

APLIKASI *TEXT TO SPEECH* DALAM SISTEM PENERJEMAH BAHASA INDONESIA-MADURA MENGGUNAKAN METODE FSA (*FINITE STATE AUTOMATA*)

Qutsiyah, Fika Hastarita Rachman, Firdaus Solihin

Program Studi Teknik Informatika, Universitas Trunojoyo Madura

Jl. Raya Telang, PO BOX 2, Kamal, Bangkalan - 69162

E-mail: qutsi92@gmail.com, hastarita.fika@gmail.com, fsolihin@gmail.com

ABSTRAK

Kurangnya penggunaan bahasa dan minimnya pelestarian Bahasa Madura menyebabkan Bahasa Madura mulai terancam punah. Bahasa Madura merupakan bahasa yang sulit untuk dipelajari, karena perbedaan dialek yang berbeda-beda dalam Bahasa Madura, sehingga banyak ditemukan perbedaan antara penulisan dan cara pengucapan.

Dalam rangka upaya pelestarian diperlukan pengembangan dan pemanfaatan teknologi. Usaha pelestarian Bahasa Madura salah satunya menggunakan aplikasi *Text to Speech* pada sistem penerjemah Bahasa Indonesia ke Bahasa Madura. Untuk mempermudah pengkonversian dari teks menjadi suara, proses pemenggalan kata menggunakan metode *Finite State Automata* dua tingkatan dimana hasil pemenggalan di tingkat pertama menjadi inputan di tingkat kedua.

Penerapan *Finite State Automata* pada aplikasi *Text to Speech* efektif digunakan dengan prosentase keberhasilan 90%. Output suara yang dihasilkan sesuai dengan hasil pemenggalan suku kata, namun pengucapan beberapa kalimat belum memiliki pola intonasi yang benar dengan prosentase keberhasilan 85%, sehingga pola intonasi hasil pengucapan suara belum sesuai dengan pengucapan aslinya.

Kata kunci: Penerjemah Bahasa Indonesia-Madura, *Text To Speech*, *Finite State Automata*.

ABSTRACT

Undercommunication of the utilization languages and minimizing of Madurese language preservation cause Madurese language start threatened in extinct. Actually, Madurese language it's difficult to learn, because of the differences of dialects inside the part of Madurese language itself, so it will find many differences between the process of writing and the way of pronunciation.

That in order to preservation efforts needs the development and technology utilization. One of the ways to preserve Madurese language it's use Text to Speech application in the system of Indonesian to Madurese translator. In order to make it easier a conversion from texts into the voice, process of separation words use Finite State Automata method in second stage where are the result of separation at first stage become second stage input.

As regards the application of Finite State Automata into Text to Speech application effectively it's a high percentage until 90%. Voice's Output that suitably with the result that separation in a syllable, however pronouncing in several sentence not including have the right form of intonation with a high percentages until 85%, with the result that the form of intonation of voice's pronunciation would have not suitably with the origin result.

Keywords: Indonesian-Madurese Translator, *Text to Speech*, *Finite State Automata*

PENDAHULUAN

Indonesia merupakan salah satu negara yang memiliki keanekaragaman budaya yang sangat tinggi, perbedaan bahasa daerah seperti Bahasa Madura keberadaannya sudah mulai menipis. Salah satu faktor penyebab kepunahan tersebut dikarenakan proses terjadinya globalisasi dan urbanisasi yang dapat memicu terjadinya asimilasi dan akulturasi budaya, terutama di daerah perkotaan. Sehingga perlu solusi untuk mempertahankan Bahasa Madura, yaitu dengan aplikasi penerjemah Bahasa Madura-Indonesia sebagai alternatif dalam upaya untuk melestarikan keberadaan dan keragaman Bahasa Madura. Bahasa Madura memiliki tiga tingkatan bahasa yaitu tingkat bahasa umum (iyâ-enjâ'), tingkat bahasa tengah (èngghi-enten), dan tingkat bahasa halus (èngghi-bunten).

