

BAB I

PENDAHULUAN

1.1. Latar Belakang

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* (Arman, 2008; Benesty, Sondhi, dan Huang, 2008; Salita dan Ary, 2015). *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* (Tritoasmoro, 2006).

Saat ini telah dikembangkan *diphone database* "id1" Bahasa Indonesia yang mana program ini dibuat menggunakan aksent Bahasa Indonesia (Rupayani, dkk, 2011:2). *Diphone database* "id1" dibuat oleh Aray 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). 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 (2019). Namun ucapan yang dihasilkan masih belum baik, dikarenakan korpus 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 (Kamiludin, 2017). 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 (Ricky, 2011). 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 disintesis untuk TTS suatu bahasa. Menurut Pijus Kasparaitis dan Margarita Beniuse (2016:573) menyatakan 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.

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 (Rao, 2012). Sehingga model durasi berbasis aturan ini cocok dengan penggunaan *diphone database*. Beberapa penelitian telah menggunakan model durasi Klatt untuk memprediksi durasi fonem. Menurut Prof. Jacob Benesty, dkk (2008:477) model Klatt telah dikembangkan untuk beberapa bahasa asing seperti Bahasa Inggris tahun 1987, Swedia tahun 1986, Jerman tahun

1994 dan Francis tahun 1987. 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.

1.2. Rumusan Masalah

Berdasarkan hal yang telah disampaikan diatas maka rumusan masalah yang muncul dalam penelitian ini adalah bagaimana menerapkan model durasi Klatt agar dapat digunakan untuk memprediksi durasi fonem yang menggunakan model *diphone* "id1" Bahasa Indonesia untuk sintesa ucapan Bahasa Melayu Pontianak.

Untuk menerapkan model durasi Klatt dengan persamaan berikut.

$$D_0 = K \cdot (D_i - D_{\min}) + D_{\min} \quad (1.1)$$

Dimana nilai D_i (durasi rata-rata fonem) dan D_{\min} (durasi minimal fonem) harus dianalisis dan nilai K (nilai konstanta fonem) harus ditentukan dengan menggunakan aturan suku kata dan tipe suku kata.

1.3. Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah untuk mengetahui apakah model durasi Klatt yang menggunakan aturan Suku Kata dan Tipe Suku Kata dapat memprediksi durasi fonem sintesa ucapan Bahasa Melayu Pontianak.

1.4. Pembatasan Masalah

Pembatasan masalah dari penelitian yang dilakukan adalah sebagai berikut.

1. Korpus Bahasa Melayu Pontianak yang digunakan untuk melakukan analisis model durasi Klatt sebanyak 576 kalimat.
2. Kata atau istilah asing selain Bahasa Melayu Pontianak tidak disertakan sebagai korpus.

3. *Database* ucapan Bahasan Melayu Pontianak menggunakan *diphone* Bahasa Indonesia (Id1).
4. Algoritma Klatt tidak melakukan proses konversi angka, simbol, waktu dan tanggal ke dalam bentuk alphabet.
5. Sintesa ucapan menggunakan *Mbrola Engine* sebagai pembangkit ucapan untuk melakukan pengujian.
6. Kualitas sintesa ucapan yang akan diuji adalah faktor *intelligibility* (kejelasan) kalimat yang diucapkan.
7. Pengujian ucapan fokus pada penelitian durasi fonem.

1.5. Sistematika Penulisan Skripsi

Adapun sistematika penulisan dari tugas akhir ini tersusun atas 5 (lima) bab yang terdiri dari Bab I Pendahuluan, Bab II Tinjauan Pustaka, Bab III Metodologi Penelitian dan Perancangan Sistem, Bab IV Hasil dan Analisis, serta Bab V Penutup.

Bab I Pendahuluan adalah bab yang berisi latar belakang penelitian, perumusan masalah, tujuan penelitian, pembatasan masalah, dan sistematika penulisan.

Bab II Tinjauan Pustaka adalah bab yang berisi landasan teori berkaitan dengan penelitian yang akan dilakukan. Beberapa teori-teori yang terkait adalah teori *Text to Speech* (TTS), fonem Bahasa Indonesia, perbedaan fonologi Bahasa Melayu Pontianak dan Bahasa Indonesia, pengertian durasi dan model durasi Klatt sebagai metode untuk melakukan penelitian, pengertian alat-alat yang diperlukan dan pengujian penelitian yaitu *Mean Opinion Score* (MOS), *The Degradation Mean Opinion Score* (DMOS), *Word Error Rate* (WER), pengujian perbandingan nilai durasi.

Bab III Metodologi Penelitian dan Perancangan Sistem adalah bab yang berisi tentang Bahan Penelitian, Alat yang Dipergunakan, Metode Penelitian, Variabel atau Data, Diagram Alir Penelitian, Arsitektur Sistem, Rencana Pengujian dengan menggunakan metode MOS, DMOS, WER, dan pengujian perbandingan nilai durasi.

Bab IV Hasil dan Analisis adalah bab yang berisi penjelasan mengenai implementasi pada proses model durasi, *screenshoot* proses analisis, serta analisis hasil uji coba. Setiap hasil yang disajikan akan dilakukan analisis untuk mengarah kepada suatu kesimpulan.

Bab V Penutup adalah bab yang berisi kesimpulan dari penelitian yang telah dilakukan dan saran/rekomendasi untuk perbaikan, pengembangan atau kesempurnaan / kelengkapan penelitian yang telah dilakukan.

BAB II

TINJAUAN PUSTAKA

2.1 Studi Literatur

Pijus Kasparaitis dan Margarita Beniušė (2016), *Department of Computer Science II, Faculty of Mathematics and Informatics, Vilnius University, Lithuania* dalam penelitian yang berjudul *Automatic Parameters Estimation of the D.Klatt Phoneme Duration Model* menjelaskan bahwa tujuan dari penelitian tersebut adalah membuat algoritma iteratif baru untuk menghitung parameter model Klatt menggunakan korpus rekaman bahasa Lithuania yang merupakan salah satu negara dan bahasa yang digunakan di bagian Eropa timur. Korpus direkam oleh 3 orang penutur. Korpus yang digunakan sebanyak 5000 kalimat dengan jumlah fonem sebanyak 160000 fonem. Kualitas prediksi durasi diuji menggunakan *Root Mean Square Error*, *Mean Absolute Error* dan *Correlation Coefficient* dengan beberapa kategori penilaian yaitu durasi rata-rata dari hasil prediksi Klatt dan durasi rata-rata dari 3 penutur. Hasil dari pengujian *Root Mean Square Error* dari kedua kategori perbandingan antara penutur dan model Klatt adalah 19.71% untuk fonem vokal dan 13.69% untuk fonem konsonan. Hasil uji *Mean Absolute Error* yaitu 19.69% untuk fonem vokal dan 11.42% untuk fonem konsonan. Sedangkan hasil *Correlation Coefficient* adalah 0.80 untuk vokal dan 0.75 untuk konsonan.

W.N. Campbell (1988), IBM (UK) *Scientific Center, Winchester, England* dalam penelitian yang berjudul *Speech-Rate Variation and The Prediction of Duration* menjelaskan bahwa tujuan dari penelitian ini adalah untuk menganalisis durasi ucapan dengan menggunakan korpus bahasa Inggris britis. Durasi dianalisis menggunakan metode model Klatt. Pengujian persentase ucapan menggunakan SPRATE dengan hasil rata-rata 100.2% dari 3959 suku kata. Pengujian SPRATE yang dilakukan berdasarkan tipe suku kata diantaranya kata benda, kata sifat, dan kata keterangan.

Caren Brinckmann & Jürgen Trouvain (2001), *Institute of Phonetics, University of the Saarland, Saarbrücken, Germany* dalam penelitian yang berjudul *On the Role of Duration Prediction and Symbolic Representation for the Evaluation of Synthetic Speech*

menjelaskan bahwa tujuan dari penelitian ini yaitu untuk mengetahui apakah perbedaan durasi yang diprediksi oleh model durasi dengan yang dihasilkan dalam durasi penutur secara pengamatan relevan dengan durasi sintesa. Korpus yang digunakan dalam penelitian menggunakan bahasa Jerman. Perbandingan durasi diuji dengan dua model prediksi durasi yaitu aturan yang dikembangkan oleh Klatt dan *Classification and Regression Trees* (CART). Hasil uji menggunakan SPRATE adalah persentase rata-rata untuk semua jenis suku kata diantaranya 100.3% untuk suku kata tanpa tekanan, 98.3% untuk suku kata dengan tekanan, dan 106.7% untuk tekanan tanpa aksen.

Antônio R. M. Simões (1990), *Department of Spanish and Portuguese The University of Kansas at Lawrence, USA* dalam penelitian ini berjudul *Predicting Sound Segment Duration in Connected Speech: an Acoustical Study of Brazilian Portuguese* menjelaskan bahwa tujuan dari penelitian ini yaitu untuk memberikan BP (Brazilian Portuguese) dengan model yang sama efektifnya dengan model D.H. Klatt (1976) untuk Bahasa Inggris Amerika. Korpus yang digunakan menggunakan bahasa brasil yang dituturkan oleh orang portugis. Korpus yang digunakan sebanyak lebih dari seribu kata. Dalam penelitian ini menggunakan 12 aturan model Klatt untuk menentukan nilai konstanta durasi fonem vokal. Aturan dirancang untuk memprediksi pemendekan dan pemanjangan 3 vokal yaitu a, i, dan u. Hasil analisis dilihat dari hasil median bukan rata-rata. Hasil median untuk durasi tanpa tanda (a,i,u) yaitu 60ms sebanyak 1275 fonem. Hasil median untuk durasi vokal posisi tertekan yaitu 120ms sebanyak 239 fonem dan tidak tertekan 52ms sebanyak 1036 fonem.

Tabel 2.1 Perbandingan Kajian Penelitian Terkait

No	Penulis/Judul	Metode	Hasil
1	<ul style="list-style-type: none"> • Pijus Kasparaitis dan Margarita Beniušė (2016), <i>Vilnius University, Lithuania</i> • <i>Automatic Parameters Estimation of the</i> 	<ul style="list-style-type: none"> • Penelitian ini mengusulkan algoritma iterative baru untuk estimasi otomatis pada model durasi 	<ul style="list-style-type: none"> • <i>Mean Absolute Error:</i> 19.69% vokal dan 11.42% konsonan. • <i>Root Mean Square error:</i> 19.71% vokal dan 13.69% konsonan. • <i>Correlation coefficient:</i> 0.80 vokal dan 0.75

No	Penulis/Judul	Metode	Hasil
	<i>D. Klatt Phonem Duration Model</i>	klatt. <ul style="list-style-type: none"> Data yang digunakan berupa korpus rekaman audio bahasa Lituania. 	konsonan.
2.	<ul style="list-style-type: none"> W.N. Campbell (1988), <i>Scientific Center, Winchester, England</i> <i>Speech-Rate Variation and The Prediction of Duration</i> 	<ul style="list-style-type: none"> Penelitian ini melakukan analisis terhadap nilai variasi tingkat bicara dan menganalisis nilai durasi minimal dan nilai durasi rata-rata. Korpus yang digunakan adalah teks bahasa Inggris British. Menggunakan model durasi Klatt untuk memprediksi durasi. 	<ul style="list-style-type: none"> Persentase rata-rata untuk semua jenis suku kata diantaranya 100.3% untuk suku kata tanpa tekanan, 98.3% untuk suku kata dengan tekanan, dan 106.7% untuk tekanan tanpa aksen.
3	<ul style="list-style-type: none"> Caren Brinckmann & Jürgen Trouvain (2001), <i>Institute of Phonetics, University of the Saarland, Saarbrücken, Germany</i> <i>On the Role of Duration Prediction and Symbolic Representation for the Evaluation of Synthetic Speech</i> 	<ul style="list-style-type: none"> Membandingkan durasi penutur dengan durasi hasil sintesa. Prediksi durasi menggunakan dua model yaitu aturan yang dikembangkan oleh Klatt dan CART. 	<ul style="list-style-type: none"> Secara signifikan hasil durasi CART lebih baik dari pada Klatt. Skor secara signifikan, untuk CART 108 dan Klatt 72.
4.	<ul style="list-style-type: none"> Antônio R. M. Simões (1990), <i>Department of Spanish and Portuguese The</i> 	<ul style="list-style-type: none"> Menentukan kenaikan dan penurunan nilai konstanta untuk fonem vokal. 	<ul style="list-style-type: none"> Median durasi vokal tanpa tekanan: 52ms. Median durasi vokal dengan tekanan: 120ms.

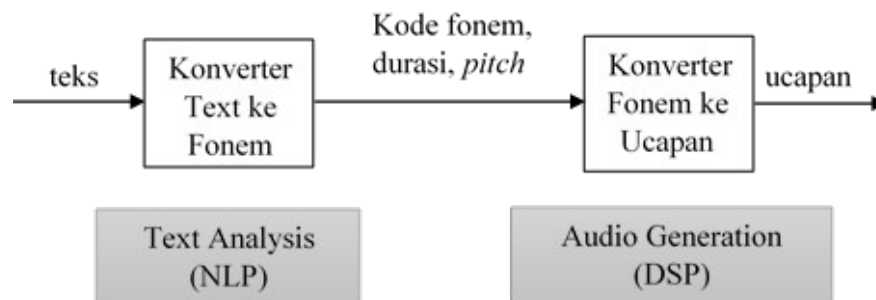
No	Penulis/Judul	Metode	Hasil
	<i>University of Kansas at Lawrence, USA</i> • <i>Predicting Sound Segment Duration in Connected Speech: an Acoustical Study of Brazilian Portuguese</i>	• Korpus yang digunakan bahasa brasil yang dituturkan oleh orang portugis lebih dari seribu kata. • Menentukan nilai konstanta (K) durasi berdasarkan aturan Klatt.	• Dan median durasi vokal tanpa tanda: 60ms.

2.2 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 (Setiawan, 2016). Suatu sistem *text to speech* prinsipnya terdiri dari dua subsistem (Arman, 2008), yaitu:

2.1 Bagian konverter teks ke fonem (text to phoneme).

2.2 Bagian konverter fonem ke ucapan (phoneme to speech).



Gambar 2.1 Sistem Text To Speech

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. 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. Menurut Arry Akhmad Arman (2008), menyatakan bahwa dua teknik yang banyak digunakan adalah *formant synthesizer*, serta *diphone concatenation* (Prini & Prihatmanto, 2015). 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.

Formant synthesizer berkerja berdasarkan suatu model matematis yang akan melakukan komputasi untuk menghasilkan sinyal ucapan yang diinginkan. *Synthesizer* jenis ini telah lama digunakan pada berbagai aplikasi. Walaupun dapat menghasilkan ucapan dengan tingkat kemudahan interpretasi yang baik, *synthesizer* ini tidak dapat menghasilkan ucapan dengan tingkat kealamian yang tinggi.

Synthesizer yang menggunakan teknik *diphone concatenation* bekerja dengan cara menggabung-gabungkan segmen-segmen bunyi yang telah direkam sebelumnya. Setiap segmen berupa *diphone* (gabungan dua buah fonem). *Synthesizer* jenis ini menghasilkan bunyi ucapan dengan tingkat kealamian (naturalness) yang tinggi.

