blok diagram sistem ditunjukkan oleh gambar 1 sementara blok diagram sistem text to speech ditunjukkan oleh gambar 2. Dengan ditanamkannya text-to-speech engine dari sisi perangkat lunaknya sebagai pensintesis suara buatan, dapat diperoleh hasil akhir berupa perangkat baru yang berfungsi sebagai alat bantu telekomunikasi bagi tunawicara.



Gambar 1. Blok diagram sistem

Text-to-speech (TTS) atau pensintesa ucapan adalah suatu sistem yang dapat mengubah deretan teks menjadi ucapan sebagai keluarannya. Sistem pensintesa ucapan pada prinsipnya terdiri dari dua bagian dasar, yaitu:



Gambar 2. Blok diagram sistem *text-to-speech* [5]

1. Bagian konverter teks ke fonem

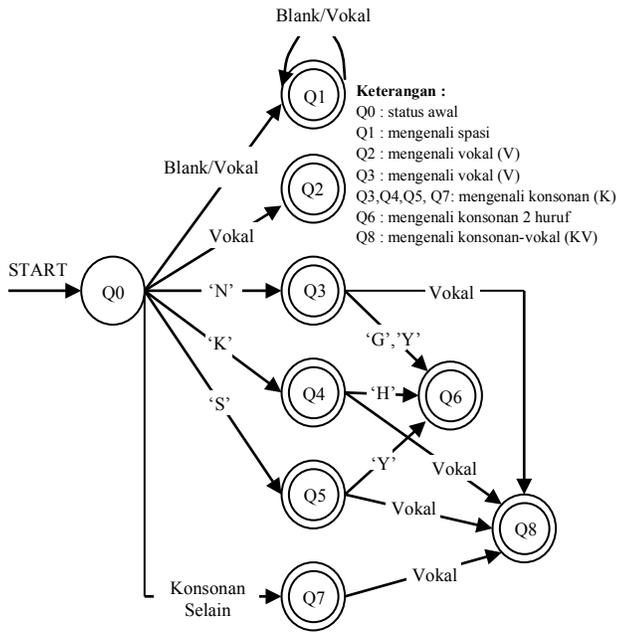
Bagian konverter teks ke fonem berfungsi untuk mengambil kalimat masukan dalam suatu bahasa tertentu yang berbentuk barisan teks dan mengubah beberapa hal seperti nomor dan tanda kedalam tulisan sesuai dengan bunyi yang seharusnya, sering disebut dengan normalisasi teks (text normalization). Kemudian menentukan kode fonetik (phonetic transcriptions) untuk tiap kata beserta durasi dan nadanya. Kode fonem adalah kode yang merepresentasikan unit bunyi yang ingin diucapkan. Pengucapan kata atau kalimat pada prinsipnya adalah urutan bunyi atau secara simbolik adalah urutan kode fonem.

2. Bagian konverter fonem ke ucapan

Bagian konverter fonem ke ucapan akan menerima masukan kode-kode fonem serta pitch dan durasi yang telah dihasilkan oleh bagian sebelumnya. Berdasarkan kode-kode tersebut bagian ini akan menghasilkan bunyi atau sinyal ucapan yang sesuai dengan kalimat yang ingin diucapkan. Ada beberapa alternatif teknik yang dapat digunakan untuk implementasi bagian ini. Salah satu teknik yang digunakan adalah penyambungan diphone (Diphone Concatenation). Pada sistem yang menggunakan teknik penyambungan diphone, sistem harus didukung oleh suatu database diphone yang berisi rekaman segmen-segmen ucapan yang berupa diphone.

Pada bagian konverter teks ke suara dapat digunakan algoritma Finite State Automata (FSA). Alur kerja dari algoritma FSA dalam mengenali dan memenggal suku kata ditunjukkan gambar 3.