Banyak masyarakat Madura yang kurang dalam tingkat kepedulian pada keragaman bahasa Madura, ketidaksadaran masyarakat bahwa Bahasa Madura memiliki tingkat keunikan tersendiri seperti *fonem-fonem* Madura dengan pengucapan kata yang dapat dihembuskan seperti: bh, dh, gh, dan jh. Melihat keunikan Bahasa Madura dengan keragaman pengucapan bahasa dan permasalahan berkurangnya seni Bahasa Madura, aplikasi penerjemah yang dilakukan oleh Rahilah[1] telah membantu untuk mempertahankan keberadaan Bahasa Madura. Tetapi penelitian tersebut sebatas dari segi tertulis, karena dialek-dialek bahasa Madura yang tidak sama dengan pengucapannya dapat mempersulit, sehingga perlu untuk melafalkan teks yang sudah diterjemah untuk mempermudah mempelajari Bahasa Madura. Dengan aplikasi *text to speech* dalam sistem penerjemah Bahasa Indonesia ke Bahasa Madura ini diharapkan membantu melestarikan keberadaan dan keragaman Bahasa Madura.

Penelitian ini bertujuan untuk membuat rancangan pengaplikasian *text to speech* dalam sistem penerjemah Bahasa Indonesia ke Bahasa Madura dengan menggunakan penerapan metode FSA (*Finite State Automata*). Program ini hanya mampu menerjemahkan kata atau kalimat yang sudah ada dalam *database* dari Bahasa Madura-Indonesia dan Indonesia-Madura. Tingkatan Bahasa yang digunakan yaitu semua tingkatan bahasa iyâ-enjâ', èngghi-enten dan èngghi-bunten.