Konversi dari teks ke fonem sangat dipengaruhi oleh aturan-aturan yang berlaku dalam suatu bahasa. Pada prinsipnya proses ini melakukan konversi dari simbol-simbol tekstual menjadi simbol-simbol fonetik yang merepresentasikan unit bunyi terkecil dalam suatu bahasa. Setiap bahasa memiliki aturan cara pembacaan dan cara pengucapan teks yang sangat spesifik. Hal ini menyebabkan implementasi unit *converter* teks ke fonem menjadi sangat spesifik terhadap suatu bahasa. Untuk dapat menghasilkan ucapan yang lebih alami, ucapan yang dihasilkan harus memiliki intonasi (prosody). Prosodi adalah perubahan nilai *pitch* (frekuensi dasar) selama pengucapan kalimat dilakukan atau *pitch* sebagai fungsi waktu. Informasi pembentuk prosodi berupa data-data *pitch* serta durasi pengucapan untuk setiap fonem yang dibangkitkan.

Konverter fonem ke ucapan berfungsi untuk membangkitkan sinyal ucapan berdasarkan kode-kode fonem yang dihasilkan dari proses sebelumnya. Sub sistem ini harus memiliki pustaka setiap unit ucapan dari suatu bahasa. Pada

sistem yang menggunakan teknik *diphone concatenation*, sistem harus didukung oleh suatu *diphone database* yang berisi rekaman segmen-segmen ucapan yang berupa diphone. Ucapan dalam suatu bahasa dibentuk dari satu set yang mungkin berbeda untuk setiap bahasa, oleh karena itu setiap bahasa harus dilengkapi dengan diphone database yang berbeda.

2.3 Fonem Bahasa Indonesia

Menurut Hasan dkk (2000), menyatakan bahwa 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/, /ŋ/, /s/, /h/, /r/, /l/, /w/, dan /y/. Kemudian terdapat penambahan empat fonem yang berasal dari bahasa asing yaitu /x/, /z/, /f/, /ʃ/ (Setyaningsih dkk, 2014). Kontras fonem konsonan bahasa Indonesia dapat dilihat pada tabel 2.2 dan kontras fonem vokal bahasa Indonesia dapat dilihat pada tabel 2.3 (Sanjoko, 2015).

Tabel 2.2 Fonem konsonan Bahasa Indonesia

No	Kontras konsonan			Contoh		
1	/p/	↔	/b/	[pola]	↔	[bola]
2	/k/	↔	/g/	[kali]	↔	[gali]
3	/c/	↔	/j/	[baca]	↔	[baja]
4	/t/	↔	/d/	[pətaŋ]	↔	[pədaŋ]
5	/m/	↔	/n/	[makam]	↔	[makan]
6	/n/	↔	/ŋ/	[saran]	↔	[saraŋ]
7	/l/	↔	/r/	[ajal]	↔	[ajar]
8	/ñ/	↔	/n/	[ñona]	↔	[nona]
9	/s/	↔	/ʃ/	[sah]	↔	[ʃah]
10	/x/	↔	/k/	[tarix]	↔	[tarik]
11	/f/	↔	/p/	[kafan]	↔	[kapan]
12	/s/	↔	/z/	[seni]	↔	[zeni]
13	/y/	↔	/r/	[sayan]	↔	[saraŋ]
14	/w/	↔	/s/	[wayan]	↔	[sayan]
15	/s/	↔	/h/	[sama]	↔	[hama]

Tabel 2.3 Fonem Vokal Bahasa Indonesia

No	Kontras Vokal			Contoh		
1	/e/	↔	/i/	[sekat]	↔	[sikit]
2	/e/	↔	/o/	[elok]	↔	[olok]
3	/a/	↔	/ə/	[baca]	↔	[kəraŋ]

2.4 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 (Arief, 2018).

Dalam banyak kosakata, bahasa Melayu Pontianak hampir sama dengan bahasa Indonesia. Hal ini tidak terlalu mengherankan, karena bahasa Indonesia memang berakar dari bahasa Melayu (Novianti, 2011).

Contoh:

1. Besok kita libur (bahasa Indonesia)
2. *Besok kite libor* (bahasa Melayu Pontianak)

Perbedaan mencolok antara bahasa Melayu Pontianak dan bahasa Indonesia adalah dalam bahasa Melayu Pontianak lebih banyak menggunakan vokal /e/ pada akhir suku kata. Perbedaan ini merujuk pada tabel 2.4.

Tabel 2.4 Perbedaan bahasa Melayu Pontianak dan bahasa Indonesia

No	Bahasa Melayu Pontianak	Bahasa Indonesia
1	<i>Semue</i>	Semua
2	<i>Pade</i>	Pada
3	<i>Kite</i>	Kita
4	<i>Suke</i>	Suka
5	<i>Kate</i>	Kata
6	<i>Bahase</i>	Bahasa
7	<i>Same</i>	Sama
8	<i>Surge</i>	Surge
9	<i>Celane</i>	celana

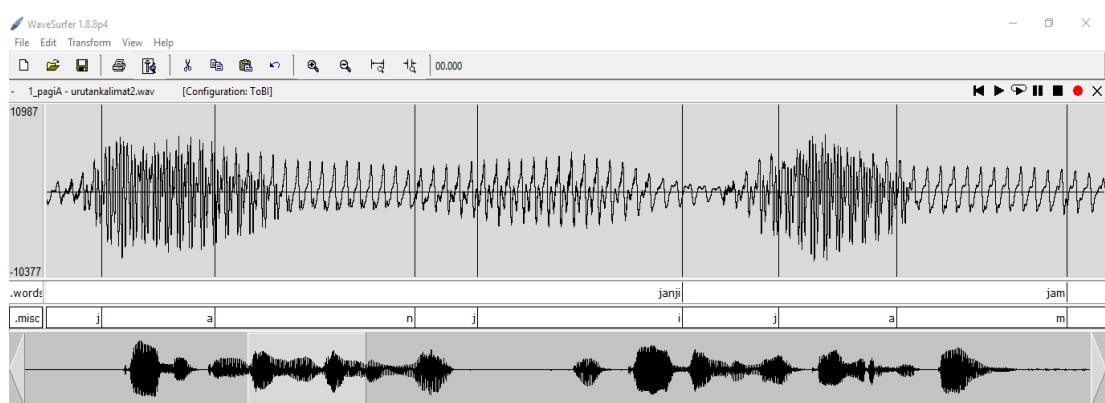
Secara fonologis, bahasa Melayu Pontianak dan bahasa Indonesia memiliki perbedaan teratur dan perbedaan tidak teratur (Dwi Putra & Hartati, 2016). Perbedaan teratur meliputi [a]-[ə], [i]-[ɛ], [u]-[o], [r]-[k], [Ø]-[k]. Terdapat enam puluh tujuh (67) kata perbedaan teratur dalam bahasa Indonesia standar 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. Contohnya [kecil] - [kecik], [kutu]-[gutu], dan [tidak]-[tadak].

2.5 Durasi

Durasi adalah panjang pendeknya waktu yang diperlukan untuk mengucapkan suatu kata sehingga pendengar dapat mengerti kata yang diucapkan. Menurut Ni Kadek Mega Ratnawati (2018), durasi sering juga disebut kuantitas (quantity) karena menyangkut jumlah atau lamanya waktu yang digunakan untuk pengucapan sebuah bunyi. Jadi, durasi berkaitan dengan lama atau tidaknya sebuah bunyi diucapkan. Contohnya kata “Tinggi” dipisah menjadi /ting/ dan /gi/. Masing-masing bagian kata dapat diucapkan dalam waktu yang sama. Tetapi dapat terjadi apabila ada pembicara yang mengucapkan bagian kata /ting/ lebih lama dari bagian kata /gi/ atau sebaliknya. Dalam penelitian ini, panjang pendeknya durasi asli setiap fonem tergantung pada ucapan penutur ahli bahasa Melayu Pontianak dengan korpus yang telah disediakan yang direkam dan disimpan dalam format *.wav.

2.6 Wavesurfer

Wavesurfer adalah *software* yang digunakan untuk menganalisis ucapan atau suara. Penggunaan *wavesurfer* dalam penelitian ini untuk melakukan pelabelan dan penandaan fonem pada *file* suara yang telah direkam. Penandaan dan pelabelan dilakukan pada anotasi “.misc” di *wavesurfer* untuk mendapatkan nilai durasi dari setiap fonem yang diucapkan. Contoh penandaan yang dilakukan pada kata “janji” dan “jam” dapat dilihat pada gambar 2.2.



Gambar 2.2 Penandaan dan Pelabelan *Wavesurfer*

Setiap fonem yang ada pada panel “.misc” nantinya akan dianalisis kemudian dikonversi ke kode SAMPA menggunakan *python* yang akan

dijalankan pada *command prompt*. *Python* sendiri merupakan bahasa pemrograman yang berisi kode-kode instruksi yang pada penelitian ini untuk membantu melakukan konversi fonem ke kode SAMPA secara otomatis.

2.7 Model Durasi Klatt

Dennis Klatt mengusulkan sebuah model berbasis aturan, yang diimplementasikan dalam sistem MITalk (Rao, 2012; Kasparaitis, 2016). Model durasi Klatt telah dikembangkan untuk beberapa bahasa, termasuk Bahasa Inggris, Swedia, Jerman dan Prancis (Benesty, Sondhi, & Huang, 2008; Rao, 2012; Kasparaitis, 2016). Model Klatt menyatakan bahwa (Klatt D. H., 1987; Campbell, 2000; Benesty, Sondhi, & Huang, 2008; Öztürk, 2014).

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 (Klatt D. H., 1973).

Menurut Kasparaitis (2016), 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 (Klatt D. H., 1987; Benesty, Sondhi, & Huang, 2008; Öztürk, 2014).

$$D_0 = K \cdot (D_i - D_{min}) + D_{min} \quad (2.1)$$

D_0 adalah nilai prediksi durasi yang dihasilkan dari persamaan (1). 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 (Klatt D. H., 1987). Nilai konstanta dalam penelitian ini tidak menggunakan 11 aturan dari Klatt, karena 11 aturan ini hanya digunakan dalam prediksi durasi Bahasa Inggris. Jadi, penelitian hanya menggunakan dua aturan untuk mendapatkan nilai konstanta diantaranya berdasarkan struktur suku kata dan tipe suku kata.

$$K = \frac{D_0 - D_{min}}{D_i - D_{min}} \quad (2.2)$$

Sebelum melakukan prediksi dari persamaan 2.1, terlebih dahulu menentukan nilai konstanta setiap fonem sesuai dengan dua aturan yang digunakan. Konstanta fonem dapat diperoleh menggunakan persamaan (2) (Simões, 1990). Untuk D_0 (durasi awal atau durasi asli), durasi minimal fonem dan durasi rata-rata fonem diperoleh dari data korpus Bahasa Melayu Pontianak.

2.8 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 (Priyowidodo & Irwan, 2010). Menurut Anggara Narullita (2009), *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.

Jumlah daftar fonem dalam suatu database *diphone* tergantung dari hasil penelitian masing-masing pembuat *database*. 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 (Mentari Putri, 2016). *Diphone* yang digunakan dalam penelitian ini adalah *diphone database* Bahasa Indonesia “id1”. *Database*

diphone bahasa Indonesia (Id1) memiliki 30 fonem termasuk spasi (Arman, 2008). Penggunaan *diphone database* ini untuk membangkitkan sinyal ucapan dari fonem-fonem bahasa Melayu Pontianak yang telah dikonversi ke kode SAMPA. Fonem yang telah dikonversi, diikuti data berupa durasi tiap-tiap fonem sehingga dapat menghasilkan suara sintesa ucapan dengan nada yang datar. Karena penelitian ini tidak melakukan analisis terhadap pola prosodi untuk nilai *pitch* setiap fonem. Table fonem untuk database Id1 dapat dilihat pada tabel 2.5 (Priyowidodo & Irwan, 2010).

Tabel 2.5 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, landau
oi	OI	Amboi
au	aU	ampau, pulau
Spasi		

Pada *database* diphone “id1” Bahasa Indonesia pada tabel 2.5, ada beberapa huruf yang tidak terdapat dalam kode sampa diantaranya huruf /x/, /q/, dan /v/. Kemudian ada beberapa fonem Bahasa Melayu Pontianak yang tidak terdapat pada fonem bahasa Indonesia yaitu /f/, /v/, /x/, /q/, dan /z/.

2.9 Pengujian Perbandingan Nilai Durasi

Pengujian dilakukan membandingkan nilai durasi terhadap kalimat yang sama dengan dua *file* suara yang berbeda. *File* suara tersebut terdiri dari *file* suara hasil sintesa ucapan dan rekaman asli penutur Bahasa Melayu Pontianak dengan format *.wav. Pengujian ini melakukan perbandingan nilai durasi setiap kalimat yang sama pada sintesa ucapan dan rekaman suara seorang penutur Bahasa Melayu Pontianak.

Hasil pengujian berupa persentase perbandingan nilai durasi sintesa ucapan dan rekaman asli seorang penutur Bahasa Melayu Pontianak. Hasil pengujian perbandingan durasi dapat disimpulkan dengan dua indikator. Pertama, jika persentase perbandingan yang didapatkan lebih dari 0% maka sintesa ucapan yang dihasilkan memiliki durasi lebih lambat dari durasi rekaman asli seorang penutur Bahasa Melayu Pontianak. Kedua, jika hasil yang didapat kurang dari 0% maka sintesa ucapan yang dihasilkan memiliki durasi lebih cepat dari durasi rekaman suara seorang penutur Bahasa Melayu Pontianak. Sehingga untuk dapat menentukan kedua indikator tersebut, dapat dilakukan dengan perhitungan berikut.

$$\text{Persentase perbandingan} = \frac{\text{durasi sintesa} - \text{durasi asli}}{\text{durasi asli}} \times 100 \quad (2.3)$$

2.10 Degradation Mean Opinion Score (MOS)

ITU-T (ITU *Telecommunication Standardization Sector*) merupakan salah satu sektor dari ITU (*International Telecommunication Union*) yang bekerja sama untuk mendefinisikan standar untuk bidang telekomunikasi. Menurut ITU-T Rec. P.10 (2006), pengujian *Degradation Mean Opinion Score* (DMOS) adalah hasil dari pengujian degradasi penilaian kategori atau yang disebut *Degradation Category Rating* (DCR). Dalam pengujian DCR, responden diperdengarkan audio

yang tidak diproses sebagai referensi sebelum mereka mendengarkan audio yang telah diproses (Benesty dkk, 2008:85).

Teknik dalam melakukan pengujian *Degradation Mean Opinion Score* adalah dengan mendengarkan suatu rekaman suara dan kemudian memberikan penilaian. Penilaian yang diberikan berbeda-beda sesuai dengan keakuratan dan kualitas suara yang diterima oleh responden. Tabel skor untuk penilaian DMOS sesuai dengan standar dari ITU-T P.800 dapat dilihat pada tabel 2.6.