Komponen sistem lainnya yang sangat penting ialah database suara. Database ini digunakan sebagai bahan referensi pembangkit sinyal suara. Database tersebut berisi kumpulan sampel suara dengan jumlah kurang lebih 4700 buah yang merupakan hasil rekaman suara suku kata baku bahasa Indonesia yang terdiri dari susunan V, VK, VKK, K, KV, KVK, KVKK, KVKKK, KKV, KKVK, KKVKK (V adalah vokal dan K adalah konsonan). Sampel suara tersebut direkam pada frekuensi 44100 Hz dalam format \*.WAV dan disimpan dalam memori SD Card. Sebagai media untuk memasukkan teks digunakan *interface* berupa *virtual keyboard* yang ditunjukkan oleh gambar 4.



Gambar 3. Diagram algoritma FSA[6]

Tombol-tombol yang digunakan merupakan tombol standar yang umum digunakan untuk pesan singkat/SMS pada smartphone. Tombol-tombol tersebut beserta fungsinya adalah sebagai berikut:

- Tombol angka 0 sampai 9, berfungsi sebagai input karakter angka.
- Tombol abjad A sampai Z, abjad a sampai z, berfungsi sebagai input karakter abjad.
- Tombol karakter titik, berfungsi sebagai input karakter titik.
- Tombol *shift*, berfungsi sebagai pengubah besar kecilnya huruf
- Tombol spasi, berfungsi sebagai input karakter spasi
- Tombol *delete*, berfungsi untuk menghapus seluruh teks yang telah di-input-kan.
- Tombol *backspace*, berfungsi sebagai menghapus satu karakter terakhir yang telah di-input-kan.
- Tombol *speak*, berfungsi untuk memroses teks yang telah diinputkan menjadi suara.



Gambar 4. Tampilan virtual keyboard

### Perancangan dan Pembuatan Algoritma

Perangkat lunak yang dimaksud ialah program komputasi untuk penyelesaian algoritma proses konversi dari teks menjadi suara. Berikut flowchart kerja dari perangkat lunak.



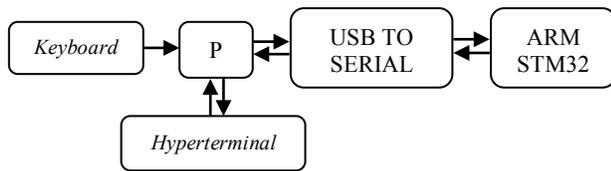
Gambar 5. Flowchart urutan proses konversi teks ke suara

### 3. Pengujian dan Analisa

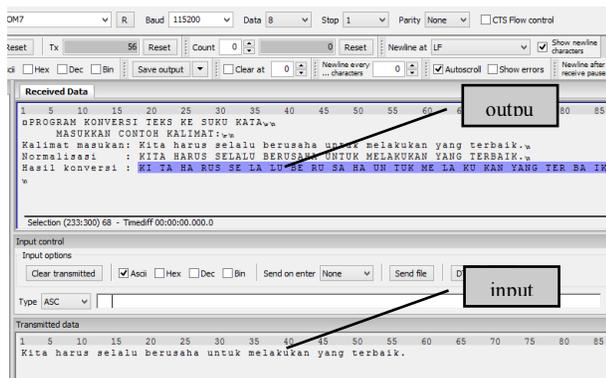
Pengujian ini dimaksudkan untuk mengetahui kehandalan perangkat lunak dalam melakukan konversi

teks ke suku kata yang blok diagramnya ditunjukkan oleh gambar 6. Pada pengujian ini digunakan mikrokontroler ARM sebagai pusat komputasinya. Kemudian hasil pemrosesan datanya yang berupa suku kata ditampilkan melalui hyperterminal pada layar PC. Hasil pengujian ditunjukkan oleh gambar 7