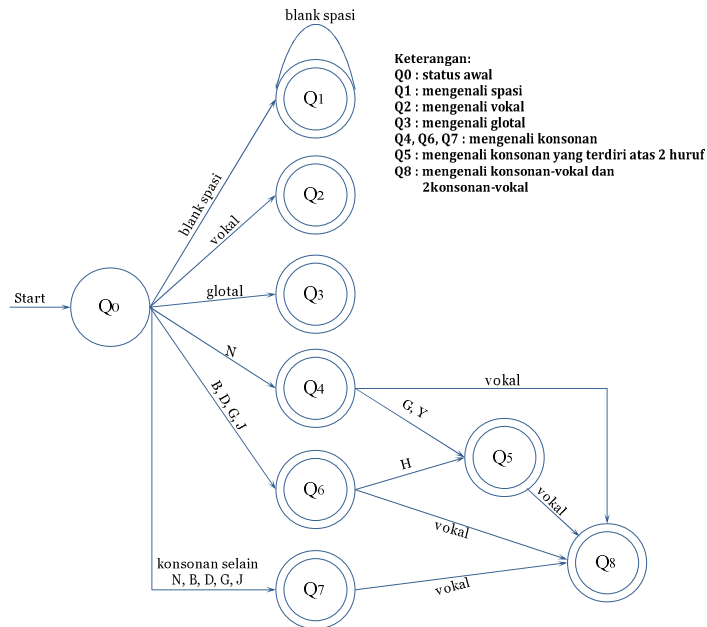
METODE

Finite State Automata

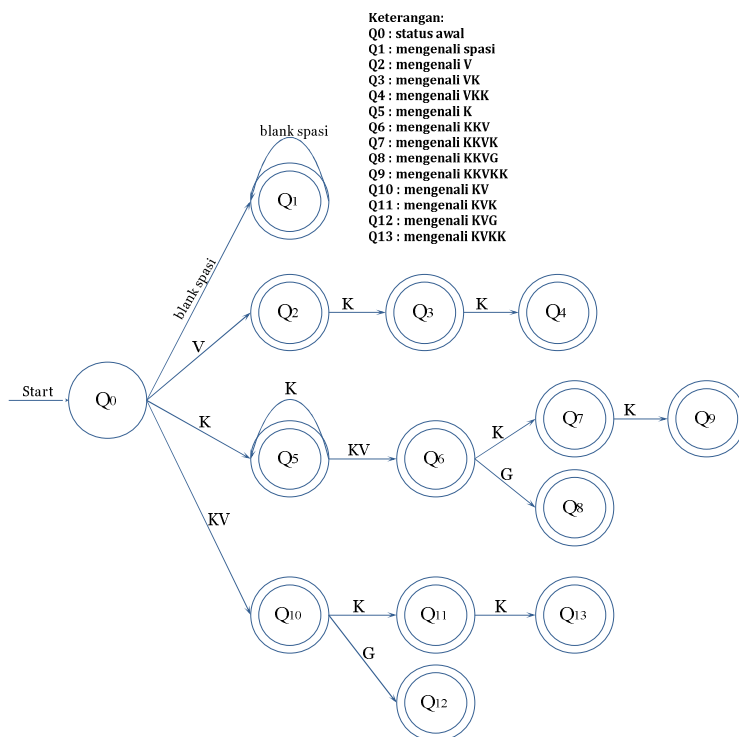
Dalam penelitian ini digunakan metode FSA (*Finite State Automata*), merupakan metode yang dapat diterapkan untuk pemisahan suku kata pada Bahasa Madura. FSA adalah salah satu sistem pemodelan matematika yang mempelajari tentang bahasa [2]. FSA berfungsi sebagai pengenalan(*recognizer*) suatu bahasa dengan melakukan pengelolaan dari masukan-masukan yang berupa *string* dan mengeluarkan keputusan YA jika *string* masukan termasuk dalam bahasa dan TIDAK jika *string* masukan tidak termasuk dalam bahasa.

FSA yang digunakan pada penelitian ini dirancang dalam dua tingkatan. Pada tingkatan pertama mengenali pola V, G, K, KV dan KKV. Hasil pengenalan FSA pada tingkatan pertama menjadi masukan bagi FSA tingkatan kedua. Pada FSA tingkatan kedua dapat mengenali semua

pola yang ada untuk bahasa Madura, yaitu pola V, VK, VKK, KV, KVK, KKV, KKVK, KVKK, KKVKK, KVG, dan KKVG.



Gambar 1. Diagram Transisi FSA tingkat pertama



Gambar 2. Diagram Transisi FSA Tingkat Kedua

Untuk penetapan pola umum Bahasa Madura berdasarkan literatur pola umum Bahasa Indonesia dan berdasarkan aturan pemenggalan suku kata pada Bahasa Madura. Berikut beberapa pola umum Bahasa Madura:

Tabel 1. Pola Umum Bahasa Madura

No.	Pola Bahasa Madura	Contoh Kata
1.	V	<i>a-tèh</i> (hati)
2.	VK	<i>ès-to</i> (sungguh)
3.	KV	<i>mè-reng</i> (mendengar)
4.	KVK	<i>dân-tos</i> (tunggu)
5.	KKV	<i>dhi-mèn</i> (dulu)
6.	VKK	<i>ong-ghâ</i> (naik)
7.	KKVK	<i>pot-trah</i> (putra)
8.	KVKK	<i>è-rèng</i> (iring),
9.	KKVKK	<i>am-bhung</i> (hidung),
10.	KVG	<i>ba-ri'</i> (kemarin)
11.	KKVG	<i>dhâ'-âr</i> (makan),

Keunikan Bahasa Madura

Bahasa Madura mempunyai ciri-ciri khas, baik dalam bidang fonologi (bunyi bahasa), morfologi (bentuk), maupun sintaksisnya (susunan kata atau kalimat). Keunikan bahasa Madura antara lain: [3]

- 1) Tidak mengenal kata ganti orang ketiga
- 2) Mempunyai *fonem-fonem* beraspirat dan tanaspirat

Tanaspirat : bâbâ (=bawah)

Aspirat : bhâbâng (=bawang)

- 3) Mempunyai fungsi morfem “Tang” atau “Sang”

Bahasa Madura “asli” yang belum terpengaruh bahasa lain, sebagai penanda milik (*possessive pronoun*) orang pertama dalam tingkat bahasa umum “enja’-iya”, dipakai istilah “tang” atau “sang”.