Tabel 2.6 Tabel Skala dan Kriteria Penilaian DMOS

Skala DMOS	Quality
5	Degradasi tidak terdengar (<i>Degradation is inaudible</i>)
4	Degradasi terdengar tapi tidak mengganggu (<i>Degradation is audible but not annoying</i>)
3	Degradasi cukup mengganggu (<i>Degradation is slightly annoying</i>)
2	Degradasi mengganggu (<i>Degradation is annoying</i>)
1	Degradasi sangat mengganggu (<i>Degradation is very annoying</i>)

Dalam pengujian DMOS, kualitas dari sinyal ucapan berdasarkan dari persepsi umum yang dimiliki oleh setiap individu pendengar dalam pengujian. (Benesty, 2008:87). Rumus perhitungan DMOS dapat dilihat pada rumus dibawah.

$$DMOS = \frac{\sum_{i=0}^n x(i) \cdot k}{N} \quad (2.4)$$

Keterangan:

$x(i)$ = Nilai sample ke i

k = Jumlah Bobot

n = Jumlah Pengamat

2.11 *Word Error Rate (WER)*

Word Error Rate diturunkan dari algoritma jarak Levenshtein. WER didasarkan pada jumlah minimum dari penambahan, penghapusan dan perubahan yang harus dilakukan untuk mengubah teks hasil pengujian (hipotesis) menjadi teks referensi (kalimat uji). Setiap kalimat dari hasil pengujian (hipotesis) dibandingkan dengan teks referensi (kalimat uji), dan setiap kata yang tidak cocok terhadap kalimat uji dihitung sebagai kesalahan dan dibagi dengan jumlah total kata dalam kalimat uji. Persamaan dari WER untuk menyimpulkan ketiga tipe kesalahan (penambahan, pengapusan, dan perubahan) dalam satu kalimat hipotesis adalah sebagai berikut.

$$WER = \frac{S+D+I}{Nr} \quad (2.5)$$

Keterangan:

S = jumlah kata yang terganti pada teks hasil sintesa,

D = jumlah kata yang terhapus pada teks hasil sintesa,

I = jumlah kata yang ditambah pada teks hasil sintesa,

Nr = jumlah kata pada teks hasil sintesa.

BAB III

METODOLOGI PENELITIAN

3.1 Bahan Penelitian

Bahan penelitian yang digunakan berupa buku Sepok Satu karangan Pay Jarot Sujarwo. Buku ini berisi teks cerita yang dikemas dalam menggunakan Bahasa Melayu Pontianak, Kalimantan Barat. 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.

3.2 Alat Penelitian

Berikut beberapa alat penelitian yang digunakan dalam penelitian ini adalah.

1. Model durasi Klatt untuk memprediksi nilai durasi fonem.
2. *Diphone database* id1, yaitu *diphone database* Bahasa Indonesia yang digunakan untuk menghasilkan sintesa ucapan yang jelas pada Bahasa Melayu Pontianak.

3.3 Spesifikasi Perangkat Keras dan Perangkat Lunak

3.3.1 Perangkat Keras

Perangkat keras atau *hardware* adalah aspek fisik dari komputer yang dapat dilihat dan disentuh. Perangkat keras yang digunakan dalam penelitian ini adalah earphone dan Laptop Asus X455LJ dengan spesifikasi sebagai berikut.

1. Processor Intel(R) Core (TM) i3-5005U 2.00 GHz.
2. Hard drive 368 GB.
3. Random Access Memory (RAM) 4GB.
4. LCD Monitor 14 inchi.
5. Speaker.

3.3.2 Perangkat Lunak

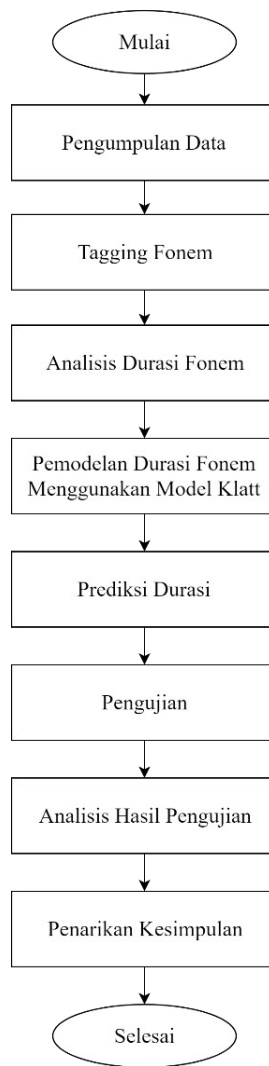
Perangkat lunak atau *software* adalah sekumpulan perintah atau program komputeryang digunakan untuk mengoperasikan perangkat keras agar dapat berinteraksi dengan pengguna. Beberapa perangkat lunak yang digunakan dalam

penelitian ini, antara lain sebagai berikut.

1. Sistem operasi *Windows 10*, merupakan sistem operasi yang digunakan dalam penelitian ini untuk menjalankan atau mengoperasikan perangkat keras komputer.
2. *WaveSurfer 1.8.8p4*, merupakan *software* yang digunakan untuk mengolah *file* suara rekaman serta melakukan pelabelan dan penandaan fonem dari *file* rekaman secara manual.
3. *Microsoft Excel 2016*, digunakan untuk melakukan analisis pengolahan data dan perhitungan terhadap nilai-nilai durasi fonem.
4. *Python 3.6.3.*, Bahasa pemrograman yang digunakan untuk mengimplementasikan perhitungan dari model durasi Klatt dalam melakukan prediksi nilai durasi tiap-tiap fonem.
5. *Sublime Text 3.0*, merupakan teks editor yang mana dalam penelitian ini digunakan untuk menulis kode-kode program dalam Bahasa pemrograman python.
6. *Command Prompt*, digunakan untuk mengeksekusi *file* python atau menampilkan hasil dari program python untuk melakukan proses serta menampilkan hasil prediksi durasi fonem dengan model Klatt.

3.4 Metode Penelitian

Metode penelitian merupakan tahapan-tahapan yang dirancang dan dijadikan sebagai panduan dalam melakukan sebuah penelitian. Metode penelitian yang dilakukan digambarkan pada diagram alir penelitian pada gambar 3.1.



Gambar 3.1 Diagram alir penelitian

Tahapan-tahapan penelitian seperti yang terlihat pada gambar 3.1 diantaranya pengumpulan data, tagging fonem, analisis durasi fonem, analisis durasi model Klatt, desain dan perancangan sistem, pengujian, analisis hasil pengujian dan penarikan kesimpulan.

3.4.1 Pengumpulan data

Penelitian ini memerlukan beberapa data untuk melakukan proses penelitian prediksi durasi menggunakan model durasi Klatt, diantaranya sebagai berikut.

1. Data Sekunder.

Data Sekunder adalah data yang diperoleh dalam bentuk yang sudah jadi, sudah dikumpulkan dan diolah oleh pihak lain, biasanya sudah dalam bentuk publikasi (Suryani dan Hendryadi, 2015). Pengumpulan data sekunder dapat menggunakan fasilitas internet, perpustakaan, publikasi lembaga-lembaga statistik, majalah dan sebagainya (Rangkuti, 2015). Data sekunder yang digunakan dalam penelitian ini adalah data teks cerita dalam Bahasa Melayu Pontianak yang diperoleh dari hasil observasi. Data teks berupa Buku yang berjudul *Sepok Satu* karangan Pay Jarot Sujarwo.

Data dari buku ini yang kemudian digunakan sebagai korpus dalam proses penelitian. Sebelum data buku *Sepok Satu* ini diolah, ada beberapa hal yang perlu disaring dalam data teks cerita tersebut. Diantaranya, tidak menyertakan kalimat yang terdapat kata atau istilah selain Bahasa Melayu Pontianak kecuali nama orang, nama tempat dan nama benda. Korpus yang digunakan terdapat nama orang yang didalamnya ada fonem /v/. Karena pada *diphone database* Bahasa Indonesia tidak terdapat kode sampa untuk fonem /v/, maka fonem /v/ dikonversikan menjadi fonem /f/. Korpus yang digunakan dari buku ini dan telah disaring untuk proses penelitian Model Durasi Klatt sebanyak 576 kalimat. Daftar kalimat yang digunakan dalam penelitian dapat dilihat pada lampiran A.

2. Data Primer.

Data primer adalah data yang dikumpulkan dan diolah sendiri oleh suatu organisasi atau perorangan langsung dari objeknya (Suryani dan Hendryadi, 2015). Pengumpulan data primer dapat dilakukan dengan cara personal *interviewing* atau wawancara pribadi (di rumah, di kantor) dengan menggunakan wawancara langsung, telepon, atau surat (Rangkuti, 2015). Data primer yang diperoleh dalam penelitian ini adalah berupa data-data rekaman suara dari buku *Sepok Satu* yang dituturkan oleh seorang penutur asli Bahasa Melayu Pontianak. Data rekaman Bahasa Melayu Pontianak disimpan dalam format audio *.wav dan bersifat *mono channel* dengan resolusi 16 bit. Durasi rekaman menyesuaikan kalimat yang diucapkan oleh seorang penutur. Total durasi rekaman yang dilakukan oleh penutur dalam merekam 576 kalimat adalah selama 32 menit 695

detik. Intonasi bicara dari seorang penutur tidak dibatasi sehingga rekaman yang dihasilkan tidak datar. Data rekaman kemudian digunakan untuk melakukan proses penandaan dan pelabelan fonem.

Pengumpulan data sekunder dan data primer yang dilakukan menggunakan metode-metode pengumpulan data berikut.

1. Observasi.

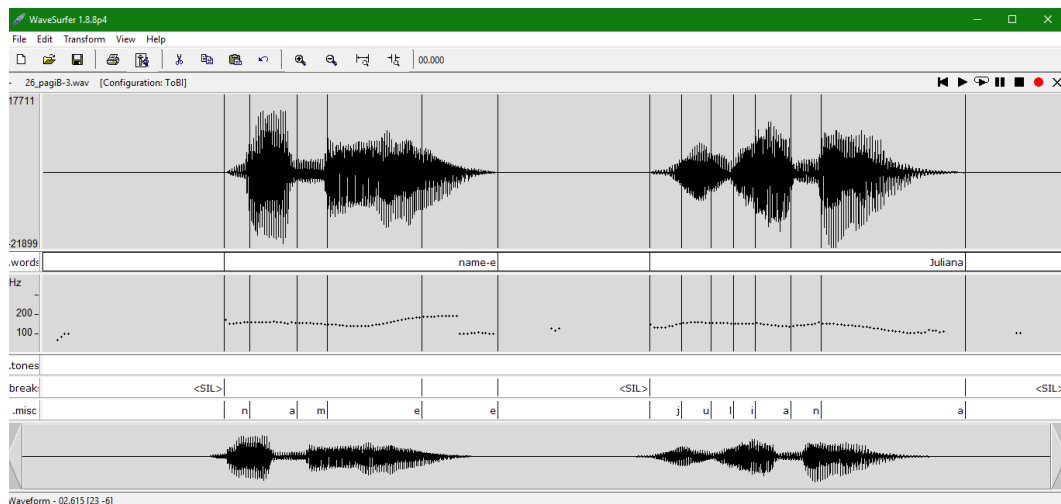
Penelitian menggunakan metode observasi terhadap penjual yang menjual buku Sepok Satu di Pontianak. Hal ini dilakukan dalam upaya pengumpulan teks cerita dalam Bahasa Melayu Pontianak.

2. Wawancara.

Setelah melakukan observasi terhadap buku Sepok Satu Bahasa Melayu Pontianak. Selanjutnya melakukan wawancara kepada narasumber. Narasumber disini adalah seorang penutur asli Bahasa Melayu Pontianak. Wawancara ini dilakukan untuk meminta ketersediaan penutur merekam data teks dari buku Sepok Satu Bahasa Melayu Pontianak guna kepentingan penelitian ini. Setelah penutur bersedia, dilakukanlah rekaman teks dari buku Sepok Satu.

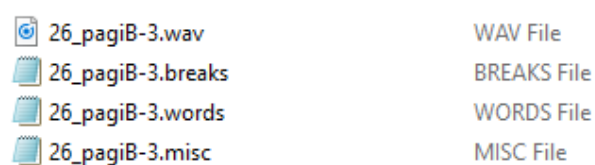
3.4.2 Tagging Fonem

Setiap *file* suara hasil rekaman dari seorang penutur asli Bahasa Melayu Pontianak yang tersimpan dalam format audio *.wav, selanjutnya dilakukan proses penandaan dan pelabelan untuk tiap fonem pada tiap-tiap kalimat. Proses ini dilakukan secara manual dengan cara mendengarkan rekaman suara pada aplikasi *Wavesurfer* 1.8.8p4. Kemudian diberi tanda diawal hingga diakhir pengucapan fonem seperti yang terlihat pada gambar 3.2. Hal ini dilakukan dengan tujuan untuk mendapatkan jumlah fonem serta nilai durasi tiap fonem yang terdapat pada 576 kalimat.



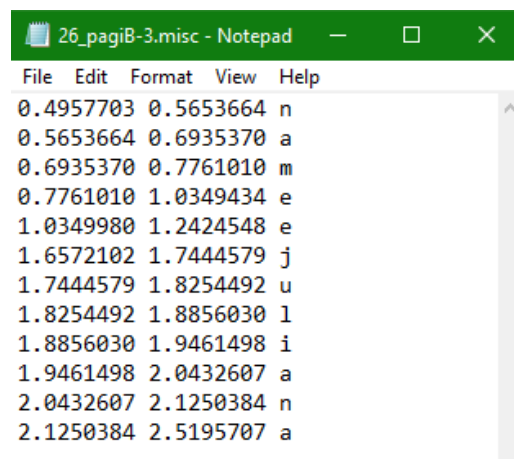
Gambar 3.2 Proses *tagging* fonem pada *wavesurfer* 1.8.8p4

Pada gambar 3.2 dilakukan proses penandaan dan pelabelan tiap fonem pada panel *.misc terhadap kalimat “name-e juliana”. Setelah melakukan penandaan pada *wavesurfer*, simpan hasil penandaan dari *file* *.wav. Setiap *file* *.wav yang diolah dalam *wavesurfer* menghasilkan empat format *file* yang berbeda tetapi dalam direktori *file* yang sama ketika disimpan hasil penandaan dan pelabelannya. Empat *file* diantaranya *.tones, *.breaks, *.misc, dan *.words. Keempatnya memiliki karakteristik isi *file* dan tujuan penggunaan yang berbeda. Pada gambar 3.2 tidak dilakukan penandaan terhadap bunyi pada format *.tones, sehingga pada gambar 3.3 tidak ada hasil dari penyimpanan *.wav dengan format *file* tersebut.



Gambar 3.3 Anotasi *file* hasil penandaan dan pelabelan

File *.breaks berisi durasi jeda, *file* *.words berisi panjangnya durasi tiap kata dan *file* *.misc berisi panjangnya durasi tiap fonem yang terjadi dalam satu kalimat atau satu *file* *.wav. *File* yang digunakan untuk proses penelitian selanjutnya yaitu *file* *.misc yang berisi fonem beserta durasi awal dan durasi akhir tiap-tiap fonem seperti pada gambar 3.4.



Gambar 3.4 *File misc*

Contoh isi dari *file *.misc* pada gambar 3.4 adalah “0.4957703 0.5653664 n”. Nilai “0.4957703” merupakan titik awal penandaan pengucapan durasi pada fonem n. Nilai “0.5653664” untuk titik akhir dari penandaan pengucapan durasi dan huruf “n” merupakan fonem yang diucapkan dari kedua nilai durasi. Satuan nilai durasi fonem yaitu milidetik (ms).

3.4.3 Analisis Durasi Fonem

Fonem yang dianalisis adalah fonem yang dapat dikonversi ke dalam kode-kode fonem dengan *database diphone* Bahasa Indonesia. Ada dua tipe fonem yang dianalisis yaitu fonem vokal dan fonem konsonan. Fonem vokal yang dianalisis sebanyak 9 fonem diantaranya, /a/, /e/, /i/, /o/, /u/, /ə/, /ai/, /au/, /oi/. Sedangkan untuk fonem konsonan yang dianalisis sebanyak 23 fonem konsonan. Fonem konsonan tersebut terdiri dari fonem, /b/, /c/, /d/, /f/, /g/, /h/, /j/, /k/, /l/, /m/, /n/, /p/, /q/, /r/, /s/, /t/, /v/, /w/, /x/, /y/, /z/, /ng/, dan /ny/.