**Blok diagram sistem pengujian**



**Gambar 6.** Blok diagram sistem pengujian pengonveri teks ke suku kata



**Gambar 7.** Hasil pengujian konversi teks ke suku kata

**Tabel 1.** Uji coba konversi teks ke suku kata

| No. | Teks masukan   | Hasil konversi         | Indikasi |
|-----|----------------|------------------------|----------|
| 1.  | Itu buku saya  | I1 TU3 BU1 KU3 SA1 YA3 | Sukses   |
| 2.  | KePolisiAN     | KE1 PO2 LI2 SI2 AN3    | Sukses   |
| 3.  | KAPAN SAJA     | KA1 PAN3 SA1 JA3       | Sukses   |
| 4.  | ImpleMEntasi   | IM1 PLE2 MEN2 TA2 SI3  | Sukses   |
| 5.  | Konteks bahasa | KON1 TEKS3 BA1 HA2 SA3 | Sukses   |
| 6.  | Khusus Anda    | KU1 SUS3 AN1 DA3       | Sukses   |
| 7.  | PROYEK AKHIR   | PRO1 YEK3 AK1 HIR3     | Sukses   |
| 8.  | menggunakan    | MENG1 GU2 NA2 KAN3     | Sukses   |
| 9.  | Ekstraksi zat  | EKS1 TRAK2 SI3 ZAT3    | Sukses   |
| 10. | memBU@L s@JA   | MEM1 BU2               | Gagal    |

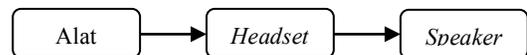
**Analisa hasil pengujian**

Pengujian ini dilakukan untuk mengonversi masukan sistem berupa teks kalimat menjadi suatu deretan suku kata. Untuk mendapatkan hasil tersebut, teks yang telah diinputkan harus dinormalisasi menjadi kumpulan string berupa deretan karakter huruf kapital, kemudian dikonversi menjadi deretan suku kata menggunakan metode Finite State Automata (FSA). Metode ini dimodifikasi sedemikian hingga mampu melakukan konversi seperti pada tabel 1. Dalam hal ini mampu mengenali posisi suku kata dalam suatu kata atau kalimat. Posisi ini menentukan pemakaian database yang akan digunakan. Terdapat tiga posisi yaitu di awal, di tengah, dan di akhir yang masing-masing diwakili dengan penomoran 1, 2, dan 3 di tiap akhir suku katanya.

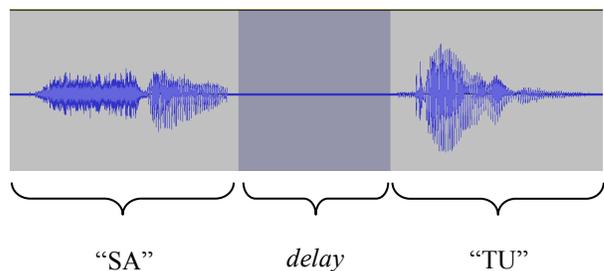
Berdasarkan hasil pengujian pada tabel 1 diketahui bahwa sembilan dari kesepuluh data pengujian, sistem mampu melakukan konversi dengan benar. Namun pada pengujian ke-10 sistem tidak dapat melakukan konversi disebabkan sistem tidak mengenali karakter @ (diluar huruf abjad) sehingga sistem mengalami eror pada posisi tersebut kemudian mengeluarkan warning berupa teks input tidak valid. Sehingga user harus memulai ulang untuk menginputkan teks dari semula.

**Pengujian sistem pengonversi suku kata ke suara**

Pengujian ini dilakukan untuk mengetahui karakteristik suara yang dihasilkan pada proses konversi suku kata menjadi suara. Blok diagram pengujian ditunjukkan oleh gambar 8 sementara hasil pengujian ditunjukkan gambar 9.



**Gambar 8.** Blok diagram sistem pengujian pengonversi suku kata ke suara



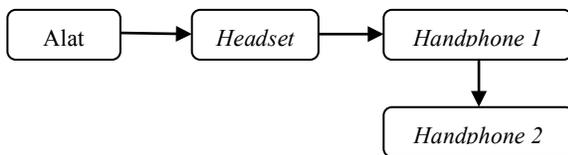
**Gambar 9.** Cuplikan spektrum sinyal suara "SATU"

**Analisa hasil pengujian**

Pada gambar 9 terdapat delay atau jeda yang memang dihasilkan oleh software dalam proses pencarian dan pencocokkan database suara suku kata yang sesuai. Lamanya jeda ini pun tergantung dari lamanya software dalam mencari file database tersebut. Sehingga menghasilkan suara yang kurang sesuai dengan yang diharapkan.