Contoh: tang buku (=buku saya) bukan : bukuna sèngko’

- 4) Mempunyai fungsi morfem (--a)

Dalam bahasa Madura untuk menyatakan kata kerja bentuk *future* “akan”, menggunakan sufiks (--a)

Contoh: Sèngko’ abinèa (=saya akan beristri)

- 5) Mempunyai fungsi prefiks (e--)

Kalimat pasif bahasa Madura mudah diketahui dengan dipakainya prefiks (e-) pada kata kerjanya, baik pelakunya orang pertama, kedua ataupun ketiga.

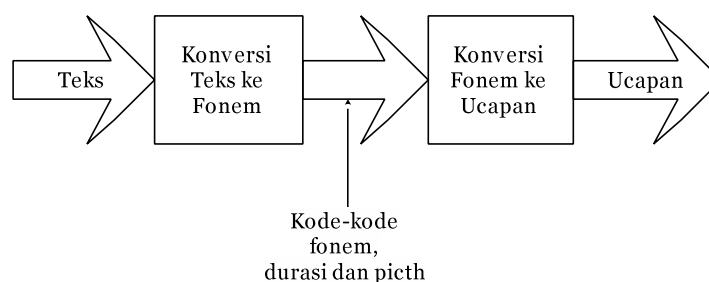
Contoh: Arèya sè èkaterroè bi' sèngko' (= Ini yang diinginkan saya).

Bahasa Madura memiliki karakter khusus dengan tujuan supaya mudah dibaca baik oleh orang Madura ataupun bukan orang Madura. Karakter tersebut adalah sebagai berikut [1].

1. a = dibaca a biasa seperti pada kata awas
2. â = dibaca e seperti pada kata belajar
3. e = dibaca e biasa seperti pada kata kertas
4. è = dibaca e seperti pada kata bebas
5. bh, dh, gh dan jh = dibaca tebal, contoh : bhâjâ, dhâbu, ghâbây dan bâjhâ.
6. tanda petik (‘) = dibaca seperti pada kata tidak.tanda petik (‘) = dibaca seperti pada kata tidak.

***Text to speech* (TTS)**

Text to speech adalah suatu proses yang mengkonversi teks menjadi ucapan yang dapat dimengerti oleh pendengar. Sistem *text to speech* memiliki banyak kegunaan. Penggunaan utama dari sistem ini memang dikhususkan bagi tunanetra sebagai suatu sistem untuk membaca teks dari sebuah buku kemudian mengkonversinya menjadi suara. Namun seiring berjalannya waktu, sistem *text to speech* banyak digunakan untuk keperluan lainnya seperti perintah untuk pengoperasian aplikasi, pembelajaran dalam melafalkan bahasa asing, dan keperluan lainnya.