Setelah semua kalimat dilakukan penandaan dan pelabelan, maka proses selanjutnya melakukan analisis nilai durasi tiap-tiap fonem yang terdapat pada 576 *file *.misc*. Proses analisis ini dilakukan menggunakan *Microsoft excel 2016*. Pada *file *.misc* terdapat nilai titik awal dan titik akhir pengucapan durasi suatu fonem seperti yang telah dijelaskan sebelumnya. Sehingga lamanya durasi suatu fonem, baik fonem vokal maupun konsonan dapat diperoleh dari hasil perhitungan yaitu durasi akhir dikurang durasi awal. Kemudian hasil perhitungan tersebut dibulatkan seperti yang terdapat pada gambar 3.5.

= (C4-B4)*1000			
B	C	D	E

Dur.Awal	Dur.Akhir	Fonem	Durasi Fonem (ms)
0,4957703	0,5653664	n	70
0,5653664	0,6935370	a	128
0,6935370	0,7761010	m	83
0,7761010	1,0349434	e	259
1,0349980	1,2424548	e	207
1,6572102	1,7444579	j	87
1,7444579	1,8254492	u	81
1,8254492	1,8856030	l	60
1,8856030	1,9461498	i	61
1,9461498	2,0432607	a	97
2,0432607	2,1250384	n	82
2,1250384	2,5195707	a	395

Gambar 3.5 Proses hitung durasi tiap fonem

Setelah mendapatkan hasil pembulatan durasi tiap fonem dari seluruh kalimat, maka setiap hasil tersebut disaring berdasarkan abjad kemudian dikelompokkan berdasarkan urutan fonem dari fonem a-z, ng, ny, ai, oi, dan au seperti pada gambar 3.6. Data lengkap pengelompokkan fonem beserta durasi penutur dapat dilihat pada lampiran B.

4	Fonem	DURASI	Fonem	DURASI	Fonem	DURASI	Fonem	DURASI
5	a	108	b	78	c	25	d	98
6	a	90	b	43	c	30	d	67
7	a	60	b	80	c	140	d	58
8	a	72	b	66	c	141	d	53
9	a	55	b	38	c	50	d	49
10	a	125	b	78	c	89	d	52
11	a	72	b	48	c	43	d	43
12	a	56	b	51	c	133	d	38
13	a	49	b	37	c	49	d	65
14	a	230	b	37	c	32	d	31
15	a	98	b	57	c	62	d	67
16	a	75	b	61	c	45	d	32
17	a	54	b	47	c	119	d	50
18	a	45	b	17	c	37	d	45
19	a	57	b	22	c	52	d	51
20	a	57	b	44	c	40	d	63
21	a	57	b	47	c	33	d	85
22	a	42	b	36	c	70	d	28
23	a	36	b	39	c	65	d	74
4317								
4318	Jlh.dur	343264	Jlh.dur	48522	Jlh.dur	10914	Jlh.dur	50105
4319	Rata2.dur	80	Rata2.dur	60	Rata2.dur	70	Rata2.dur	65
4320	MIN.dur	12	MIN.dur	4	MIN.dur	10	MIN.dur	7
4321	MAX.dur	848	MAX.dur	551	MAX.dur	822	MAX.dur	922
4322	Jlh.fonem	4312	Jlh.fonem	810	Jlh.fonem	157	Jlh.fonem	766

Gambar 3.6 Pengelompokan fonem dan durasi penutur

Tujuan dikelompokkannya fonem seperti pada gambar 3.6 yaitu untuk mendapatkan nilai durasi minimal, maksimal dan durasi rata-rata dari tiap-tiap fonem vokal dan konsonan. Tabel rencana hasil analisis dapat dilihat pada table 3.1.

Tabel 3.1 Tabel Rencana Hasil Analisis Nilai Durasi Penutur

Tipe Fonem	Fonem	Jumlah Fonem	Total Jumlah Durasi Fonem (ms)	Durasi Minimal/ Dmin (ms)	Durasi Maksimal (ms)	Durasi Rata-rata/ Di (ms)
Vokal	a					
	e					
	i					
	o					
	u					
	e'					
	ai					
	au					
	oi					
Konsonan	b					
	c					
	d					
	f					
	g					
	h					
	j					
	k					
	l					
	m					
	n					
	p					
	q					
	r					
	s					
	t					
	v					
	w					
	x					
	y					
	z					
	ng					
	ny					
Total Jumlah Fonem						

Pada tabel 3.1 ini terdiri dari 7 kolom diantaranya.

- a. Tipe fonem: kolom ini berisi tipe fonem yang terdiri dari dua tipe yaitu vokal dan konsonan.
- b. Fonem: berisi 9 fonem vokal (a, i, e, o, u, e', ai, au, oi) dan 23 fonem konsonan (b, c, d, f, g, h, j, k, l, m, n, p, q, r, s, t, v, w, x, y, z, ng, ny).
- c. Jumlah fonem: berisi jumlah tiap-tiap fonem yang terdapat pada 576 kalimat dan jumlah fonem-fonem ini diperoleh dari hasil analisis fonem seperti pada tabel 3.5.
- d. Total jumlah durasi fonem: berisi jumlah total durasi masing-masing fonem yang diperoleh dari analisis durasi fonem.
- e. Durasi minimal: berisi nilai durasi terkecil dari masing-masing fonem.
- f. Durasi maksimal: berisi nilai durasi terbesar dari masing-masing fonem.
- g. Durasi rata-rata: berisi nilai rata-rata durasi masing-masing fonem yang diperoleh dari jumlah total durasi fonem dibagi banyaknya jumlah fonem tersebut.

3.4.4 Pemodelan Durasi Fonem Menggunakan Model Klatt

Pemodelan durasi fonem dengan model Klatt adalah proses yang dilakukan untuk dapat memprediksi nilai durasi pada tiap-tiap fonem. Pada tahap analisis ini yang perlu diketahui sebelumnya adalah persamaan dari model durasi Klatt untuk memprediksi durasi fonem. Persamaan model durasi Klatt sebagai berikut.

$$D_0 = K \cdot (D_i - D_{\min}) + D_{\min} \quad (3.1)$$

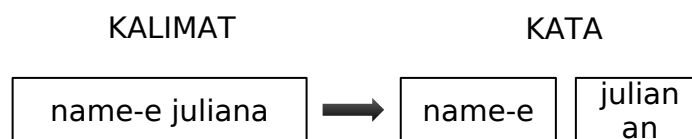
Persamaan model durasi Klatt terdiri dari nilai konstanta (K) fonem vokal dan konsonan, nilai durasi minimal (D_{\min}) fonem dari penutur, nilai durasi rata-rata (D_i) fonem dari penutur dan untuk D_0 adalah durasi Klatt. Hasil nilai D_{\min} dan D_i setiap fonem vokal dan konsonan diambil dari hasil analisis durasi fonem. Sedangkan untuk mendapatkan nilai konstanta fonem perlu dianalisis terlebih dahulu.

3.4.4.1 Analisis Nilai Konstanta Fonem

Konstanta fonem ditentukan menggunakan aturan suku kata dan tipe suku kata. Berikut tahapan dalam menganalisis nilai konstanta fonem.

1. Segmentasi kata

Seluruh korpus sebanyak 576 kalimat dilakukan segmentasi kata. Segmentasi kata yaitu mengubah teks kalimat menjadi blok kata per kata seperti contoh pada gambar 3.7.



Gambar 3.7 Segmentasi Kata

Pada gambar 3.7 contoh kalimat yang dipenggal menjadi blok kata yaitu "name-e juliana". Kalimat tersebut terdiri dari 2 kata yaitu kata "name-e", dan "juliana". Kedua kata inilah yang merupakan hasil dari segmentasi kata.

2. Segmentasi Suku Kata

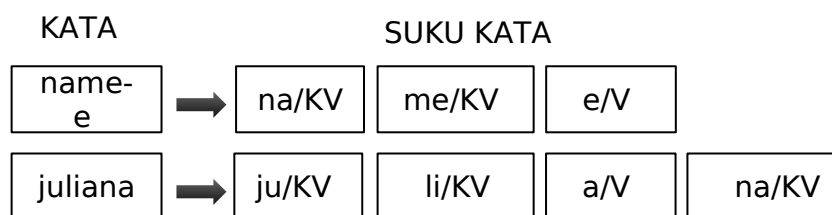
Langkah selanjutnya, melakukan pemenggalan kata menjadi suku kata berdasarkan struktur suku kata. Struktur suku kata yang digunakan merupakan struktur suku kata Bahasa Indonesia yang mengacu pada tabel 3.2 (Yuwono, 2007:164).

Tabel 3.2 Struktur Suku Kata Bahasa Indonesia

Struktur Suku Kata	Contoh Kata
V	<i>a</i> -ku, su- <i>a</i> -tu, go- <i>a</i>
VK	<i>an</i> -ti, ber- <i>in</i> -ves-ta-si, ma- <i>in</i>
KV	<i>be</i> -sar, bang- <i>sa</i>
KVK	<i>man</i> -di, be- <i>lum</i>
KKV	<i>slo</i> -gan, <i>pra</i> -sas-ti, <i>tra</i> -ge-di
KKVK	<i>prak</i> -tis, <i>prok</i> -la-ma-si
VKK	<i>eks</i> -tra, <i>ing</i> -kar
KVKK	<i>teks</i> -til, kon- <i>teks</i> -tu-al, mo- <i>dern</i>
KKVKK	<i>trans</i> -por-ta-si
KKKV	<i>stra</i> -te-gi
KKKVK	<i>struk</i> -tur

Pada tabel 3.2 Suatu kata dipenggal menjadi suku kata berdasarkan aturan pemenggalan kata Bahasa Indonesia. Berikut aturan pemenggalan kata menurut Rudiyant (2015:36-37).

- a. Jika ditengah kata ada huruf vokal yang berurutan, maka pemenggalannya dilakukan diantara kedua huruf vokal itu. Contohnya pada kata **bu-ah**, **ma-in**, **ni-at** dan **sa-at**.
- b. Huruf diftong ai, au, dan oi tidak boleh dipenggal. Bunyi diftong tersebut merupakan kaidah dari aturan Bahasa Indonesia. Artinya, jika pada korpus terdapat kata yang didalamnya diikuti huruf diftong maka huruf diftong tersebut tidak boleh dipenggal. Contohnya pada kata pan-**dai**, **au**-la, **sau**-da-ra dan am-**boi**.
- c. Jika ditengah kata dasar terdapat huruf konsonan (termasuk gabungan huruf konsonan) diantara dua buah huruf vokal, pemenggalannya dilakukan sebelum huruf konsonan tersebut. Contohnya pada kata ba-**pak**, la-**wan**, ke-**nyang**, pe-**nger**-ti-an.
- d. Jika di tengah kata dasar ada dua huruf konsonan yang berurutan, pemenggalannya dilakukan di antara kedua huruf konsonan tersebut. Contohnya **man-di**, **sang-gup**, **som-bong** dan swas-**ta**.
- e. Jika di tengah kata dasar ada tiga huruf konsonan atau lebih yang masing-masing melambangkan satu bunyi, pemenggalannya dilakukan di antara huruf konsonan yang pertama dan huruf konsonan yang kedua. Contohnya **ul-tra** dan **ben-trok**.
- f. Apabila telah melakukan pemenggalan kata, maka kata yang dipenggal tersebut langsung diberikan struktur suku katanya seperti contoh pada gambar 3.8. Pemberian struktur suku kata pada tiap kata yang telah dipenggal bertujuan untuk mempermudah proses analisis selanjutnya.



Gambar 3.8 Segmentasi Suku Kata

Pada gambar 3.8, kata yang dipenggal menjadi suku kata adalah kata "name-e" dan "juliana". Hasil pemenggalan kata tersebut terdiri dari dua suku kata yaitu "na", "me", "e", "ju", "li", "a" dan "na". Kemudian kedua suku kata ini diberikan struktur suku kata yang sesuai pada tabel 3.2. Suku kata "na", "me", "ju", dan "li" terdiri dari satu konsonan dan satu vokal. Sedangkan suku kata "e" dan "a" terdiri dari satu vokal. Hasil pemberian struktur suku kata untuk suku kata "na", "me" dan "e" adalah "na/KV", "me/KV", "e/V".

3. Segmentasi Fonem

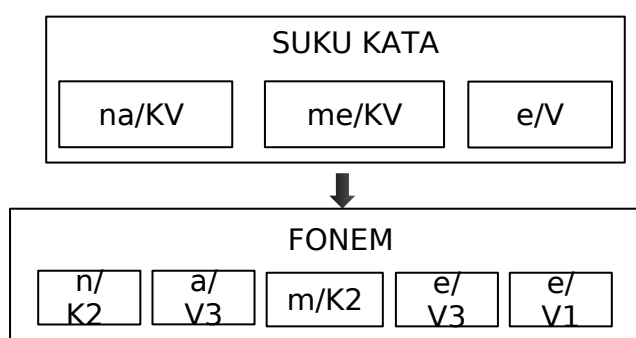
Setelah melakukan segmentasi suku kata, maka dipecah lagi menjadi blok-blok fonem berserta tipe suku kata baik konsonan maupun vokal. Tipe suku kata beserta contoh dapat dilihat pada tabel 3.3.

Tabel 3.3 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

Pada tabel 3.3 terdapat tipe suku kata dari 11 struktur suku kata. Tipe suku kata dibuat untuk mengelompokkan fonem-fonem vokal dan fonem-fonem konsonan berdasarkan posisi struktur suku kata fonem tersebut. Hal ini untuk memudahkan dalam melakukan analisis nilai konstanta tiap-tiap fonem berdasarkan tipe suku katanya. Fonem vokal terdiri dari 11 tipe suku kata, sedangkan untuk fonem konsonan terdiri dari 25 tipe suku kata sesuai posisi struktur suku kata Bahasa Indonesia. Segmentasi fonem dapat dilihat pada gambar 3.9.



Gambar 3.9 Segmentasi Fonem

Gambar 3.9 merupakan pemenggalan suku kata menjadi fonem-fonem serta pemberian tipe suku kata fonem sesuai dengan posisi struktur suku kata. Suku kata yang dipenggal yaitu "na", "me", dan "e" yang mana pada proses segmentasi suku kata telah diberikan struktur suku kata pada kata yang dipenggal menjadi suku kata. Hasil pemenggalan suku kata tersebut limat fonem yaitu "n", "a", "m", "e", dan "e". Kemudian fonem diberikan tipe suku kata seperti pada tabel 3.3 berdasarkan struktur suku kata yang telah diberikan sebelumnya. Hasil pemberian tipe suku kata dapat dilihat pada gambar 3.9 yang mana fonem konsonan "n" untuk K2 dilihat dari struktur suku kata KV yang diikuti oleh fonem

vokal "a" dengan tipe suku kata V3. Fonem yang sama memiliki tipe suku kata yang berbeda seperti pada gambar 3.9 pada fonem vokal "e", hal ini dikarenakan hasil pemenggalan suku kata. Dimana fonem vokal "e" sebelumnya terdapat fonem konsonan yaitu fonem "m" dan fonem "e" yang satu tidak diikuti konsonan sebelum atau setelahnya. Sehingga fonem "e" yang sebelumnya terdapat fonem konsonan memiliki tipe suku kata V3 dilihat dari suku kata KV. Sedangkan untuk fonem "e" dengan satu pola saja yaitu vokal memiliki tipe suku kata V1 dilihat dari suku kata V.

