



## Analisis Implementasi Model Durasi Klatt pada Bahasa Melayu Pontianak Menggunakan *Diphone* Bahasa Indonesia

Arif Bijaksana Putra Negara<sup>#1</sup>, Numrotul Hatimah<sup>#2</sup>, Rudy Dwi Nyoto<sup>#3</sup>

<sup>#</sup>Fakultas Teknik Jurusan Informatika, Universitas Tanjungpura  
Jl. Prof. Hadary Nawawi, Pontianak

<sup>1</sup>arifbpn@informatika.untan.ac.id

<sup>2</sup>numrotul64@gmail.com

<sup>3</sup>rudydn@informatika.untan.ac.id

**Abstrak**— Bahasa Melayu Pontianak merupakan salah satu bahasa daerah yang ada di Provinsi Kalimantan Barat, khususnya banyak digunakan sebagai alat komunikasi oleh orang Melayu yang ada di Pontianak sebagai bahasa sehari-hari. Salah satu usaha untuk melestarikan Bahasa Melayu Pontianak yaitu dengan mengembangkan TTS. Salah satu model yang dapat dikembangkan untuk membangun TTS adalah model *diphone concatenation*. Untuk mengembangkan TTS model *diphone* diperlukan nilai durasi tiap fonem yang diucapkan. Penelitian ini bertujuan untuk mengetahui apakah model durasi Klatt yang menggunakan aturan suku kata dan tipe suku kata dapat memprediksi durasi fonem sintesa ucapan Bahasa Melayu Pontianak. Jumlah kalimat yang digunakan untuk analisis adalah 576 kalimat yang terdiri dari 5133 kata. Pengujian awal dilakukan untuk mencari rentang nilai kecepatan durasi yang masih dapat diterima oleh pendengar yang dinyatakan dalam rentang persentase. Hasil pengujian awal ini diperoleh nilai rentang kecepatan durasi adalah -20% hingga 20%. Pengujian selanjutnya adalah mencari nilai akurasi prediksi durasi fonem sintesa yang menggunakan metode Klatt terhadap durasi fonem dari 10 kalimat uji yang terdiri dari 320 fonem. Hasil pengujian durasi fonem dari 10 kalimat uji diperoleh 30% fonem yang kecepatan durasinya masih dapat diterima. Sedangkan hasil pengujian untuk mencari nilai akurasi prediksi durasi fonem terhadap total durasi kalimat pada 100 kalimat uji, diperoleh 94% kalimat yang kecepatan durasinya masih dapat diterima. Berdasarkan hasil pengujian tersebut model durasi Klatt yang menggunakan aturan suku kata dan tipe suku kata dapat memprediksi durasi fonem total pada ucapan sintesa bahasa Melayu Pontianak.

**Kata kunci**— durasi fonem, model durasi Klatt, melayu pontianak, *diphone* bahasa indonesia, aturan durasi

### I. PENDAHULUAN

Text to speech (TTS) adalah pengkonversi teks menjadi ucapan yang mirip dengan ucapan manusia berbasis komputer. Kualitas TTS dinilai dari dua hal penting yang selalu dijadikan tolak ukur, yaitu *naturalness* dan

*intelligibility* [1]-[3]. *Naturalness* menyatakan tingkat kealamian atau kemiripan sintesa ucapan yang dihasilkan dengan suara manusia. Sedangkan *intelligibility* dinilai dari kemampuan untuk memahami kata atau kalimat yang diucapkan. Parameter terpenting dalam kualitas keluaran sistem TTS adalah kejelasan suara yang diucapkan.

Sulitnya membangun TTS untuk suatu bahasa tentu memerlukan *database* yang besar. Pengembangan TTS saat ini banyak menggunakan model statistik. Membangun TTS pada model ini memerlukan usaha dan data yang sangat besar. Salah satu cara untuk mengurangi penyimpanan data yang besar yaitu dengan menggunakan *diphone database*. *Diphone* merupakan unit ucapan yang paling sering digunakan sebagai unit *database* [4].

Saat ini telah dikembangkan *diphone database* "id1" Bahasa Indonesia yang mana program ini dibuat menggunakan aksent Bahasa Indonesia [5]. *Diphone database* "id1" dibuat oleh Arry Akhmad Arman pada tahun 2000 di Belgia. *Software* pendukung yang dikembangkan untuk *database diphone* Bahasa Indonesia yaitu MBrola. Mbrola berfungsi sebagai pembangkit sinyal ucapan dengan menggabungkan fonem yang telah diberi nilai-nilai durasi dan pitch.

Penelitian sebelumnya yang menggunakan *diphone* Bahasa Indonesia dilakukan oleh Mizky Dwi Mentari Putri (2016) [6]. Pada penelitian tersebut, dilakukan pengembangan text to speech generik Bahasa Bugis Wajo menggunakan *diphone* Bahasa Indonesia. Sintesa ucapan pada penelitian ini menghasilkan bunyi yang jelas dengan nilai rata-rata kejelasan kata uji sebesar 92.71%. Tak hanya itu, pengembangan TTS Bahasa Melayu Pontianak sudah pernah dikembangkan. Penelitian ini telah dilakukan oleh Erwin Guntoro (2018) [7]. Namun ucapan yang dihasilkan masih belum baik, dikarenakan data yang digunakan dalam pengembangan TTS Bahasa Melayu Pontianak masih kurang.