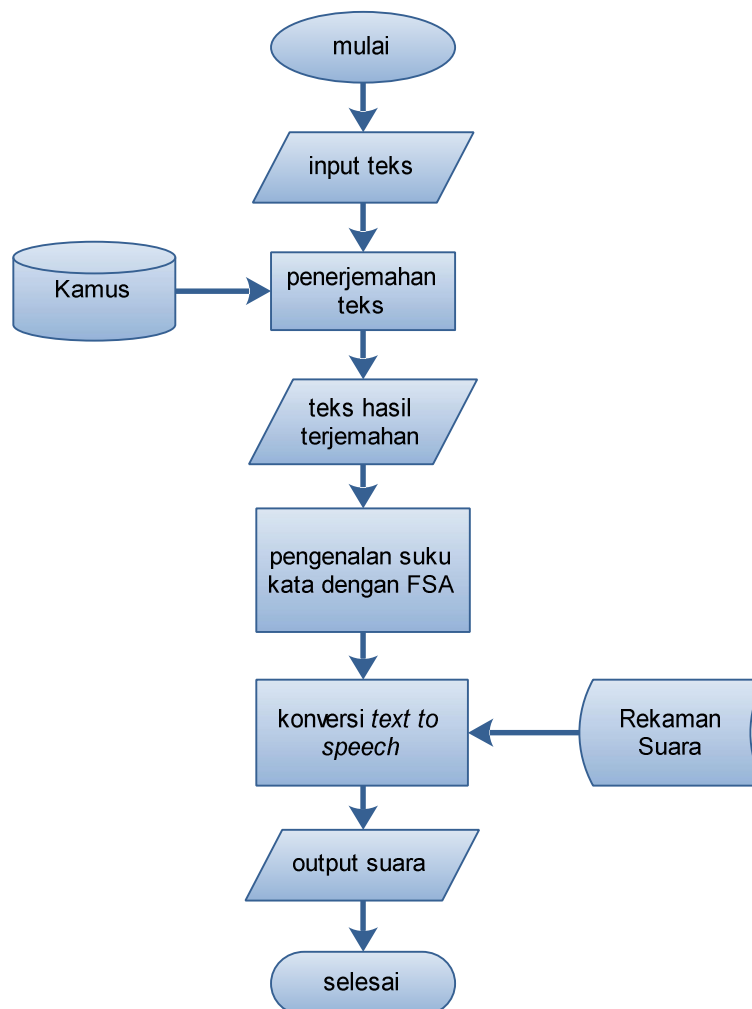


Gambar 3. Alur *Text to Speech*

Sistem konversi *text to speech* ini merupakan sistem yang mampu memproduksi sinyal ucapan secara otomatis melalui transkripsi *grafem* ke *fonem* untuk kalimat yang diucapkan. Perbedaan sistem *text to speech* dengan *talking machine* adalah keotomatisannya dalam mengucapkan kata-kata baru. Oleh karena itu sistem ini memungkinkan untuk diimplementasikan pada bidang aplikasi yang beragam seperti aplikasi sms bicara, buku digital dan pembaca *email* otomatis[4].

Rancangan Sistem

Flowchat Umum Sistem

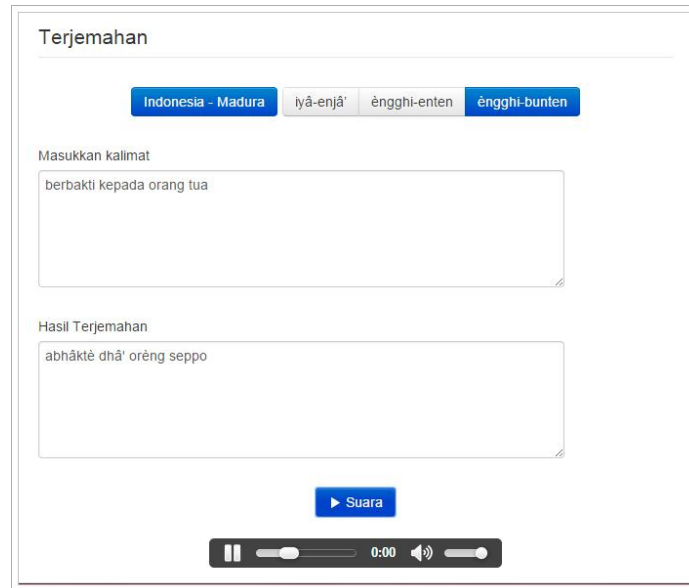


Gambar 4. Flowchart Umum Sistem

Proses *flowchat* pada Gambar 1 diawali dengan memasukkan teks Bahasa Indonesia, setelah itu menerjemahkan teks Bahasa Indonesia dalam Bahasa Madura. Hasil teks dari penerjemahan tersebut, dilakukan pemisahan suku kata dengan menggunakan FSA. Suku kata yang telah diperoleh, kemudian dicocokkan dengan rekaman suara. Maka suara hasil terjemahan dalam Bahasa Madura akan keluar sehingga dapat didengarkan.