4. Mencari Nilai Konstanta Fonem

Setelah memberikan tipe suku kata pada tiap fonem sesuai dengan posisi struktur suku kata. Selanjutnya mencari nilai konstanta fonem vokal dan fonem konsonan dari 576 kalimat. Nilai konstanta diperoleh dari persamaan berikut, yaitu

$$K = \frac{D_0 - D_{\min}}{D_i - D_{\min}} \quad (3.2)$$

K adalah konstanta tiap fonem vokal dan konsonan yang ditentukan berdasarkan tipe suku kata. D_{\min} adalah nilai durasi minimal fonem dari penutur. D_i adalah nilai durasi rata-rata fonem dari penutur dan D_0 adalah nilai durasi fonem penutur. Contoh perhitungan konstanta tiap fonem pada kalimat ke 106 pada data dapat dilihat pada gambar 3.10. Data lengkap nilai konstanta fonem dari korpus sebelum dianalisis dapat dilihat pada lampiran C.

		=[@[D0-Dmin]]/[@[Di-Dmin]]					
fonem	Tipe SK	Dur.asli (D0)	Di	Dmin	D0-Dmin	Di-Dmin	Konstanta
n	K2	70	86	9	61	77	0,79
a	V3	128	80	12	116	68	1,71
m	K2	83	81	10	73	71	1,02
e'	V3	259	96	4	255	92	2,77
e'	K1	207	96	4	203	92	2,21
j	K2	87	63	13	74	50	1,48
u	V3	81	105	8	73	97	0,75
l	K2	60	60	8	52	52	1,00
i	V3	61	94	8	53	86	0,61
a	V1	97	80	12	85	68	1,25
n	K2	82	86	9	73	77	0,95
a	V3	395	80	12	383	68	5,63

Gambar 3.10 Proses hitung konstanta fonem

Gambar 3.10 terdiri dari kolom fonem, D_0 atau nilai durasi fonem penutur, tipe suku kata, D_i atau durasi rata-rata fonem, D_{min} atau durasi minimal fonem dan K atau nilai konstanta fonem yang diperoleh dari hasil perhitungan dari persamaan 3.2. Nilai konstanta tiap fonem seperti gambar 3.10 disaring berdasarkan fonem. Fonem vokal dari tipe V1 sampai V11 dan fonem konsonan dari tipe K1 sampai K25. Contohnya fonem "a" pada gambar 3.10 terdiri dari 3 fonem "a", yang mana masing-masing fonem "a" ada di tipe suku kata yang berbeda-beda yaitu satu fonem "a" ditipe V1 dan dua fonem "a" ditipe V3. Kemudian nilai konstanta tiap fonem dikelompokkan berdasarkan fonem dan tipe suku kata seperti yang terlihat pada gambar 3.11.

4	fonem	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
5	a	1,25	0,83	1,79	0,70	1,02	0,84	1,03	0,00	1,00	0,00	0,00
6	a	2,13	0,86	1,73	0,71	1,16	0,99			0,39		
7	a	0,93	2,05	4,28	0,99	1,02	1,81					
8	a	1,94	2,06	1,28	0,46	1,18	1,48					
9	a	3,81	0,66	1,61	0,50	0,49	1,14					
10	a	2,13	1,46	0,98	0,91	0,53	1,68					
11	a	1,86	2,89	2,83	1,74	1,65	1,01					
12	a	1,71	1,29	1,91	0,98	1,44	0,59					
13	a	1,76	1,60	1,75	0,55	0,96	0,70					
14	a	1,11	1,00	2,27	1,15	0,70	0,69					
15	a	0,87	1,29	2,16	1,27	0,74	0,87					
16	a	1,96	0,50	1,71	0,42	0,41	1,42					
17	a	0,78	0,84	5,63	0,73	1,77	1,49					
2151	RATA-RATA	1,43	1,05	1,12	0,88	0,95	1,02	1,03	0,00	0,70	0,00	0,00

Gambar 3.11 Pengelompokan nilai konstanta tiap fonem

Data pengelompokan konstanta fonem berdasarkan fonem dan tipenya dapat dilihat pada lampiran **D**. Nilai konstanta fonem dari korpus tersebut dirata-ratakan untuk mendapatkan nilai konstanta dari tiap-tiap fonem vokal maupun fonem konsonan sesuai dengan tipe suku kata. Tabel rencana hasil rata-rata nilai konstanta fonem vokal dan fonem konsonan dapat dilihat pada tabel 3.4 dan 3.5.

Tabel 3.4 Tabel Rencana Hasil Analisis Nilai Konstanta Fonem Vokal

[illegible]

Tabel 3.5 Tabel Rencana Hasil Analisis Nilai Konstanta Fonem Konsonan

[illegible]

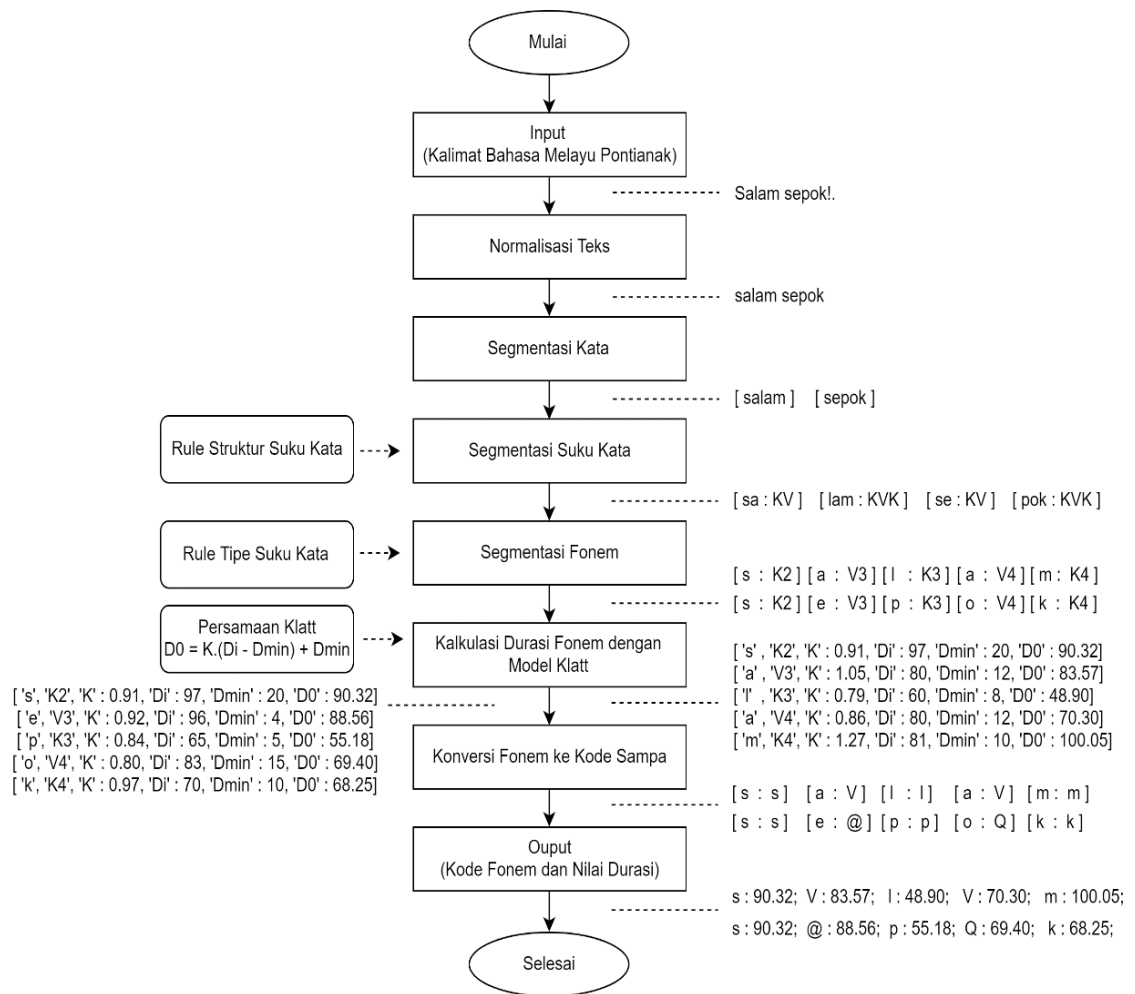
3.4.5 Prediksi Durasi

Setelah melakukan beberapa tahap dalam analisis durasi dari korpus Bahasa Melayu Pontianak Kalimantan Barat. Langkah selanjutnya yaitu melakukan prediksi durasi suatu fonem menggunakan model durasi Klatt. Model durasi Klatt memiliki persamaan untuk memprediksi fonem, persamaan tersebut dapat dilihat pada persamaan (3.1). Prediksi durasi dapat dilakukan dengan cara manual seperti contoh pada gambar 3.12.

Kalimat	Kata	Suku Kata	Struktur SK	Fonem	Tipe K/V	K	Di	Dmin	Di-Dmin	K.(Di - Dmin)	Dur.Klatt
salam sepok	salam	sa	KV	s	K2	0,9478478	97	20	77	72,98428333	93
				a	V3	1,6142181	80	12	68	109,7668304	122
	lam	KVK	KVK	l	K3	0,6853945	60	8	52	35,64051364	44
				a	V4	1,0458594	80	12	68	71,11844118	83
				m	K4	1,6841485	81	10	71	119,5745429	130
	se'pok	se'	KV	s	K2	0,9478478	97	20	77	72,98428333	93
				e'	V3	1,2092034	96	4	92	111,2467117	115
		pok	KVK	p	K3	0,4277504	65	5	60	25,665025	31
				o	V4	1,0610317	83	15	68	72,15015882	87
				k	K4	0,6238232	70	10	60	37,42938929	47

Gambar 3.12 Contoh Prediksi Durasi Model Klatt dengan Cara Manual

Hal ini membutuhkan lebih banyak waktu untuk memprediksi durasi fonem dengan model Klatt secara manual. Sehingga untuk memudahkan prediksi durasi dibuatlah algoritma kalkulasi model durasi Klatt dari persamaan (3.1). Pengkodean algoritma menggunakan Bahasa pemrograman *Python* dan dijalankan menggunakan *Command Prompt*. Diagram alir proses prediksi durasi fonem dengan model Klatt dapat dilihat pada gambar 3.10.



Gambar 3.13 Diagram Alir Proses Prediksi Model Klatt

Tahapan proses prediksi durasi menggunakan model Klatt pada gambar 3.13 terdiri dari input kalimat, normalisasi teks, segmentasi kata, segmentasi suku kata, segmentasi fonem, kalkulasi durasi fonem dengan model Klatt, konversi fonem ke kode SAMPA serta output yang berupa kode fonem dan nilai durasi.

3.4.5.1 Input

Proses yang pertama kali dilakukan dalam algortima adalah melakukan *input* teks berupa teks kalimat bahasa Melayu Pontianak.

3.4.5.2 Normalisasi Teks

Tahap normalisasi teks merupakan proses untuk mengubah teks masukan menjadi teks yang dapat diolah untuk proses selanjutnya. Normalisasi teks dilakukan terhadap huruf kapital dan tanda baca. Huruf kapital yang terdapat pada

teks masukan diubah menjadi huruf kecil atau *lower case*. Hal ini untuk mempermudah melakukan konversi teks ke kode fonem. Huruf yang dapat dikonversi ke kode fonem hanya huruf kecil saja. Kemudian melakukan penghapusan karakter angka dan tanda baca atau simbol lainnya karena tidak ditangani oleh sistem. Tanda baca serta simbol lainnya dalam proses ini dihilangkan karena tidak disertakan dalam analisis durasi fonem pada model durasi Klatt.

3.4.5.3 Segmentasi Kata

Selanjutnya melakukan segmentasi kata yaitu membagi teks kalimat menjadi blok kata per kata. Segmentasi dilakukan untuk mempermudah penerapan segmentasi kata ke dalam suku kata.

3.4.5.4 Segmentasi Suku Kata

Segmentasi suku kata merupakan proses untuk memenggal kata menjadi suku kata berdasarkan aturan Bahasa Indonesia. Struktur suku kata sendiri berupa susunan fonem-fonem baik fonem konsonan (K) maupun fonem vokal (V) yang menjadi bagian dari suatu kata. Terdapat sebelas struktur suku kata dalam Bahasa Indonesia yang digunakan dalam proses segmentasi suku kata, dapat dilihat pada tabel 3.1 yang telah dijelaskan pada subbab sebelumnya. Sebelas struktur ini digunakan karena tidak semua korpus berbahasa Melayu Pontianak. Seperti nama tempat, nama orang dan nama benda. Contohnya pada kata Amsterdam, William, pesawat, dan imigrasi.

3.4.5.5 Segmentasi Fonem

Suku kata dipenggal lagi menjadi satuan fonem-fonem. Tujuan dari proses ini, untuk mempermudah melakukan prediksi durasi fonem serta mempermudah melakukan pengkonversian fonem-fonem ke dalam kode-kode fonem. Proses ini tidak hanya melakukan segmentasi fonem, tetapi juga memberikan tipe suku kata terhadap setiap fonem. Contohnya kata “salam”, pada kata ini terdapat dua suku kata dengan polanya yaitu “sa/KV” dan “lam/KVK”.

Tipe suku kata diberikan berdasarkan tabel 3.3. Suku kata “sa” dan “lam” dipenggal menjadi satuan fonem yaitu “s”, “a”, “l”, “a”, “m” juga diikuti tipe suku

katanya. Pola “KV” untuk “sa” memiliki tipe suku kata “K2V3” dan “KVK” untuk “lam” memiliki tipe “K3V4K4”. Kata “salam” terdiri dari lima fonem yang diikuti tipe suku katanya yaitu “s/K2”, “a/V3”, “l/K3”, “a/V4”, “m/K4”. Tipe suku kata dan struktur suku kata dimasukkan ke dalam *array* sehingga mempermudah proses selanjutnya dalam pemberian nilai konstanta tiap fonem.

3.4.5.6 Kalkulasi Durasi Fonem dengan Model Klatt

Selanjutnya adalah proses perhitungan untuk melakukan prediksi durasi fonem dengan model Klatt menggunakan persamaan 3.2. Pada proses ini, setiap fonem diikuti tipe suku kata sesuai dengan pemenggalan suku kata berdasarkan struktur suku kata. Selain itu fonem juga diikuti komponen-komponen dari persamaan 3.2, diantaranya nilai konstanta (K), nilai durasi minimal (D_{\min}) dan nilai durasi rata-rata (D_i). ketiga komponen tersebut dilakukan analisis terlebih dahulu. Hasil dari analisis ketiga komponen itu dimasukkan kedalam *array* pada algoritma prediksi durasi Klatt. Contohnya kata “salam”, pada fonem “s” diproses ini menjadi [s , 'K2', 'K': 0.91, 'Di': 97, 'Dmin': 20, 'D0': 90.32]. “s” merupakan fonem konsonan dari kata “salam” dengan tipe suku kata “K2”. Nilai konstanta diambil dari tipe suku kata “K2” untuk fonem “s” dengan nilai konstanta sebesar 0.91ms. Fonem “s” memiliki nilai D_i sebesar 97ms dan D_{\min} sebesar 20ms. D_0 yang merupakan hasil perhitungan durasi Klatt untuk fonem “s” adalah sebesar 90.32ms.