Bahasa Melayu Pontianak merupakan satu diantara banyak bahasa yang ada di Provinsi Kalimantan Barat,

khususnya banyak digunakan sebagai alat komunikasi oleh orang Melayu yang ada di Pontianak sebagai Bahasa sehari-hari [8]. Namun tidak hanya penduduk asli yang menggunakan Bahasa Melayu Pontianak, tetapi ada juga orang Cina dan Dayak yang merupakan salah satu suku yang ada di Kota Pontianak. Kosakata Bahasa Melayu Pontianak hampir sama dengan Bahasa Indonesia. Bahasa Indonesia merupakan satu dialek temporal dari Bahasa Melayu, yang struktur dan khazanahnya sebagian besar masih sama atau mirip dengan temporal terdahulu [9]. Kemiripan tersebut dikarenakan akar dari Bahasa Indonesia adalah Bahasa Melayu, yang mana kedua bahasa ini masih serumpun. Set pos fonem Bahasa Indonesia ada pada fonem Bahasa Melayu Pontianak sehingga *diphone database* "id1" bisa digunakan pada Bahasa Melayu.

Untuk membangkitkan sintesa ucapan TTS Bahasa Melayu Pontianak menggunakan *diphone* diperlukan data berupa kode fonem, durasi dan nilai pitch. Salah satu indikator terpenting terhadap kejelasan ucapan yang disintesiskan untuk TTS suatu bahasa yaitu durasi tiap fonem. Durasi fonem sangat bervariasi tergantung pada penutur yang berbeda-beda sehingga karakteristik durasi rata-rata yang diperkirakan oleh penutur ahli bahasa dan orang awam pada umumnya juga berbeda [10].

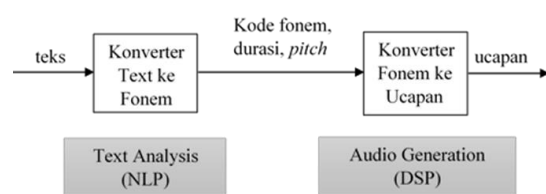
Pemberian durasi untuk tiap fonem dapat menggunakan beberapa model prediksi durasi. Salah satunya yaitu model durasi Klatt yang mana model ini menetapkan persentase kenaikan atau penurunan nilai durasi fonem. Model durasi Klatt termasuk model durasi berbasis aturan, model ini dapat dibangun pada *database* yang lebih kecil [11]. Sehingga model durasi berbasis aturan ini cocok dengan penggunaan *diphone database*. Beberapa penelitian telah menggunakan model durasi Klatt untuk memprediksi durasi fonem. Model Klatt telah dikembangkan untuk beberapa bahasa asing seperti Bahasa Inggris tahun 1987, Swedia tahun 1986, Jerman tahun 1994 dan Prancis tahun 1987 [2]. Penelitian terbaru prediksi durasi fonem menggunakan model Klatt dikembangkan pada bahasa Lituania pada tahun 2016. Suatu bahasa bersifat dependen, sehingga suatu metode bisa diterapkan untuk suatu bahasa tertentu tetapi belum tentu untuk bahasa yang lain. Oleh karena itu, penelitian yang akan dilakukan adalah memprediksi durasi fonem menggunakan model durasi Klatt pada ucapan Bahasa Melayu Pontianak dengan *diphone database* Bahasa Indonesia.

## II. TINJAUAN PUSTAKA

### A. Text To Speech

Sistem text to speech adalah suatu sistem yang mampu memproduksi sinyal ucapan secara otomatis melalui transkripsi grafem ke fonem untuk kalimat yang diucapkan [12]. Suatu sistem text to speech prinsipnya terdiri dari dua subsistem [1], yaitu:

1. Bagian konverter teks ke fonem (text to phoneme).
2. Bagian konverter fonem ke ucapan (phoneme to speech).



Gambar. 1 Sistem Text To Speech [1]

Bagian konverter teks ke fonem berfungsi untuk mengolah kalimat masukan dalam suatu bahasa tertentu yang berbentuk teks menjadi urutan kode-kode bunyi yang direpresentasikan dengan kode fonem, durasi serta pitch. Bagian konverter fonem ke ucapan menerima masukan kode-kode fonem serta pitch dan durasi yang telah dihasilkan oleh bagian sebelumnya.

### B. Fonem Bahasa Indonesia

Bahasa Indonesia memiliki enam buah fonem vokal, yaitu /i/, /e/, /ə/, /a/, /u/, dan /o/ dan delapan belas fonem konsonan diantaranya /p/, /t/, /c/, /k/, /b/, /d/, /j/, /g/, /m/, /n/, /ñ/, /ŋ/, /s/, /h/, /r/, /l/, /w/, dan /y/ [13]. Kemudian terdapat penambahan empat fonem yang berasal dari bahasa asing yaitu /x/, /z/, /f/, /ʃ/ [14].

### C. Perbedaan Fonologi Bahasa Melayu Pontianak dan Bahasa Indonesia

Bahasa Melayu Pontianak tidak memiliki bunyi /f/, /v/, /z/, /x/, dan /q/. Ciri khas Bahasa Melayu Pontianak ditemukan pada bunyi /e/ pepet seperti pada kata "lemari". Setiap kata yang memiliki akhiran bunyi a dalam bahasa melayu dialek lain dan Bahasa Indonesia akan berubah menjadi /e/ pepet [15].

Secara fonologis, bahasa Melayu Pontianak dan bahasa Indonesia memiliki perbedaan teratur dan perbedaan tidak teratur [16]. Perbedaan teratur meliputi [a]-[ə], [i]-[ε], [u]-[O], [r]-[k], [Ø]-[k]. Terdapat enam puluh tujuh (67) kata perbedaan teratur dalam bahasa Indonesia standar dengan bahasa Melayu Pontianak. Perbedaan tidak teratur terdiri dari [l]-[k], [k]-[g], [i]-[a]. Terdapat 3 perbedaan tidak teratur dalam bahasa Indonesia dengan bahasa Melayu Pontianak. Terdapat perbedaan diakhir kata dan juga diawal kata.