HASIL DAN PEMBAHASAN

Aplikasi *text to speech* dalam penelitian ini telah di uji coba pada web browser dengan layanan host local (*localhost*). Aplikasi *text to speech* dalam penelitian ini telah diselesaikan menjadi suatu aplikasi berbasis web. Untuk memulai implementasi program dapat dilakukan dengan menjalankan aplikasi pada *localhost*.



Gambar 5. Running Aplikasi

Untuk mengenali suku kata Bahasa Madura menggunakan pemodelan FSA dua tingkat yang berfungsi mempermudah pemisahan suku kata. Pada aplikasi ini, telah dilakukan uji coba beberapa kata dan kalimat. Uji coba pertama yaitu mengetahui hasil penerapan FSA untuk pemisahan kata pada Bahasa Madura. Uji coba kedua dan ketiga yaitu mengetahui pengucapan kata dan kalimat Bahasa Madura dengan benar.

Tabel 1. Hasil Uji Coba Pemisahan Suku Kata

No.	Input	FSA		Hasil
		Tingkat 1	Tingkat 2	
1.	antèng	an-tè-ng	an-tèng	benar
2.	neddhâ	neddhâ	neddhâ	benar
3.	ramah	ra-ma-h	ra-mah	benar
4.	kaule	ka-u-le	ka-u-le	salah
5.	paras	pa-ra-s	pa-ras	benar
6.	rabu	ra-bu	ra-bu	benar
7.	saè	sa-è	sa-è	benar
8.	rajih	ra-ji-h	ra-jih	benar
9.	be'na	be-'-na	be'na	benar
10.	sèra	sè-ra	sè-ra	benar
11.	atèh	a-tè-h	a-tèh	benar
12.	socah	so-ca-h	so-cah	benar
13.	sèdhâh	sè-dhâ-h	sè-dhâh	benar
14.	mèbis	mè-bi-s	mè-bis	benar
15.	seppo	se-p-po	sep-po	benar
16.	o-ro'	o-ro-'	o-ro'	benar
17.	si-yam	si-ya-m	si-yam	benar
18.	serrang	se-r-ra-ng	ser-rang	benar
19.	sarèn	sa-rè-n	sa-rèn	benar
20.	sepprei	se-p-p-r-e-i	sep-pre-i	salah

Hasil uji coba yang telah dilakukan untuk pemisahan suku kata menggunakan FSA (*Finite State Automata*) pada Tabel 1 sudah cukup efektif digunakan dengan presentase keberhasilan 90%. Namun metode FSA dua tingkat tidak dapat membedakan gabungan vokal au dan ai untuk kata yang mengandung diftong (dua vokal yang diucapkan sekaligus) maupun tidak diftong. Semua gabungan vokal au dan ai dibaca tidak diftong. Sehingga pemisahan suku kata untuk kata yang mengandung diftong tidak sesuai dengan aturan pemisahan suku kata.