**Pengujian integrasi sistem dengan telepon seluler**

Pengujian sistem dari alat yang dtunjukkan oleh gambar 11 ini ditujukan untuk membandingkan tingkat kejelasan suara yang dihasilkan oleh alat setelah dilakukannya integrasi sistem secara keseluruhan. Blok diagram dari pengujian ditunjukkan oleh gambar 10 sementara hasil pengujian ditunjukkan oleh tabel 2



**Gambar 10.** pengujian pengonversi suku kata ke suara



**Gambar 11.** Dokumentasi pengujian sistem

**Tabel 2.** uji coba tingkat kejelasan suara hasil sintesis

| NO. | Kata yang disintesis | Respon            |
|-----|----------------------|-------------------|
| 1.  | Fajar                | Suara cukup jelas |
| 2.  | Terang               | Suara cukup jelas |
| 3.  | Sumber               | Suara samar       |
| 4.  | Panah                | Suara cukup jelas |
| 5.  | Semua                | Suara samar       |
| 6.  | Terka                | Suara cukup jelas |
| 7.  | Tidak                | Suara cukup jelas |
| 8.  | Sekarang             | Suara cukup jelas |
| 9.  | Mendapatkan          | Suara cukup jelas |
| 10. | Selamannya           | Suara samar       |

Pada pengujian ini, alat dipasang pada handphone 1 melalui headset. Sedangkan handphone 2 dipakai oleh responden. Melalui komunikasi telepon, responden kemudian diperdengarkan hasil sintesisnya. Parameter yang dipakai acuan ialah tingkat kejelasan suara. Sesuai dengan tabel 2, dari kesepuluh kata yang dipilih secara acak, tidak semua hasil suara sintesis terdengar dengan jelas. Hal ini dikarenakan kualitas database suara memiliki karakteristik yang kurang sesuai dengan yang diharapkan. Maksudnya ialah kesesuaian pengucapan bacaan dalam proses perekaman database. Diantaranya suku kata yang masih mengalami kendala ialah kata yang di dalamnya mengandung unsur huruf ‘n’, ‘ny’, ‘ng’, atau ‘m’.

**4. Kesimpulan**

Setelah melakukan pengujian dan analisa, maka dapat diambil beberapa kesimpulan tentang kinerja sistem yang telah dibuat, sebagai berikut :

1. Hasil pengujian sistem dalam pengenalan suku kata diperoleh tingkat keberhasilan 90% dari 10 macam teks yang diujikan.
2. Hasil pengujian sistem dalam pengonversian suku kata menjadi suara juga diperoleh tingkat keberhasilan maksimal 75% dari 20 responden.
3. Dari pengujian sistem secara keseluruhan, kualitas suara yang dihasilkan dinilai 80% cukup dan 20% bagus dari respon terhadap 5 subyek tunawicara.
4. Kata-kata yang sering mengalami kendala suara samar adalah kata yang di dalamnya mengandung unsur huruf ‘n’, ‘ny’, ‘ng’, atau ‘m’.

**Referensi**

[1] [http://alatbantuanlb.blogspot.com/2010/11/alat-bantu-tuna-wicara.h tml?m=1](http://alatbantuanlb.blogspot.com/2010/11/alat-bantu-tuna-wicara.html?m=1), dikunjungi tanggal 20 Mei 2012.

[2] [http://en.wikipedia.org/wiki/Telecommunications\\_devices\\_for\\_the\\_deaf](http://en.wikipedia.org/wiki/Telecommunications_devices_for_the_deaf), dikunjungi tanggal 20 Mei 2012.

[3] <http://www.uic.edu/depts/accc/telecom2.0/phone/deafevices.shtml>, dikunjungi tanggal 20 Mei 2012.

[4] Lisnasari, Dwi. 2010. Perancangan dan implementasi komunikasi data Text To Speech (TTS) dalam bahasa Indonesia. Proyek Akhir PENS-ITS. Surabaya

[5] Arman, Arry Akhmad. 2004. Konversi dari Teks ke Ucapan. Bandung

[6] Basuki, Thomas Anung. 2000. Pengenalan Suku Kata Bahasa Indonesia Menggunakan Finite-State Automata. Bandung.