### D. Durasi

Menurut Ni Kadek Mega Ratnawati (2018) [17], durasi sering juga disebut kuantitas (quantity) karena menyangkut jumlah atau lamanya waktu yang digunakan untuk pengucapan sebuah bunyi.

### E. Model Durasi Klatt

Dennis Klatt mengusulkan sebuah model berbasis aturan, yang diimplementasikan dalam sistem MITalk [11][10]. Model durasi Klatt telah dikembangkan untuk beberapa bahasa, termasuk Bahasa Inggris, Swedia, Jerman dan Prancis [2][10][11]. Model Klatt menyatakan bahwa [2],[18]-[20].

1. Setiap jenis segmen fonetik memiliki durasi asli yang ditentukan sebagai salah satu sifat yang khas,
2. Setiap aturan mencoba untuk mempengaruhi persentase kenaikan atau penurunan durasi segmen, tetapi

3. Segmen tidak dapat dipersingkat kurang dari durasi minimum tertentu [21].

Menurut Kasparaitis (2016) [10], Klatt menentukan sejumlah efek kontekstual yang signifikan yang mempengaruhi durasi fonem. Setiap bahasa memiliki kumpulan efek kontekstualnya sendiri, yang mempengaruhi durasi fonem dan harus dibentuk dibawah pengawasan ahli bahasa. Contoh efek yang mempengaruhi durasi fonem menurut Klatt, sebagai berikut.

1. Apakah fonem tersebut termasuk vokal atau konsonan;
2. Apakah konsonan disuarakan atau tidak disuarakan;
3. Posisi fonem dalam sebuah kalimat, posisi fonem dalam satu kata;
4. Panjang kata;
5. Dampak dari fonem yang berdekatan.

Oleh karena itu, prediksi durasi suatu fonem dinyatakan dengan persamaan [2][18][20].

$$D_0 = K \cdot (D_i - D_{\min}) + D_{\min} \tag{1}$$

Persamaan (1) merupakan persamaan dari model durasi Klatt yang mana  $D_0$  adalah nilai prediksi durasi.  $D_i$  (durasi inheren) diperoleh dari hasil rata-rata durasi setiap fonem.  $D_{\min}$  (durasi minimal) diperoleh dari hasil analisis setiap fonem.  $K$  (konstanta) adalah persentase pemendekan yang ditentukan dengan menerapkan 11 aturan Klatt [18].

$$K = \frac{D_0 - D_{\min}}{D_i - D_{\min}} \tag{2}$$

Sebelum melakukan prediksi dari persamaan (1), terlebih dahulu menentukan nilai konstanta setiap fonem. Konstanta fonem diperoleh dari rumus yang mengacu pada (2). Untuk  $D_0$  (durasi awal atau durasi asli), durasi minimal fonem dan durasi rata-rata fonem diperoleh dari data kalimat Bahasa Melayu Pontianak.

F. Database Diphone Mbrola

Mbrola engine adalah software yang digunakan untuk mensistesa ucapan berkualitas tinggi. Mbrola menggunakan *database diphone* yang direkam dalam format 16bit sehingga kualitas suara yang dihasilkan sangat baik [23]. Menurut Anggara Narullita (2009) [24], *diphone* merupakan sebuah *database* yang digunakan oleh Mbrola engine untuk melakukan pemeriksaan string yang akan dirubah ke dalam bentuk suara, teks yang akan dirubah dalam bentuk suara harus melewati penyaringan kata yang sesuai dengan *diphone* terlebih dahulu, tanpa adanya *diphone* maka proses pengubahan teks menjadi suara tidak akan bisa dilakukan.

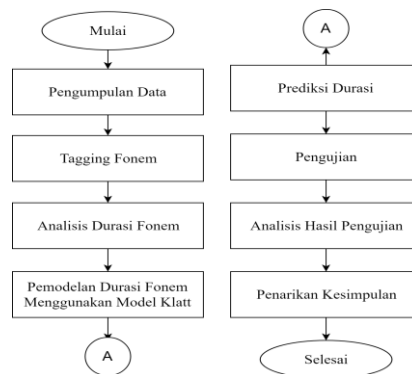
Saat ini telah tersedia *diphone database* dalam berbagai macam bahasa, antara lain bahasa Inggris Amerika, Inggris British, Jerman, Yunani, Hindi, Italia, Spanyol, Indonesia, Korea, Jepang, Turki, dan lainnya [7]. *Diphone* yang digunakan dalam penelitian ini adalah *diphone database* Bahasa Indonesia "id1". *Database diphone* bahasa Indonesia (Id1) memiliki 30 fonem termasuk spasi [1]. Table fonem untuk *database* Id1 dapat dilihat pada tabel 1 [23].

TABEL I  
FONEM DATABASE DIPHONE ID1

Fonem	Sampa	Contoh Kata
p	p	panas, apa, siap
b	b	baca, sebab
t	t	tanam, atap, penat
d	d	datang, tadi
k	k	kami, akan, tidak
g	g	gula, agar, reog
c	tS	cari, rencana
j	dZ	jadi, bejana
f	f	faktor, kafan, alif
s	s	saran, kasar, beras
z	z	zebra, azimuth
h	h	hitam, sehat, sah
m	m	musin, aman, hitam
n	n	nusa, anda, hutang
ng	N	angka, siang
r	r	robot, armada, sonar
l	l	lima, alamat, kesal
w	w	waktu, bawa
y	j	yakin, bahaya
ny	nY	nyaris, kenyang
a	V	akhir, bagus, nira
e	@	embun, petang
e	e	edisi, medan, hore
i	I	ikal, bila, bali
o	Q	orang, mohon, joko
u	U	usaha, bukan, biru
ai	aI	santai, landai
oi	OI	amboi
au	aU	ampau, pulau
Spasi		