Tabel 2. Hasil Uji Coba Pengucapan Kalimat dengan Pola Intonasi

No.	Indo	Madura	Output Suara	Pola intonasi
1.	Saya makan nasi kemarin	engko' ngakan nasè' bâri'	benar	benar
		bulâ neddhâ nasè' bâri'	benar	benar
		abdhina dhâ'âr nasè' bâri'	benar	benar
2.	Ibu sudah sholat subuh	embu' marè sholat sobbhu	benar	benar
		embu' ampon sholat sobbhu	benar	benar
		èbhuh lastarè sholat sobbhu	benar	benar
3.	Rumah yang bagus	bengkoh sè bhâghus	benar	benar
		compo' sè bhâghus	benar	benar
		compo' sè bhâghus	benar	benar
4.	Bersama lebih baik	abhâreng lebbi bhâgus	benar	benar
		asareng lebbi bhâgus	benar	benar
		asareng langkong bhâgus	benar	benar
5.	Memberi makan anak yatim	aberri' ngakan ana' jatèm	benar	benar
		abherri' neddhâ ana' jatèm	benar	benar
		aparèng dhâ'âr ana' jatèm	benar	benar
6.	Pergi ke sekolah naik sepeda	èntar dâ' sakolah naè' sapèda	benar	benar
		mèyos dâ' sakolah onghâ sapèda	benar	benar
		mèyos dâ' sakolah onghâ sapèda	benar	benar
7.	Berbakti kepada orang tua	abhâktè ka orèng towah	benar	benar
		abhâktè dhâ' orèng tuwah	benar	benar
		abhâktè dhâ' orèng seppo	benar	benar
8.	Rindu setengah mati	kerrong satengnga matèh	benar	benar
		kerrong satengnga sidhâh	benar	benar
		kerrong satengnga sèdhâh	benar	benar
9.	Panutan semua murid	panodhân kabbhi morèd	benar	benar
		panodhân sadhâjâh morèd	benar	benar
		panodhân samandhâjâh morèd	benar	benar
10.	Berangkat sekolah bersama teman	mangkat sakolah abhâreng kancah	benar	benar
		mangkat sakolah asareng kancah	benar	benar
		mangkat sakolah asareng kancah	benar	benar
11.	Mencuci baju di sungai	nyassa kalambhi è songay	benar	benar
		nyassa rasok è songay	benar	benar
		nyassa rasok è songay	benar	benar
12.	Belajar berhitung bersama teman	âjhâr abitong abhâreng kancah	benar	benar
		âjhâr abitong asareng kancah	benar	benar
		âjhâr abitong asareng kancah	benar	benar
13.	Uangnya bapak dicuri maling	pèssèna eppa' èkèco' malèng	benar	benar
		obângna bhâpa' èkèco' malèng	benar	benar

		obângna ramah èkèco' malèng	benar	benar
14.	Bayi itu dipeluk ibunya	bhâji' jrowa èdhekkep embu'na	benar	benar
		bhâji' ghruwâ èdhekkep embu'na	benar	benar
		bhâji' ghâpanèkah èdhekkep èbhuhna	benar	benar
15.	Ada dimana sekarang?	bâdâh èdimmah satèya?	benar	salah
		bâdâh èghumma samangkèn?	benar	salah
		bâdâh èka'dimma satonto?	benar	salah
16.	Apakah kamu tahu alamat rumah itu?	apah bâ'na taoh alamat bengkoh jrowa?	benar	salah
		napè sampèyan onèng alamat compo' ghruwâ?	benar	salah
		panapa panjhennengngan mèyarsa alamat compo' ghâpanèkah?	benar	salah
17.	Inikah yang kamu cari?	arèya sè bâ'na sarèh?	benar	salah
		anèko sè sampèyan sarèh?	benar	salah
		panèka sè panjhennengngan sarèh?	benar	salah
18.	Pulang ke kampung halaman	molè dà' kampong halaman	benar	benar
		palèman dà' kampong halaman	benar	benar
		bhubhâr dà' kampong halaman	benar	benar
19.	Saya belajar ilmu agama	engko' âjhâr èlmo aghâma	benar	benar
		bulâ âjhâr èlmo aghâma	benar	benar
		abdhina âjhâr èlmo aghâma	benar	benar
20.	Memperbanyak amal ibadah	mabânnya' amal èbâdâ	benar	benar
		mabânnya' amal èbâdâ	benar	benar
		mabânnya' amal èbâdâ	benar	benar