3.4.5.7 Konversi Fonem ke Kode SAMPA

Fonem yang telah diprediksi nilai durasinya kemudian dikonversikan ke dalam kode fonem. Kode fonem berupa kode SAMPA yang terdapat dalam *Tools Mbrola Engine* dengan menggunakan *database diphone Bahasa Indonesia (id1)*. Kode SAMPA ini diperlukan untuk membangkitkan ucapan pada *tools Mbrola Engine* yang digunakan dalam tahap pengujian. Kode sampa juga dimasukkan kedalam *array* untuk mempermudah konversi fonem. Kode sampa dapat dilihat pada tabel 2.8. Contoh pada kata “salam” jika dikonversi kedalam kode SAMPA menjadi “sVIVm”.

3.4.5.8 Ouput

Output dari algoritma model durasi Klatt adalah berupa kode fonem hasil konversi dari kode SAMPA database *diphone* Bahasa Indonesia (id1) dan nilai durasi fonem hasil prediksi menggunakan model durasi Klatt.

3.4.6 Pengujian

Pengujian dilakukan untuk menguji kualitas hasil sintesa dari analisis durasi fonem dengan menggunakan model durasi Klatt. Pengujian ini menggunakan beberapa metode pengujian, diantaranya *Different Mean Opinion Score* (DMOS), *Word Error Rate* (WER), perbandingan durasi kalimat, perbandingan durasi fonem dan pengujian standar kecepatan durasi. Pengujian DMOS, WER dan perbandingan durasi fonem dilakukan secara terpisah dan responden yang berbeda dengan jumlah responden masing-masing pengujian 30 orang. Responden berasal dari kalangan akademisi yaitu mahasiswa program studi Informatika Fakultas Teknik Universitas Tanjungpura yang memahami dialek Bahasa Melayu Pontianak Kalimantan Barat. Pengujian DMOS dan WER masing-masing menggunakan sepuluh kalimat uji yang berbeda-beda yang diambil dari korpus. Kalimat uji DMOS dan perbandingan durasi fonem dapat dilihat pada tabel 3.6.

Tabel 3.6 Kalimat uji DMOS dan Perbandingan Durasi Fonem

No	No Kalimat Uji	Kalimat	Perbandingan Durasi Fonem
1	281	Aok ingat benar aku.	4%
2	268	Manelah negare kau tuh nak maju kalok tebiat pemerintahe tak tentu rudu macam itu.	-4%
3	317	Pade suatu hari kamek pegi ke kubu.	-5%
4	380	Bende ini ni ba tanggong jawab kitak semue.	-15%
5	534	Mang jago mang gugel ni.	1%
6	130	Aku sengaje turon lamak-lamak.	-2%
7	158	Lalu aku pon keluar dari pesawat.	26%
8	23	Jeket lelong kote baru ni kebetolan pulak warne-e jingge.	10%
9	307	Tiket pon udah dibeli.	5%
10	546	Sehe karang kau kenak pangkong pakek kayuk same orang belande.	22%

Pada saat melakukan pengujian DMOS dan WER, kalimat uji tidak disertakan dalam lembar pengujian dan tidak diperlihatkan atau diberitahukan kepada responden. Skenario audio pengujian DMOS dan WER dapat dilihat pada Tabel 3.7.

Tabel 3.7 Audio Pengujian DMOS dan WER

No	Audio DMOS	Audio WER
1	Durasi Klatt dan Durasi Penutur	Durasi Klatt
2	Durasi Klatt dan Durasi Penutur	Durasi Klatt
3	Durasi Klatt dan Durasi Penutur	Durasi Klatt
4	Durasi Klatt dan Durasi Penutur	Durasi Klatt
5	Durasi Klatt dan Durasi Penutur	Durasi Klatt
6	Durasi Klatt dan Durasi Penutur	Durasi Klatt
7	Durasi Klatt dan Durasi Penutur	Durasi Klatt
8	Durasi Klatt dan Durasi Penutur	Durasi Klatt
9	Durasi Klatt dan Durasi Penutur	Durasi Klatt
10	Durasi Klatt dan Durasi Penutur	Durasi Klatt

Lembar pengujian standar kecepatan durasi, DMOS dan WER dapat dilihat pada lampiran E. Pengujian MOS, DMOS dan WER memiliki kalimat uji yang sama sebanyak 10 kalimat namun pada saat pengujian, audio yang digunakan berbeda. Sebelum melakukan pengujian, sepuluh kalimat diprediksi menggunakan model Klatt untuk mendapatkan durasi dari masing-masing fonem agar dapat dibangkitkan pada *tools Mbrola Engine* menggunakan *diphone database* Bahasa Indonesia. Setelah dibangkitkan, *file* audio kalimat durasi Klatt disimpan dalam format *.wav . Kemudia disiapkan skenario untuk masing-masing pengujian seperti pada tabel 3.7. Audio yang diperdengarkan kepada responden berupa audio yang telah disimpan dengan format *.wav. Audio ini adalah hasil *converter* dari *file mbrola engine* (.pho). Sintesa ucapan dari hasil penelitian yang dibangkitkan di *mbrola engine* hanya berisi kode fonem dan nilai durasi hasil prediksi fonemnya saja, sehingga audio yang dihasilkan datar. Begitu pula dengan audio rekaman asli, hanya berisi kode fonem dan nilai durasi asli fonem.

Nilai *pitch* pada audio rekaman asli tidak disertakan untuk pengujian DMOS dan WER karena dalam penelitian ini hanya fokus pada analisis durasi fonem. *File* audio yang tadinya hanya berisi satu kali pengucapan, direkam ulang

sesuai banyaknya pengulangan pada masing-masing pengujian tersebut. Perekaman dilakukan menggunakan *mbrola engine*. Sehingga dalam satu *file *.wav* dapat berisi satu kalimat ucapan dengan banyaknya pengulangan. Skenario lama jeda pengulangan kalimat uji dapat dilihat pada tabel 3.8.

Tabel 3.8 Tabel Skenario Lama Jeda Pengulangan Kalimat Uji

Pengujian	Jenis Audio Uji	Pengulangan	Jeda Pengulangan
DMOS	Sintesa Klatt dan Sintesa penutur	3 Kali	3 detik
WER	Sintesa Klatt	4 Kali	2 Detik

Tabel 3.8 pengujian DMOS menggunakan audio sintesa klatt dan dilanjutkan dengan audio sintesa penutur. Pengulangan audio sintesa klatt dan penutur sebanyak tiga kali setiap kalimatnya. Masing-masing pengulangan diberi jeda tiga detik. Pengujian WER menggunakan audio sintesa klatt, audio diulang sebanyak empat kali untuk setiap kalimat. Masing-masing pengulangan diberi jeda dua detik.

3.4.6.1 Pengujian Standar kecepatan durasi

Pengujian standar kecepatan durasi dilakukan untuk mencari rentang nilai kecepatan durasi yang masih dapat diterima oleh pendengar yang dinyatakan dalam rentang persentase. Pengujian ini dilakukan terhadap satu kalimat yang digunakan sebagai standar untuk menentukan kualitas kecepatan ucapan. Kalimat yang digunakan sebagai standar berupa audio rekaman kalimat Bahasa Melayu Pontianak yang diambil dari korpus. Kalimat tersebut berbunyi "Lalulah die balek jadi pahlawan Belande" yang mana pengujian ini dilakukan terhadap 30 responden. Kecepatan audio rekaman suara kalimat ini dipercepat dan diperlambat sebesar 10%, 20%, 30%, 40%, dan 50%. Pengujian ini menentukan sampai batas berapa persen pendengar dapat menerima kecepatan durasi normal orang berbicara, apabila kalimat yang diucapkan diperlambat maupun dipercepat.

Responden diperdengarkan rekaman suara penutur dengan kecepatan yang berbeda-beda. Persentase kecepatan audio tidak diberitahukan kepada responden. Audio rekaman suara tiap kecepatan akan diputar sebanyak tiga kali. Setelah itu responden memberikan nilai sesuai dengan kriteria penilaian yang telah

ditentukan pada lembar pengujian. Urutan kecepatan audio rekaman untuk pengujian dapat dilihat pada tabel 3.9.

Tabel 3.9 Rencana Urutan Kecepatan Audio

Nomor	1	2	3	4	5	6	7	8	9	10
Kecepatan Ucapan	30%	-50%	10%	-20%	40%	-10%	50%	-30%	20%	-40%

Nomor satu, rekaman suara penutur diperlambat 30%, nomor dua dipercepat -50%, nomor tiga diperlambat 10%, nomor empat dipercepat -20%, nomor lima diperlambat 40%, nomor enam dipercepat -10%, nomor tujuh diperlambat 50%, nomor delapan dipercepat -30%, nomor sembilan diperlambat 20%, dan nomor sepuluh dipercepat -40%. Rencana pengujian standar kecepatan durasi dapat dilihat pada tabel 3.10.

Tabel 3.10 Rencana Pengujian Standar Kecepatan Durasi.

Nomor	Dapat Diterima	Tidak Dapat Diterima
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Tabel 3.10 terdiri dari tiga kolom yaitu kolom nomor dan dua kolom berikutnya merupakan kriteria penilaian diantaranya dapat diterima dan tidak dapat diterima. Kedua kolom tersebut yang diisi oleh responden dengan tanda centang (✓) pada tiap nomor sesuai jawaban yang responden berikan terhadap pengujian kecepatan ucapan normal. Hasil pengujian ini yakni menghitung rata-rata dari data jawaban responden. Mencari nilai rata-rata dilakukan dengan rumus sebagai berikut.

$$x^- = \frac{\sum_{i=0}^n x_i}{N} \times 100\% \quad (3.1)$$

[illegible]

Respon Den	Inisial	Nomor									
		1	2	3	4	5	6	7	8	9	10
Kecepatan Ucapan											
5											
6											
7											
8											
9											
10											
...											
30											
Rata-rata (%)											

Tabel 3.12 terdiri dari tiga kolom diantaranya.

- Responden: berisi nomor urut responden.
- Inisial: huruf pertama kata dari nama responden.
- Nomor: terdiri dari 10 kolom yang berisi penilaian dari opini responden.

Setelah mendapatkan hasil pengujian, selanjutnya hasil tersebut dianalisis guna menginterpretasikan persentase hasil rata-rata pengujian berdasarkan tabel 3.11 yaitu tabel kriteria penilaian. Rencana hasil analisis dapat dilihat pada tabel 3.13.

Tabel 3.13 Rencana Analisis Hasil Uji Standar Kecepatan Durasi

Range Hasil Uji Responden	Kecepatan Durasi Ucapan (%)	Rata-rata Hasil Uji (%)
80% - 100% [sangat baik]		... Rata-rata hasil uji per kecepatan ucapan...
70% - 79% [baik]		
60% - 69% [cukup baik]		
50% - 59% [tidak baik]		
0% - 49% [sangat tidak baik]		

Analisis hasil uji diurutkan berdasarkan persentase rata-rata hasil uji tertinggi hingga terendah. Dari hasil analisis inilah diketahui persentase kecepatan durasi ucapan yang dapat diterima maupun tidak dapat diterima oleh pendengar.

3.4.6.2 Pengujian Perbandingan Nilai Durasi

Pengujian perbandingan durasi terdiri dari dua bagian yaitu perbandingan durasi berdasarkan durasi fonem terhadap 10 kalimat uji dan perbandingan durasi berdasarkan durasi kalimat terhadap 100 kalimat uji.

3.4.6.2.1 Pengujian Perbandingan Nilai Durasi Berdasarkan Durasi Fonem

Pengujian perbandingan durasi fonem bertujuan untuk mencari nilai akurasi prediksi durasi fonem sintesa. Pengujian dilakukan terhadap 10 kalimat uji pada tabel 3.6. Durasi fonem setiap kalimat uji diprediksi terlebih dahulu baik fonem vokal maupun fonem konsonan untuk model durasi Klatt. Setelah mendapatkan durasi klatt, kemudian memberikan durasi penutur seperti contoh pada gambar 3.14. Pengujian dilakukan menggunakan tabel pada *Microsoft Excel* 2016.

fonem	tipe sk	klatt (ms)	penutur (ms)
a	V1	109	96
o	V2	109	128
k	K1	63	92
i	V1	91	49
ng	K3	75	65
a	V4	73	45
t	K4	81	64
b	K2	58	57
e	V3	89	35
n	K3	63	78
a	V4	73	99
r	K4	87	60
a	V1	109	91
k	K2	74	125
u	V3	106	127

Gambar 3.14 Contoh pemberian durasi fonem klatt dan penutur

Gambar 3.14 merupakan contoh satu kalimat pemberian durasi klatt dan penutur, yang mana durasi klatt diperoleh dari hasil prediksi menggunakan *python* klatt. Kalimat tersebut berbunyi "aok ingat benar aku" dan hasil pemenggalan suku kata pada kalimat tersebut yaitu "a/V - ok/VK", "i/V - ngat/KVK", "be/KV - nar/KVK", dan "a/V - ku/KV". Tipe fonem pada kalimat tersebut diperoleh dari

hasil konversi pemenggalan suku kata seperti yang terdapat pada tabel 3.3 yaitu tabel tipe suku kata. Pada kalimat ini terdiri dari empat kata, delapan suku kata dan 15 fonem. Pada gambar tersebut juga terdapat tipe fonem berdasarkan hasil pemenggalan suku kata guna untuk mengetahui perbedaan durasi pada fonem yang sama terhadap tipe suku kata yang berbeda.

Selanjutnya setiap fonem pada kalimat uji dilakukan perbandingan nilai durasi Klatt dan durasi penutur. Untuk menghitung persentase perbandingan durasi klatt dan durasi penutur dapat menggunakan rumus berikut.

$$\text{Persentase perbandingan} = \frac{\text{durasi Klatt} - \text{durasi Penutur}}{\text{durasi penutur}} \times 100 \quad (3.2)$$

Hasil perbandingan nilai durasi berupa persentase. Selanjutnya diurutkan berdasarkan nilai perbandingan dari yang terkecil hingga terbesar guna untuk mengetahui fonem yang masuk dalam rentang kecepatan durasi yang dapat diterima oleh pendengar. Kemudian 10 kalimat uji yang digunakan untuk membandingkan durasi fonem, dibandingkan juga berdasarkan total nilai durasi pada kalimat untuk mengetahui apakah hasil kedua perbandingan ini simetris. Tahap pengujian ini yaitu menjumlahkan durasi fonem masing-masing kalimat untuk durasi Klatt dan durasi penutur. Untuk menghitung perbandingan dapat menggunakan rumus pada persamaan 3.2. Rencana hasil perbandingan durasi klatt dan durasi penutur berdasarkan kalimat dapat dilihat pada tabel 3.14.

Tabel 3.14 Rencana Hasil Perbandingan Nilai Durasi Kalimat

Nomor	Kalimat Uji	Jumlah Durasi Klatt (ms)	Jumlah Durasi Penutur (ms)	Perbedaan (%)
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Setelah melakukan perbandingan durasi berdasarkan fonem dan kalimat terhadap 10 kalimat uji. Hasil perbandingan nilai durasi, dianalisis berdasarkan perbandingan nilai durasi untuk durasi fonem dan durasi kalimat. Analisis hasil uji digunakan untuk mengetahui total kalimat dan total fonem terhadap 10 kalimat uji yang masuk dalam rentang kecepatan durasi yang dapat diterima oleh pendengar berdasarkan pengujian standar kecepatan durasi.