III. METODE PENELITIAN

Langkah-langkah penelitian yang dilakukan pada penelitian ini dapat dilihat pada gambar 2 sebagai berikut:



Gambar. 2 Metode penelitian

Berdasarkan gambar 2 dapat dijelaskan metodologi penelitian sebagai berikut:

A. Pengumpulan Data

Pengumpulan data berupa data teks cerita dalam Bahasa Melayu Pontianak yang diperoleh dari hasil observasi. Data teks berupa Buku yang berjudul Sepok Satu karangan Pay Jarot Sujarwo. Jumlah kalimat yang digunakan dalam penelitian ini sebanyak 576 kalimat. Teks cerita dari buku tersebut direkam oleh seorang penutur ahli Bahasa Melayu Pontianak. Hasil rekaman disimpan dalam *file* suara mono, resolusi 16bit, dan sampling rate 44100 dengan format \*.wav.

B. Tagging Fonem

File rekaman suara yang tersimpan diproses menggunakan *wavesurfer* untuk diberi tanda atau label diawal hingga diakhir pengucapan fonem guna mendapatkan nilai durasi tiap-tiap fonem dalam kalimat. Setiap rekaman yang telah diberi label disimpan dalam *file* dengan format *.misc*.

C. Analisis Durasi Fonem

Fonem yang dianalisis adalah fonem yang dapat dikonversi ke dalam kode-kode fonem dengan *database diphone* Bahasa Indonesia. Pada *file \*.misc* terdapat nilai titik awal dan titik akhir pengucapan durasi suatu fonem. Sehingga lamanya durasi suatu fonem, baik fonem vokal maupun konsonan dapat diperoleh dari hasil perhitungan yaitu durasi akhir dikurang durasi awal. Hasil akhir analisis durasi fonem berupa nilai durasi dari penutur yang terdiri dari nilai durasi minimal, maksimal dan durasi rata-rata dari tiap-tiap fonem vokal dan konsonan.

D. Pemodelan Durasi Fonem Menggunakan Model Klatt

Pada tahap ini mencari nilai konstanta fonem vokal maupun konsonan. Variabel konstanta diperoleh dari persamaan (3) model durasi Klatt sebagai berikut:

$$D_0 = K \cdot (D_i - D_{min}) + D_{min} \tag{3}$$

Keterangan:

$D_0$  = Nilai durasi Klatt

$K$  = Nilai konstanta fonem konsonan/vokal

$D_{min}$  = Nilai durasi minimal penutur

$D_i$  = Nilai durasi rata-rata penutur

Nilai  $D_i$  dan  $D_{min}$  diperoleh dari hasil analisis durasi fonem. Nilai konstanta fonem diperoleh dari persamaan (4) sebagai berikut:

$$K = \frac{D_0 - D_{min}}{D_i - D_{min}} \tag{4}$$

Keterangan:

$K$  = Nilai konstanta fonem konsonan/vokal

$D_0$  = Nilai durasi fonem penutur

$D_{min}$  = Nilai durasi minimal penutur

$D_i$  = Nilai durasi rata-rata penutur

Konstanta fonem ditentukan berdasarkan dua aturan yang digunakan yaitu berdasarkan struktur suku kata dan tipe suku kata. Struktur suku kata yang digunakan merupakan struktur suku kata Bahasa Indonesia yang mengacu pada tabel 2 [25].

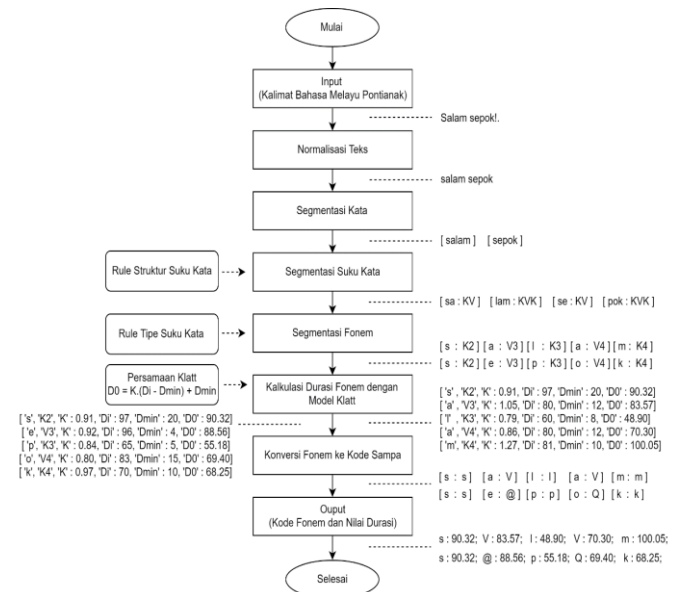
TABEL II  
STRUKTUR SUKU KATA BAHASA INDONESIA DAN TIPE SUKU KATA

Struktur Suku Kata	Tipe Suku Kata	Contoh Kata
V	V1	a V1 - ku
VK	V2; K1	ma - a V2; f K1
KV	K2; V3	bang - s K2; a V3
KVK	K3; V4; K4	a - k K3; a V4; n K4
KKV	K5; K6; V5	t K5; r K6; a V5 - ge - di
KKVK	K7; K8; V6; K9	p K7; r K8; a V6; k K9 - tis
VKK	V7; K10; K11	e V7; k K10; s K11 - tra
KVKK	K12; V8; K13; K14	kon - t K12; e V8; k K13; s K14 - tu - al
KKVKK	K15; K16; V9; K17; K18	t K15; r K16; a V9; n K17; s K18 - por - ta - si

Struktur Suku Kata	Tipe Suku Kata	Contoh Kata
KKKV	K19; K20; K21; V10	s K19; t K20; r K21; a V10 - te - gi
KKKVK	K22; K23; K24; V11; K25	s K22; t K23; r K24; u V11; k K25 - tur

E. Prediksi Durasi Fonem

Tahap ini memprediksi durasi fonem menggunakan model Klatt. Prediksi durasi dijalankan pada *Command Prompt*. Berikut diagram alir yang menggambarkan proses prediksi durasi fonem menggunakan model durasi klatt dapat dilihat pada gambar 3.