Dari 20 hasil uji coba pengucapan kalimat Bahasa Madura tiga tingkatan yang terdiri dari kalimat sederhana, kalimat aktif, kalimat pasif, kalimat tanya dan kalimat majemuk pada Tabel 2 dengan prosentase keberhasilan pola intonasi sebesar 85%. Hasil analisa uji coba disimpulkan bahwa hasil pengucapan suara dari kalimat pada sistem dapat berjalan dengan baik. Kesesuaian hasil pengucapan suara dengan pengucapan aslinya bergantung pada hasil pemisahan suku kata. Untuk beberapa kalimat pola intonasinya kurang menyerupai ucapan asli dan masih kurang alami, seperti pengucapan untuk kalimat tanya yang tidak sesuai dengan pola intonasi sebagaimana pengucapan aslinya.

SIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pengujian dalam pembuatan Aplikasi *Text to Speech* dalam Sistem Penerjemah Bahasa Indonesia-Bahasa Madura Menggunakan FSA (*Finite State Automata*), dapat disimpulkan:

1. Metode FSA (*Finite State Automata*) dua tingkat efektif digunakan pada pemisahan suku kata bahasa Madura dengan prosentase 90%.
2. Pengenalan suku kata dalam aplikasi ini tidak dapat membedakan gabungan dua vokal au dan ai yang diftong dengan tidak diftong.

3. Aplikasi *Text to Speech* ini secara umum dapat mengucapkan kata/kalimat bahasa Madura dengan benar, mulai dari tingkatan bahasa kasar (iyâ-enjâ'), tingkat bahasa tengah (èngghi-enten), dan tingkat bahasa halus (èngghi-bhunten).

4. Keberhasilan pengucapan suara untuk kata/kalimat yang tidak diftong adalah 100%, dan prosentase tingkat kebenaran pola intonasi 85%.

5. Suara yang dihasilkan dalam Aplikasi *Text to Speech* ini tergantung dari input suara pada database, dan belum memiliki pola intonasi sebagaimana pengucapan aslinya.

Pada tugas akhir ini ingin diberikan beberapa saran yang mungkin berguna untuk pengembangan lebih lanjut untuk pelaksanaan belajar bagi masyarakat umum, yaitu :

1. Mengembangkan metode supaya dapat membedakan gabungan dua vokal au, ai yang diftong dan bukan diftong.
2. Perlu adanya penyempurnaan dalam pengucapan dan pola intonasi.
3. Aplikasi *Text to Speech* ini masih dapat mengucapkan kata/kalimat bahasa Madura, dalam penelitian berikutnya dapat dikembangkan dengan menambah pengucapan bahasa Indonesia hasil terjemahan dari bahasa Madura.

DAFTAR PUSTAKA

- [1] Rahilah. *Aplikasi Penerjemah Bahasa Madura-Indonesia dan Indonesia-Madura menggunakan Free Contex Parsing Algorithm*. Skripsi Jurusan Teknik Informatika Fakultas Teknik Universitas Trunojoyo Madura. 2013.
- [2] Wasista, S., Astin, N. Algoritma sistem pembaca teks Bahasa Indonesia menggunakan Metode FSA (*Finite State Automata*). *Industrial Electronics Seminar*. IES. Volume 1: 314-320. 3 November 2010.
- [3] Pawitra, A. *Kamus Lengkap Bahasa Madura Indonesia*. Jakarta: Dian Rakyat. 2009.
- [4] Astuti, E., S., Naba, E., A., prijono, W., A. Sistem TTS dalam Bahasa Indonesia menggunakan metode FSA dan database diphone. *Seminar Nasional Electrical, Informatics and It's Educations*. B1:139-144. 2009.