1. Rencana hasil analisis perbandingan nilai durasi berdasarkan durasi kalimat dapat dilihat pada tabel 3.15.

Tabel 3.15 Rencana Hasil Analisis Perbandingan Nilai Durasi Kalimat

<i>Range Kecepatan Durasi Ucapan</i>	Jumlah Kalimat	Persentase Kalimat
41% - 50%		
31% - 40%		
21% - 30%		
11% - 20%		
1% - 10%		
0%		
(-1%) - (-10%)		
(-11%) - (-20%)		
(-21%) - (-30%)		
(-31%) - (-40%)		
(-41%) - (-50%)		
Total Jumlah		

2. Rencana hasil analisis perbandingan nilai durasi berdasarkan durasi fonem dapat dilihat pada tabel 3.16.

Tabel 3.16 Rencana Hasil Analisis Perbandingan Nilai Durasi Fonem

<i>Range Kecepatan Durasi Ucapan</i>	Jumlah Fonem	Total Fonem	Persentase Total Fonem
(-10%) – (10%)			
(-20%) – (10%)			
(-20%) – (20%)			
(-30%) – (20%)			
(-30%) – (30%)			
(-40%) – (30%)			
(-40%) – (40%)			
(-50%) – (40%)			
(-50%) – (50%)			

Setelah menganalisis perbandingan nilai durasi berdasarkan durasi fonem, kemudian menganalisis jumlah fonem dari hasil perbandingan nilai durasi fonem berdasarkan kecepatan durasi fonem yang dapat diterima dan tidak dapat diterima oleh pendengar. rencana tabel hasil analisis dapat dilihat pada tabel 3.17.

Tabel 3.17 Rencana Hasil Analisis Jumlah Fonem

Fonem (diterima)		Fonem (tidak dapat diterima)	
Vokal (%)	Konsonan (%)	Vokal (%)	Konsonan (%)

3.4.6.2.2 Pengujian Perbandingan Durasi Berdasarkan Kalimat

Pengujian perbandingan durasi berdasarkan kalimat antara durasi penutur dan durasi klatt. Pengujian dilakukan untuk melihat seberapa besar tingkat persentase perbedaan yang dihasilkan oleh durasi penutur dan durasi Klatt dalam kalimat yang sama. Tahap-tahap pengujian ini sama dengan tahap pengujian perbandingan durasi berdasarkan fonem. Hanya saja jumlah kalimat yang digunakan pada pengujian ini sebanyak 100 kalimat uji yang diambil dari 576 kalimat Bahasa Melayu Pontianak. Kalimat uji untuk perbandingan nilai durasi dapat dilihat pada lampiran F. Rencana tabel pengujian perbandingan ini dapat dilihat pada tabel 3.18.

Tabel 3.18 Rencana Pengujian Perbandingan Durasi Berdasarkan Kalimat

No	Nomor korpus	Kalimat Uji	Jumlah Durasi Penutur (ms)	Jumlah Durasi Klatt (ms)	Perbedaan (%)	Keterangan
1						> 0%
2						> 0%
3						> 0%
4						> 0%
5						> 0%
...						0%
100						< 0%

Pada tabel 3.18 terdiri dari tujuh kolom yaitu nomor, nomor korpus yang dipilih secara acak dari 576 kalimat, kalimat uji sesuai dengan nomor korpus,

durasi penutur merupakan jumlah durasi fonem dari penutur tiap kalimat uji, dan kolom durasi model klatt merupakan jumlah durasi Klatt tiap kalimat uji yang diperoleh dari hasil diprediksi. Kolom persentase perbandingan diperoleh dari perhitungan persamaan 3.3. Setelah melakukan pengujian perbandingan durasi klatt dan durasi penutur. Selanjutnya hasil pengujian 100 kalimat uji dirangkum seperti pada tabel 3.19.

Tabel 3.19 Rencana Hasil Pengujian Perbandingan Nilai Durasi

Keterangan	Persentase Tertinggi	Persentase Terendah	Jumlah Kalimat
< 0%			
0%			
> 0%			
Total Jumlah Kalimat			

Tabel 3.19 merupakan tabel rencana hasil pengujian perbandingan yang bertujuan untuk mengetahui kualitas kecepatan durasi Klatt terhadap durasi penutur dari 100 kalimat uji. Analisis hasil uji sama dengan pengujian perbandingan nilai durasi fonem. Hasil analisis digunakan untuk memilah kalimat uji WER berdasarkan persentase standar kecepatan ucapan yang dapat diterima oleh pendengar.

Tabel 3.20 Rencana Hasil Analisa Pengujian Perbandingan Nilai Durasi

Range Kecepatan Durasi Ucapan	Jumlah Kalimat	Persentase Kalimat
41% - 50%		
31% - 40%		
21% - 30%		
11% - 20%		
1% - 10%		
0%		
(-1%) - (-10%)		
(-11%) - (-20%)		
(-21%) - (-30%)		
(-31%) - (-40%)		
(-41%) - (-50%)		
Rata-rata		

Tabel 3.20 merupakan tabel rencana hasil analisa pengujian

perbandingan nilai durasi berdasarkan 100 kalimat uji. Tabel terdiri dari tiga kolom diantaranya kolom *range* kecepatan durasi ucapan, jumlah kalimat dan persentase kalimat. Range kecepatan durasi ucapan berisi rentang yang digunakan untuk menentukan jumlah kalimat yang berada pada rentang tersebut berdasarkan kecepatan durasi yang dihasilkan dari tabel 3.17. Tabel 3.19 juga digunakan untuk menentukan kalimat uji pengujian *Word Error Rate* (WER). *Range* kecepatan durasi ucapan untuk kalimat uji WER ditentukan berdasarkan hasil pengujian standar kecepatan durasi. Dari hasil tersebut diambil lima kalimat tercepat dan lima kalimat terlambat untuk pengujian WER.

3.4.6.3 *Degradation Mean Opinion Score (DMOS)*

Pengujian DMOS bertujuan untuk membandingkan kecepatan durasi sintesa klatt dan durasi sintesa penutur berdasarkan pendapat dari responden. Tahap pengujian perbandingan kecepatan durasi hanya dilakukan terhadap 10 kalimat uji. Pengujian DMOS berupa pengujian dengan mendengarkan audio hasil sintesa klatt dan sintesa penutur. Responden diperdengarkan audio hasil sintesa klatt dan sintesa penutur tanpa memberikan daftar kalimat yang diujikan. Audio sintesa klatt dan sintesa penutur diperdengarkan sebanyak tiga kali seperti pada tabel 3.6. Tujuan ini dilakukan karena tiap responden memiliki tingkat daya tanggap yang berbeda-beda sehingga diberikan kesempatan untuk mendengarkan sebanyak jumlah yang telah ditetapkan.

Setelah mendengarkan 10 kalimat uji dengan tiga kali pengulangan audio tiap kalimatnya, responden memberikan penilaian sesuai dengan skala yang telah ditentukan pada lembar kuesioner. Skala ini ditentukan berdasarkan teknik *Degradation Mean Opinion Score (DMOS)*, yang mana teknik ini digunakan untuk pengujian *intelligibility* atau kejelasan suara yang dihasilkan. Penilaian pada setiap kategori menggunakan skala yang dapat dilihat pada tabel 3.13. Skala dalam penilaian DMOS dapat dilihat pada tabel 3.21.

Tabel 3.21 Penilaian DMOS

Skala DMOS	Kualitas
4	Durasi Klatt sangat mendekati durasi penutur

Skala DMOS	Kualitas
3	Durasi Klatt mendekati durasi penutur
2	Durasi Klatt tidak mendekati durasi penutur
1	Durasi Klatt sangat tidak mendekati durasi penutur

Skala penilaian DMOS pada tabel 3.20 dikonversi ke skala 1-5 sesuai dengan standar ITU-P.800.1. Skala penilaian DMOS pada saat pengujian menggunakan skala 1-4 dikarenakan skala 3 dengan kualitas penilaian sedang atau cukup pada skala penilaian DMOS dengan standar ITU-P.800.1 dapat menyebabkan kebingungan terhadap responden. Kebingungan yang terjadi disebabkan responden cenderung memberi nilai sedang atau cukup terhadap objek pengujian yang belum pasti baik ataupun buruk dimana penilaian sedang atau cukup tidak memiliki batas yang jelas. Konversi dilakukan dengan rumus:

$$DMOS_{ITU-P.800.1} = \frac{5}{4} DMOS \quad (3.3)$$

Pengujian DMOS dilakukan dengan rumus:

$$DMOS = \frac{\sum_{i=0}^n x(i) \cdot k}{N} \quad (3.4)$$

Dengan $x_{(i)}$ merupakan nilai sampel ke- i , k merupakan jumlah bobot yang diberikan oleh responden dan telah dikonversi pada rumus 3.3 dan N merupakan jumlah responden. Nilai DMOS dihitung dengan merata-ratakan nilai DMOS dengan jumlah responden N . Hasil rata-rata nilai DMOS inilah yang menjadi hasil akhir dari pengujian perbandingan kecepatan durasi ini. Untuk menginterpretasikan nilai hasil rata-rata DMOS yang telah dikonversi dapat dilihat pada tabel 3.22.

Tabel 3.22 Kriteria Penilaian Hasil Rata-rata DMOS

Skala	Range Nilai Rata-Rata DMOS	Kualitas
5	4.01 – 5.00	Durasi Klatt sangat mendekati durasi penutur
4	3.01 – 4.00	Durasi Klatt mendekati durasi penutur
3	2.01 – 3.00	Durasi Klatt cukup mendekati durasi penutur

2	1.01 – 2.00	Durasi Klatt tidak mendekati durasi penutur
1	0.00 – 1.00	Durasi Klatt sangat tidak mendekati durasi penutur

Kalimat uji yang digunakan pada pengujian DMOS dapat dilihat pada tabel 3.6. Responden mengisi tabel pengujian dengan memberikan tanda centang (✓) pada tiap nomor kalimat sesuai tanggapan responden dan skala yang telah diberikan. Rencana tabel pengujian DMOS dapat dilihat pada tabel 3.23.

Tabel 3.23 Rencana Pengujian DMOS

Bobot	Keterangan	Nomor Kalimat									
		1	2	3	4	5	6	7	8	9	10
4	Durasi Klatt sangat mendekati durasi penutur										
3	Durasi Klatt mendekati durasi penutur										
2	Durasi Klatt tidak mendekati durasi penutur										
1	Durasi Klatt sangat tidak mendekati durasi penutur										

Tabel 3.23 yang merupakan tabel rencana pengujian DMOS terdiri dari 3 kolom, yaitu.

- Bobot: berupa nilai dari kualitas kalimat yang diuji.
- Keterangan: merupakan penjelasan dari setiap bobot sebagai tolak ukur penilaian oleh responden.
- Nomor Kalimat: terdiri dari 10 kolom yang berisi penilaian yang diberikan oleh setiap responden berdasarkan opini responden serta bobot dan keterangan yang telah diberikan pada table pengujian.

Nilai DMOS direkap dan disimpan pada tabel hasil pengujian DMOS yang dapat dilihat pada tabel 3.24.

Tabel 3.24 Rencana Hasil Pengujian DMOS

Nomor Responden	Nomor Kalimat									
	1	2	3	4	5	6	7	8	9	10

Nomor	Nomor Kalimat									
Responden										
2										
3										
4										
5										
6										
7										
8										
...										
30										
Rata-rata										
Total rata-rata										

Tabel 3.24 yang merupakan tabel rencana hasil pengujian DMOS terdiri dari lima kolom, yaitu.

- Nomor responden: berupa nomor urut responden.
- Inisial: huruf pertama kata dari nama responden.
- Jenis kelamin: menunjukkan responden laki-laki atau perempuan.
- Nomor kalimat: terdiri dari 10 kolom yang berisi bobot-bobot yang diberikan responden.
- DMOS: hasil DMOS dari setiap responden berdasarkan persamaan 3.4.

3.4.6.4 Word Error Rate (WER)

Pengujian WER digunakan menghitung akurasi kalimat yang diujikan dari model durasi Klatt. Tujuan dilakukannya pengujian WER ini untuk mengetahui tingkat kejelasan (*intelligibility*) kalimat yang dihasilkan oleh sintesa ucapan dari model durasi Klatt. Tahap pengujian ini memperdengarkan 10 kalimat uji, dimana masing-masing kalimat diulang sebanyak empat kali. Kalimat uji yang digunakan diambil dari 100 kalimat uji sebelumnya berdasarkan hasil pengujian standar kecepatan durasi. Kalimat uji terdiri dari lima kalimat dengan durasi kalimat tercepat dan lima kalimat dengan durasi kalimat terlambat yang masuk dalam standar kejelasan ucapan. Rencana tabel kalimat uji WER dapat dilihat pada tabel 3.25.

Tabel 3.25 Rencana Tabel Kalimat Uji WER

No	Kalimat Uji	Kecepatan Durasi Kalimat (%)
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Setelah didengarkan, responden diminta untuk menuliskan kembali kalimat yang telah didengarkan. Kalimat yang telah ditulis kembali oleh responden kemudian dibandingkan dengan transkrip kalimat uji yang ada pada tabel 3.7. Tabel rencana pengujian WER dapat dilihat pada tabel 3.26.

Tabel 3.26 Rencana Pengujian *Word Error Rate*

No	Kalimat
1kalimat hasil dengar.....
2	
3	
4	
5	
6	
7	
8	
9	
10	

Hasil pebandingan tersebut digunakan untuk menghitung tingkat kejelasan hasil sintesa. Perhitungan WER menggunakan rumus 3.5 dan tingkat kejelasan hasil sintesa direpresentasikan dalam bentuk akurasi WER dengan mennggunakan rumus 3.6.

$$\text{WER} = \frac{S+D+I}{N_r} \quad (3.5)$$

$$\text{Akurasi WER} = 1 - \text{WER} \quad (3.6)$$

S atau singkatan dari *substitution* merupakan jumlah kata yang terganti pada kalimat yang telah didengarkan dan ditulis kembali, D atau *deletion*

merupakan jumlah kata yang terhapus, I atau *insertion* merupakan jumlah kata yang bertambah dan N_r merupakan jumlah kata pada kalimat yang didengarkan. Nilai WER didapatkan dari hasil penjumlahan jumlah kata terganti, terhapus dan bertambah kemudian dibagi dengan jumlah kata dari kalimat. Jumlah nilai WER 3.27.

Tabel 3.27 Tabel Rencana Hasil Pengujian WER

No. Res ponden	No. Kal i mat	Hasil dikte	Jumlah			Total Kata	WER
			I	U	D		
1	1						
	2						
	3						
	...						
	10						
...	...						
30	1						
	2						
	3						
	..						
	10						
Rata-rata WER							
Akurasi WER							

3.4.7 Analisis Hasil Pengujian

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

3.4.8 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.

BAB IV

HASIL DAN ANALISA

4.1 Hasil Penelitian

Hasil dari penelitian ini diantaranya terdiri dari analisis durasi, analisis durasi dengan model Klatt, implementasi prediksi durasi fonem dengan model Klatt dan pengujian. Output dari hasil penelitian adalah berupa nilai durasi tiap-tiap fonem yang diprediksi menggunakan model durasi Klatt serta fonem yang telah dikonversi ke dalam kode-kode fonem.