Gambar. 3 Diagram Alir Proses Prediksi Model Klatt

F. Pengujian

Pengujian ini menggunakan beberapa metode pengujian, diantaranya pengujian standar kecepatan durasi, pengujian perbandingan durasi berdasarkan durasi fonem terhadap 10 kalimat uji dan durasi kalimat terhadap 100 kalimat uji, *Different Mean Opinion Score* (DMOS) berdasarkan 10 kalimat uji dan *Word Error Rate* (WER) berdasarkan 10 kalimat uji yang diambil dari hasil pengujian perbandingan durasi kalimat.

G. Analisis Hasil Pengujian

Pada tahap ini, hasil pengujian dari prediksi durasi fonem menggunakan model Klatt dianalisis secara keseluruhan untuk mempermudah dalam penarikan kesimpulan.

H. Penarikan Kesimpulan

Kesimpulan dirumuskan berdasarkan tahapan-tahapan yang telah dilakukan sebelumnya yaitu apakah prediksi durasi fonem dengan model Klatt pada Bahasa Melayu Pontianak dapat menghasilkan sintesa ucapan yang jelas.

IV. HASIL DAN PEMBAHASAN

A. Hasil Analisis Durasi Fonem

Hasil analisis durasi fonem dari 576 kalimat diperlukan untuk melakukan prediksi durasi fonem

menggunakan model durasi Klatt. Hasil analisis yang digunakan berupa nilai durasi minimal ( $D_{min}$ ) dan durasi rata-rata ( $D_i$ ) yang terdapat pada tabel 3.

TABEL III  
HASIL ANALISIS DURASI FONEM

Fonem	Jumlah Fonem	Total Jumlah Durasi Fonem (ms)	Durasi Minimal/ $D_{min}$ (ms)	Durasi Maksimal (ms)	Durasi Rata-rata/ $D_i$ (ms)
a	4312	343264	12	848	80
e	602	47730	8	452	79
i	1234	116061	8	580	94
o	853	70642	15	938	83
u	1195	125447	8	586	105
e'	2032	139831	4	729	96
ai	76	9811	25	320	129
au	61	11885	41	500	195
oi	7	2616	121	788	374
b	810	48522	4	551	60
c	157	10914	10	822	70
d	766	50105	7	922	65
f	2	82	39	43	41
g	350	22111	11	180	63
h	479	41189	8	598	86
j	378	23716	13	250	63
k	1992	138694	10	593	70
l	932	55845	8	761	60
m	810	65508	10	808	81
n	1356	116234	9	1234	86
p	754	48915	5	569	65
q	0	0	0	0	0
r	846	56582	3	365	67
s	773	75151	20	677	97
t	1310	81711	4	719	62
v	0	0	0	0	0
w	169	10028	17	146	59
x	0	0	0	0	0
y	183	13061	20	293	71
z	0	0	0	0	0
ng	832	83073	13	470	100
ny	80	7216	30	154	90
Total Jumlah	23351				

B. Hasil Analisis Konstanta Fonem

Hasil analisis konstanta dari 576 kalimat diperlukan untuk memprediksi durasi dengan model Klatt. Hasil nilai konstanta fonem vokal dan konsonan dapat dilihat pada tabel 4 untuk konstanta fonem vokal, tabel 5 dan 6 untuk konstanta fonem konsonan.

TABEL IV  
HASIL ANALISIS KONSTANTA FONEM VOKAL

Tipe K Vokal	Fonem Vokal								
	a	e	i	o	u	e'	ai	au	oi
V1	1,4	1,3	1,0	1,2	0,9	1,9	0,0	0,0	2,6
V2	1,0	1,5	0,6	1,4	1,1	0,5	0,0	0,0	0,0
V3	1,1	1,2	1,2	1,2	1,0	0,9	1,0	0,9	0,7
V4	0,9	1,1	0,7	0,8	0,7	0,6	0,0	0,1	0,0
V5	1,0	0,8	0,9	0,8	0,6	1,0	0,9	0,0	0,0
V6	1,0	0,8	0,7	0,6	0,2	0,7	0,0	0,0	0,0
V7	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
V8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
V9	0,7	0,0	0,4	0,0	0,0	0,0	0,0	0,0	0,0

Tipe K Vokal	Fonem Vokal								
	a	e	i	o	u	e'	ai	au	oi
V10	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
V11	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

TABEL V  
HASIL ANALISIS KONSTANTA FONEM KONSONAN (B - M)

K Konsonan	Fonem Konsonan									
	b	c	d	f	g	h	j	k	l	m
K1	1,2	0,0	0,0	0,0	0,0	1,3	0,0	0,9	1,4	1,5
K2	1,0	0,9	1,0	0,0	0,9	0,6	1,0	1,1	0,9	0,8
K3	1,2	1,0	0,9	0,0	1,0	0,6	1,0	0,8	0,8	0,9
K4	1,7	0,0	1,6	0,0	0,5	1,1	0,0	1,0	2,2	1,3
K5	0,8	0,0	0,0	0,0	0,9	0,0	0,0	0,6	0,0	1,0
K6	0,5	0,6	0,8	0,0	0,8	0,0	0,8	0,7	0,6	0,0
K7	0,7	0,0	0,0	0,0	1,9	0,0	0,0	0,9	0,0	1,4
K8	0,4	0,0	0,6	0,0	0,0	0,0	0,5	0,6	0,5	1,0
K9	0,0	0,0	0,0	0,0	0,0	0,6	0,0	1,1	1,5	1,2
K10	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,1
K11	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K12	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K13	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K14	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K15	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K16	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K17	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,1	0,0	0,0
K18	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
...	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K25	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

TABEL VI  
HASIL ANALISIS KONSTANTA FONEM KONSONAN (N - NG)

K Konsonan	Fonem Konsonan									
	n	p	r	s	t	v	w	y	ng	ny
K1	1,3	1,1	1,5	1,5	1,5	0,0	0,0	0,0	1,6	0,0
K2	0,7	1,0	0,9	0,9	1,0	0,0	1,2	1,5	0,6	0,9
K3	0,7	0,8	0,7	0,9	0,8	1,5	1,0	0,9	0,7	1,1
K4	1,1	1,3	1,3	1,2	1,3	0,0	0,0	0,0	1,1	0,0
K5	1,0	1,0	0,8	1,3	0,7	0,0	0,0	0,0	0,6	0,0
K6	0,8	0,7	0,5	0,0	0,6	0,0	0,5	0,0	0,0	0,0
K7	1,0	0,4	1,1	1,7	0,8	0,0	0,0	0,0	0,0	0,0
K8	0,0	0,4	0,9	0,0	0,6	0,0	0,0	0,0	0,0	0,0
K9	1,2	1,1	1,5	0,0	0,9	0,0	0,0	0,0	0,9	0,0
K10	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K11	0,0	0,0	0,0	1,4	0,0	0,0	0,0	0,0	0,0	0,0
K12	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K13	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K14	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K15	0,0	0,0	0,0	0,0	0,7	0,0	0,0	0,0	0,0	0,0
K16	0,0	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K17	0,4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K18	0,0	0,0	0,0	1,1	0,0	0,0	0,0	0,0	0,0	0,0
...	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
K25	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

C. Implementasi Model Durasi Klatt

Prediksi durasi fonem menggunakan model durasi Klatt ditampilkan pada *command prompt* seperti yang terlihat pada gambar 4.

```

C:\Users\N.K\python-klatt>python sentence_to_SAMPA.py "Salam sepokl."
{ 'phoneme': 's', 'type': 'K2', 'K': 0.9, 'DMin': 20, 'Di': 97, 'D0': 89.3}
{ 'phoneme': 'a', 'type': 'V3', 'K': 1.05, 'DMin': 12, 'Di': 80, 'D0': 83.4}
{ 'phoneme': 'l', 'type': 'K3', 'K': 0.79, 'DMin': 8, 'Di': 60, 'D0': 49.08}
{ 'phoneme': 'a', 'type': 'V4', 'K': 0.9, 'DMin': 12, 'Di': 80, 'D0': 73.2}
{ 'phoneme': 'm', 'type': 'K4', 'K': 1.27, 'DMin': 10, 'Di': 81, 'D0': 100.17}
{ 'phoneme': 's', 'type': 'K2', 'K': 0.9, 'DMin': 20, 'Di': 97, 'D0': 89.3}
{ 'phoneme': 'e', 'type': 'V3', 'K': 0.92, 'DMin': 4, 'Di': 96, 'D0': 88.64}
{ 'phoneme': 'p', 'type': 'K3', 'K': 0.84, 'DMin': 5, 'Di': 65, 'D0': 55.4}
{ 'phoneme': 'o', 'type': 'V4', 'K': 0.8, 'DMin': 15, 'Di': 83, 'D0': 69.4}
{ 'phoneme': 'k', 'type': 'K4', 'K': 1.0, 'DMin': 10, 'Di': 70, 'D0': 70.0}

s 89.3
V 83.4
l 49.08
V 73.2
m 100.17
s 89.3
e 88.64
p 55.4
o 69.4
k 70.0
    
```

Gambar. 4 Hasil prediksi durasi fonem model Klatt

D. Hasil dan Analisis Pengujian

1. Pengujian Standar Kecepatan Durasi

Pengujian standar kecepatan durasi dilakukan untuk menentukan kualitas kecepatan kalimat yang diucapkan. Tujuan pengujian ini yaitu kemampuan untuk menerima kecepatan ucapan (kejelasan dan pemahaman) yang masih bisa diterima oleh pendengar. Pengujian ini dilakukan terhadap satu kalimat yang digunakan sebagai standar untuk penentuan kualitas kecepatan ucapan. Audio rekaman kalimat ini dipercepat -10% hingga -50% dan diperlambat 10% hingga 50%. Analisis hasil uji terdapat pada tabel 7.

TABEL VII  
ANALISIS HASIL UJI STANDAR KECEPATAN DURASI

Range Hasil Uji Responden	Kecepatan Durasi Ucapan (%)	Rata-rata Hasil Uji Responden (%)
80% - 100% [sangat baik]	-10%	100%
	10%	100%
	-20%	93%
60% - 79% [baik]	20%	83%
	-30%	63%
40% - 59% [cukup baik]	-40%	47%
	30%	47%
20% - 39% [tidak baik]	-	-
0% - 19% [sangat tidak baik]	-50%	13%
	50%	0%
	40%	0%

Tabel 5 menunjukkan hasil dari pengujian standar kecepatan durasi diperoleh hasil bahwa responden dapat menerima ucapan yang dipercepat maupun diperlambat dalam rentang kecepatan -20% hingga 20% dalam kategori sangat baik.

2. Pengujian Perbandingan Durasi

Pengujian perbandingan durasi dilakukan dua kali perbandingan. Pengujian ini bertujuan untuk membandingkan durasi penutur dan durasi klatt. Menghitung perbandingan durasi dapat dilakukan menggunakan rumus berikut:

$$\text{Persentase perbandingan} = \frac{\text{durasi klatt} - \text{durasi penutur}}{\text{durasi penutur}} \times 100$$

Keterangan hasil persentase perbandingan nilai durasi.  
 Nilai durasi (+) : durasi Klatt lebih lambat dari durasi penutur  
 Nilai durasi (-) : durasi Klatt lebih cepat dari durasi penutur  
 Nilai durasi (0) : durasi Klatt sama dengan durasi penutur

a. Berdasarkan Durasi Fonem Terhadap 10 Kalimat Uji

Pengujian ini, terlebih dahulu dilakukan analisis perbandingan nilai durasi berdasarkan kalimat. Hasil perbandingan durasi Klatt dan durasi penutur terhadap 10 kalimat uji dapat dilihat pada tabel 8. Kemudian 10 kalimat uji diuraikan lagi perbandingan durasi berdasarkan durasi fonem tiap-tiap kalimat. Hasil analisis perbandingan durasi fonem dapat dilihat pada tabel 9.

TABEL VIII  
HASIL PERBANDINGAN DURASI KALIMAT PADA 10 KALIMAT UJI

Range Kecepatan Durasi Ucapan	Jumlah Kalimat	Persentase Kalimat
41% - 50%	-	-
31% - 40%	-	-
21% - 30%	2	20%
11% - 20%	-	-
1% - 10%	3	30%
0%	1	10%
(-1%) - (-10%)	3	30%
(-11%) - (-20%)	1	10%
(-21%) - (-30%)	-	-
(-31%) - (-40%)	-	-
(-41%) - (-50%)	-	-
Total Jumlah	10	100%

TABEL IX  
ANALISIS HASIL PERBANDINGAN DURASI FONEM

Range Kecepatan Durasi Ucapan	Jumlah Fonem	Total Fonem	Persentase Total Fonem
(-10%) - (-10%)	320	42	13%
(-20%) - (-10%)	320	73	23%
(-20%) - (-20%)	320	95	30%
(-30%) - (-20%)	320	149	47%
(-30%) - (-30%)	320	166	52%
(-40%) - (-30%)	320	211	66%
(-40%) - (-40%)	320	223	70%
(-50%) - (-40%)	320	270	84%
(-50%) - (-50%)	320	278	87%

Hasil analisis pengujian perbandingan durasi berdasarkan fonem dilihat dari dua sisi perbandingan terhadap 10 kalimat uji.

- 1) Hasil perbandingan durasi berdasarkan durasi kalimat  
 Tabel 8 diperoleh hasil kecepatan durasi yang dapat diterima oleh pendengar berdasarkan hasil uji standar kecepatan durasi pada tabel 7 dalam range kecepatan durasi 20% hingga -20% pada kategori sangat baik adalah sebanyak 80% kalimat dari 10 kalimat uji.
- 2) Hasil perbandingan durasi berdasarkan durasi fonem  
 Tabel 9 menampilkan analisis hasil perbandingan durasi fonem yang masuk dalam kategori sangat baik berdasarkan hasil pengujian standar kecepatan durasi. Maka dari hasil pengujian perbandingan durasi Klatt dan durasi penutur berdasarkan durasi fonem, kecepatan durasi fonem dari model Klatt yang termasuk dalam range kategori sangat baik sebesar 95 fonem (vokal maupun

konsonan) dengan persentase 30% fonem dari 320 fonem terhadap 10 kalimat uji.

b. *Pengujian Perbandingan Durasi Berdasarkan Durasi Kalimat Terhadap 100 Kalimat Uji*

Hasil analisis perbandingan durasi berdasarkan durasi kalimat dapat dilihat pada tabel 10.

TABEL X  
HASIL ANALISI PENGUJIAN PERBANDINGAN NILAI DURASI KALIMAT

Range Kecepatan Durasi Ucapan	Jumlah Kalimat	Persentase Kalimat
41% - 50%	-	-
31% - 40%	-	-
21% - 30%	3	3%
11% - 20%	9	9%
1% - 10%	34	34%
0%	4	4%
(-1%) - (-10%)	30	30%
(-11%) - (-20%)	17	17%
(-21%) - (-30%)	3	3%
(-31%) - (-40%)	-	-
(-41%) - (-50%)	-	-
Total Jumlah	100	100%

Tabel 10 diperoleh hasil analisis pengujian perbandingan nilai durasi kalimat yang kecepatan durasinya dapat diterima oleh pendengar berdasarkan hasil uji standar kecepatan durasi pada tabel 7 dalam range kecepatan durasi -20% hingga 20% pada kategori sangat baik adalah sebanyak 94% kalimat dari 100 kalimat uji.

KESIMPULAN

Berdasarkan hasil analisis dan pengujian terhadap prediksi durasi yang dilakukan menggunakan model Klatt pada bahasa Melayu Pontianak, maka dapat ditarik kesimpulan sebagai berikut.

- Berdasarkan hasil pengujian standar kecepatan durasi, pendengar dapat menerima ucapan yang dipercepat maupun diperlambat dalam rentang kecepatan -20% hingga 20% dalam kategori sangat baik.
- Berdasarkan pengujian perbandingan durasi. Hasil perbandingan nilai durasi berdasarkan durasi fonem pada 10 kalimat uji dari 320 fonem diperoleh 95 fonem atau 30% yang termasuk dalam rentang kecepatan durasi ucapan pada kategori sangat baik. Perbandingan jumlah fonem vokal dan konsonan terhadap 95 fonem pada kategori sangat baik yaitu fonem vokal sebanyak 42.1% (40 fonem) dan fonem konsonan sebanyak 57.8% (55 fonem). Hasil perbandingan durasi kalimat terhadap 100 kalimat uji diperoleh 94% kalimat yang termasuk dalam rentang kecepatan durasi ucapan pada kategori sangat baik.
- Berdasarkan pengujian perbandingan model durasi Klatt yang menggunakan aturan suku kata dan tipe suku kata dapat memprediksi durasi fonem total pada ucapan sintesa bahasa Melayu Pontianak namun masih perlu penelitian lanjut untuk memprediksi durasi masing-masing fonem.

REFERENSI

- Arman, a. A. 2008, Maret 14. Konversi Teks ke Ucapan. Retrieved from Teknologi Bahasa: <https://teknologibahasa.wordpress.com/category/text-to-speech/>
- Benesty, J., Sondhi, M. M., dan Huang, Y. 2008. *Springer Handbook of Speech Processing*. Berlin: Springer-Verlag.
- Salita Ulitia Prini dan Ary Setijadi Prihatmanto. 2015. Penambahan Emosi Menggunakan Metode Manipulasi Prosodi Untuk Sistem Text To Speech Bahasa Indonesia. Research Gate, Halaman 7.
- Tritoasmoro, I. I. 2006. Text To Speech Bahasa Indonesia Menggunakan Concatenation Synthesizer Berbasis Fonem. Seminar Nasional Sistem dan Informatika, Halaman 3.
- Rupayani, N. L., Antara Kesiman, M. W., dan Agus Wirawan, I. M. 2011. Pengembangan Media Text To Speech Untuk Aplikasi Tokoh Virtual. Prosiding Seminar Nasional Pendidikan Teknik Informatika, Halaman 2.
- Mentari Putri, M. D. 2016. Text to Speech Engine Generik Bahasa Bugis Wajo. Research Gate, Halaman 12.
- Guntoro, Erwin. 2019. Development of Text-to-Speech for Malay Pontianak Dialect Based Hidden Markov Model with Phoneme. TELKOMNIKA.
- Kamiludin, M. I. 2017. Prediksi Jeda Pada Ucapan Bahasa Melayu Pontianak dengan Menggunakan Metode Shallow Parsing. Jurnal Sistem dan Teknologi Informasi JUSTIN, Halaman 1.
- Ricky. 2011. Kamus Digital Dengar Tulis Ucap Bahasa Indonesia. eprints, 1.
- Kasparaitis, Piju; dan Beniušė, Margarita. 2016. *Automatic Parameters Estimation of the D. Klatt Phoneme Duration Model*. ResearchGate.
- Rao, K. S. 2012. *SpringerBriefs in Electrical and Computer Engineering*. Kharagpur: Amy Neustein.
- Setiawan, A. F. 2016. Text To Speech Bahasa Indonesia Menggunakan Metode *Diphone* Concatenation. Seminar Nasional Inovasi dan Aplikasi Teknologi di Industri, Halaman B.38.
- Hasan, A., Lapoliwa, H., dan Soenjono, D. 2000. Tata Bahasa Baku Bahasa Indonesia Edisi Ketiga. Jakarta: Departemen Pendidikan dan Kebudayaan. Republik Indonesia.
- Setyaningsih, Y., dan R, K. R. 2014. Fonologi Bahasa Indonesia : Mengkaji Tata Bunyi Dalam Perspektif Linguistik Edukasi. Yogyakarta: Universitas Sanata Dharma.
- Arief, H. A. 2018, Juni 24. Deskripsi Bunyi Bahasa Melayu Pontianak. Retrieved from Repository Universitas Airlangga: <http://repository.unair.ac.id/id/eprint/14435>
- Dwi Putra, E. P., dan Hartati, U. 2016. Perbedaan Bahasa Melayu Pontianak Kalimantan Barat dengan Bahasa Indonesia Standar. CARAKA, Halaman 149-155.
- Mega Ratnawati, N. K. 2018, Mei 26. Fonem Suprasegmental / Ciri-ciri Prosodi. Retrieved from Jani Indonesia : <https://janindonesia.files.wordpress.com/2016/05/fonem-suprasegmental1.pdf>
- Klatt, D. H. 1987, May 1. *Review of Text-To-Speech Conversion for English*. Cambridge: Acoustical Society of America.
- Campbell, N. 2000. *Timing In Speech: A Multi-Level Process*. In M. Horne, Text, Speech and Language Technology p. Halaman 289. Sweden: Kluwer Academic Publishers in 2000.
- Öztürk, Ö. 2014. *Modeling Phoneme Durations and Fundamental Frequency Contours in Turkish Speech*. Research Gate, Halaman 26.
- Klatt, D. H. 1973, June 26. *Interaction Between two factors that influence vowel duration*: Acoustical Society of America.
- Simões, A. 1990. *Predicting Sound Segment Duration in Connected Speech: An Acoustical Study of Brazilian Portuguese*. International Speech Communication Association ISCA-Speech, Halaman 175.
- Priyowidodo, S., dan Irwan, D. 2010. Speech Synthesizer Bahasa Indonesia Berbasis *Diphone* MBROLA. Seminar Nasional Ilmu Komputer SNIKOM, Halaman 305.
- Narullita, A. 2009. Aplikasi Pensintesa Ucapan Bahasa Indonesia Sebagai Pembaca Email. Eprints, Halaman 1.
- Sanjoko, Y. 2015. Perbandingan Karakteristik Fonem Bahasa Indonesia Dengan Bahasa Lasalimu. KANDAI, Halaman 61.