4.1.1 Hasil Analisis Durasi Fonem

Hasil analisis fonem vokal dan konsonan terdapat empat fonem konsonan yang tidak ada dalam data Bahasa Melayu Pontianak. Empat fonem itu diantaranya fonem /q/, /v/, /x/, dan /z/, sehingga hasil analisis durasi dari fonem-fonem ini bernilai 0. Namun Bahasa Melayu Pontianak juga tidak memiliki fonem /f/, tetapi terdapat fonem tersebut pada data Bahasa Melayu Pontianak. Fonem tersebut berupa nama orang. Kata untuk fonem /f/ yang terdapat pada korpus adalah "fan". Total jumlah fonem vokal dan fonem konsonan dari 576 korpus kalimat adalah sebanyak 23351 fonem. Jumlah fonem yang diperoleh merupakan hasil dari jumlah 9 fonem vokal dan 23 fonem konsonan. Perbandingan jumlah data fonem vokal dan konsonan yang digunakan untuk penelitian terhadap 23351 fonem pada tabel 4.1 yaitu fonem konsonan 56% dan fonem vokal 44%.

Setiap fonem yang dianalisis memiliki total jumlah durasi fonem, durasi minimal fonem, durasi maksimal dan durasi rata-rata tiap fonem. Durasi tersebut didapatkan dari hasil analisis masing-masing durasi. Jumlah fonem yang paling sedikit dari korpus kalimat yaitu fonem /v/ dengan jumlah fonem tiga buah. Hasil analisis durasi fonem yang diperoleh dari 576 korpus kalimat Bahasa Melayu Pontianak dapat dilihat pada tabel 4.1.

Tabel 4.1 Hasil Analisis Durasi Fonem

Tipe Fonem	Fonem	Jumlah Fonem	Total Jumlah Durasi Fonem (ms)	Durasi Minimal/ Dmin(ms)	Durasi Maksimal (ms)	Durasi Rata-rata/ Di(ms)
Vokal	a	4312	343264	12	848	80
	e	602	47730	8	452	79
	i	1234	116061	8	580	94

Tipe Fonem	Fonem	Jumlah Fonem	Total Jumlah Durasi Fonem (ms)	Durasi Minimal/ Dmin(ms)	Durasi Maksimal (ms)	Durasi Rata-rata/ Di(ms)
	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
Konsonan	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 Fonem Vokal		10372 (44%)				
Total Fonem Konsonan		12979 (56%)				
Total Jumlah Fonem		23351				

4.1.2 Hasil Analisis Konstanta Fonem

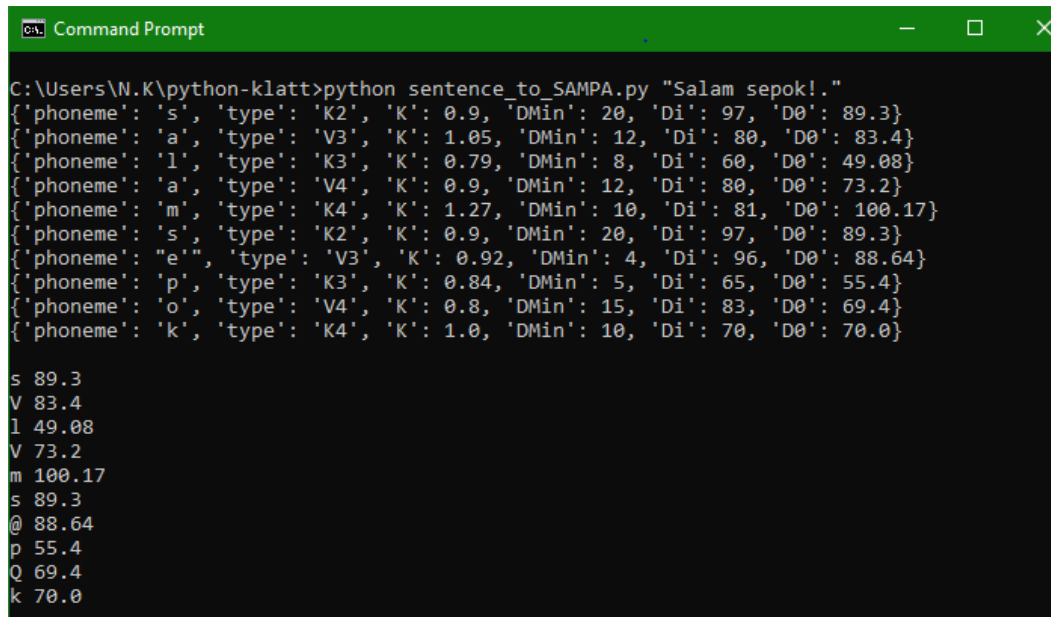
Hasil nilai konstanta fonem vokal dan konsonan dapat dilihat pada lampiran G. Konstanta fonem merupakan salah satu komponen dari model Klatt. Konstanta fonem terbagi menjadi dua berdasarkan tipe fonem yaitu konstanta fonem vokal dan konsonan. Konstanta fonem vokal terdiri dari 11 tipe dan konsonan 25 tipe suku kata, yang mana konstanta fonem ini dianalisis berdasarkan

tipe suku kata yang terdapat pada tabel 3.3 pada sub bab analisis nilai konstanta. Dari hasil analisis konstanta fonem vokal dan konsonan terdapat fonem yang semua tipe suku katanya bernilai 0.

Tipe suku kata untuk fonem vokal yang memiliki nilai konstantanya 0 yaitu tipe /V8/, /V10/, dan /V11/. Sedangkan tipe suku kata untuk fonem konsonan yang memiliki nilai konstanta 0 sebanyak 10 tipe, diantaranya ada tipe /K12/, /K13/, /K14/, /K19/, /K20/, /K21/, /K22/, /K23/, /K24/, dan /K25/. Hal ini menunjukkan bahwa ada beberapa struktur suku kata yang tidak terdapat pada korpus kalimat Bahasa Melayu Pontianak. Struktur suku kata tersebut yaitu /KVKK/, /KKKV/ dan /KKKVK/.

4.1.3 Implementasi Model Klatt

Model durasi Klatt adalah salah satu model dengan aturan-aturan tertentu untuk memprediksi nilai durasi suatu fonem pada suatu bahasa. Pada penelitian ini, bahasa yang digunakan adalah Bahasa Melayu Pontianak yang merupakan salah satu bahasa daerah yang ada di Kalimantan Barat. Proses prediksi durasi fonem dengan model Klatt terhadap Bahasa Melayu Pontianak ini menggunakan dua aturan yaitu struktur suku kata dan tipe suku kata. Struktur suku kata digunakan untuk memenggal suatu kata menjadi suku kata. Sedangkan tipe suku kata digunakan untuk dapat menentukan nilai tetap atau nilai konstanta setiap fonem berdasarkan tipe yang telah diberikan sesuai posisi suku kata. Prediksi durasi menggunakan *python* yang dijalankan pada *command prompt* seperti yang terlihat pada gambar 4.1.



```

C:\Users\N.K\python-klatt>python sentence_to_SAMPA.py "Salam sepok!"
{'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
@ 88.64
p 55.4
Q 69.4
k 70.0

```

Gambar 4.1 Hasil prediksi durasi fonem dengan model Klatt

Prediksi durasi fonem dengan model Klatt pada gambar 4.1 dimulai dengan memasukkan kalimat Bahasa Melayu Pontianak. Kalimat yang diinputkan yaitu "Salam sepok!".

Kode Program 4.1 Proses prediksi durasi fonem

```

1.  if (len(sys.argv) < 2):
2.      quit()
3.
4.  sentence = sys.argv[1]
5.
6.  #mengubah huruf kapital menjadi huruf kecil
7.  sentence = [letter.lower() for letter in sentence]
8.
9.  #hapus tanda baca
10. sentence = [letter for letter in sentence if letter
    in
    ALLOWED_CHARACTERS]
11. sentence = "".join(sentence)
12.
13. #ubah kalimat menjadi kata
14. sentence = sentence.split(" ")
15.
16. #pemenggalan kata menjadi suku kata
17. sentence = [split_into_syllables(word) for word in
    sentence]
18.
19. #memisahkan suku kata menjadi fonem

```

```

20.         sentence = [[syllable_to_phonemes(syllable) for
    syllable
        in syllables] for syllables in sentence]
21.
22.         #mengkonversi setiap fonem kedalam kamus
23.         sentence = [[convert_phonemes(phonemes) for phonemes
    in
        word] for word in sentence]
24.
25.         for word in sentence:
26.             for syllable in word:
27.                 for phoneme in syllable:
28.                     print(phoneme)
29.
30.         print("")

```

Berikut tahap-tahap proses prediksi durasi fonem menggunakan model Klatt dari gambar 4.1 yang terdapat pada kode program 4.1.

- Setelah kalimat diinputkan, proses yang terjadi mengubah huruf kapital "S" menjadi huruf kecil "s". Perintah untuk melakukan proses ini ada pada baris ke tujuh.
- Kemudian menghapus tanda baca, yang mana pada kalimat yang diinputkan terdapat dua tanda baca yaitu tanda seru (!) dan titik (.). perintah yang digunakan untuk menghapus tanda baca terdapat pada baris ke sepuluh. Karakter-karakter yang hanya digunakan untuk dapat dilakukan pada proses selanjutnya yaitu /a/, /b/, /c/, /d/, /e/, /f/, /g/, /h/, /i/, /j/, /k/, /l/, /m/, /n/, /o/, /p/, /r/, /s/, /t/, /u/, /w/, dan /y/.
- Setelah melakukan penghapusan tanda baca, selanjutnya dilakukan segmentasi kata. Kalimat "salam sepok" dipisah menjadi kata perkata yaitu "salam" dan "sepok". Perintah yang digunakan untuk melakukan proses ini terdapat pada baris ke 14.
- Proses selanjutnya yaitu pemenggalan kata menjadi suku kata. Kata "salam" dan "sepok" dipenggal menjadi suku kata yaitu ['sa', 'lam'] dan ['se', 'pok']. Perintah yang digunakan proses pemenggalan kata terdapat pada baris ke 17.
- Berikutnya memisahkan suku kata menjadi fonem-fonem. Suku kata ['sa', 'lam'] dan ['se', 'pok'] dipisah menjadi fonem-fonem yaitu ['s', 'a', 'l', 'a', 'm'] dan ['s', 'e', 'p', 'o', 'k']. Kalimat "salam sepok" terdiri dari sepuluh

fonem. Perintah untuk melakukan proses ini dapat dilihat pada baris ke 20 dan implementasi suku kata menjadi fonem dapat dilihat pada kode program 4.2.

Kode Program 4.2 Suku kata menjadi fonem-fonem

```

1. def syllable_to_phonemes(syllable):
2.     phonemes = []
3.
4.     pos = 0
5.     while (pos < len(syllable)):
6.         possible_phonemes = malay_phonemes[syllable[pos]]
7.
8.         skip = 1
9.         for phoneme in possible_phonemes:
10.            if pos + len(phoneme) > len(syllable):
11.                continue
12.
13.            if syllable[pos:pos+len(phoneme)] == phoneme:
14.                phonemes.append(phoneme)
15.                skip = len(phoneme)
16.                break
17.
18.        pos += skip
19. return phonemes

```

- f. Setelah memisahkan suku kata menjadi fonem, fonem akan diberikan data tiap-tiap fonem. Data tersebut berupa tipe fonem (type), nilai konstanta fonem (K), nilai minimal fonem (Dmin), nilai rata-rata fonem (Di) dan durasi fonem hasil prediksi dari model Klatt (D0) yang terdapat pada kode program 4.3.

Kode Program 4.3 Memberikan data pada tiap fonem

```

1. def convert_phonemes(phonemes):
2.     phoneme_type = get_phoneme_type(phonemes)
3.
4.     result = []
5.     for i, phoneme in enumerate(phonemes):
6.
7.         constant = duration_constants[phoneme]
8.         [phoneme_type[i]]
9.         d_i = durations[phoneme]['Di']
10.        d_min = durations[phoneme]['Dmin']
11.
12.        result.append({
13.            'phoneme': phoneme,
14.            'type': phoneme_type[i],
15.            'K': constant,
16.            'DMin': d_min,
17.            'Di': d_i,

```

```

17.         'D0': calculate_duration(constant, d_i, d_min)
18.     })
19.
20.     return result

```

Tiap fonem dari kalimat inputan kemudian diperiksa jenis fonemnya seperti yang terda. Jenis fonem terdiri dari fonem vokal dan fonem konsonan. Fonem vokal ada sembilan fonem diantaranya /a/, /e/, /i/, /u/, /o/, /e'/, /ai/, /au/, dan /oi/. Sedangkan fonem konsonan ada 17 fonem diantara /b/,/c/, /d/, /f/, /g/, /h/, /j/, /k/, /l/, /m/, /n/, /p/, /r/, /s/, /t/, /w/, dan /y/. Pengecekan jenis fonem dapat dilihat pada kode program 4.4.

Kode Program 4.4 Pengecekan jenis fonem

```

1. def phoneme_to_code(phoneme):
2.     if phoneme in VOCALS:
3.         return 'V'
4.     else:
5.         return 'K'

```

Selah melakukan pengecekan jenis fonem, selanjutnya jenis fonem dicek berdasarkan posisi suku kata untuk mendapatkan tipe suku kata tiap-tiap fonem pada kode program 4.3 baris ke 13. Tipe suku kata dapat dilihat pada kode program 4.5.

Kode Program 4.5 Tipe suku kata

```

1. indo_phoneme_types = {
2.     "V" : ["V1"],
3.     "VK" : ["V2", "K1"],
4.     "KV" : ["K2", "V3"],
5.     "KVK" : ["K3", "V4", "K4"],
6.     "KKV" : ["K5", "K6", "V5"],
7.     "KKVK" : ["K7", "K8", "V6", "K9"],
8.     "VKK" : ["V7", "K10", "K11"],
9.     "KVKK" : ["K12", "V8", "K13", "K14"],
10.     "KKVKK" : ["K15", "K16", "V9", "K17", "K18"],
11.     "KKKV" : ["K19", "K20", "K21", "V10"],
12.     "KKKVK" : ["K22", "K23", "K24", "V11", "K25"],
13. }

```

Setelah mendapatkan tipe suku kata pada tiap fonem, selanjutnya memberikan nilai konstanta tiap-tiap fonem, nilai durasi rata-rata (Di) dan nilai durasi minimal (Dmin) pada kode program 4.3. baris ke 14 sampai ke 16. Kode program untuk nilai konstanta tiap fonem dapat dilihat pada lampiran H.

Pada program 4.3 dibaris ke 17, durasi fonem menggunakan model klatt (D0) diperoleh dari fungsi perhitungan model Klatt pada kode program.

Kode Program 4.6 Perhitungan prediksi durasi model Klatt

```
1. def calculate_duration(constant, d_i, d_min):
    return constant * (d_i - d_min) + d_min
```

- g. Proses terakhir yaitu mengkonversi fonem kedalam kode SAMPA yang terdapat pada kode program 4.6. Hasil konversi fonem akan diikuti oleh nilai durasi hasil prediksi menggunakan model Klatt.

Kode Program 4.7 Konversi fonem ke kode SAMPA

```
1. sampa = [[[ sampa_codes[phoneme['phoneme']] + " " +
    str(phoneme["D0"]) + " " for phoneme in syllable] for
    syllable in word] for word in sentence]
2.
3. for word in sampa:
4.     for syllable in word:
5.         for phoneme in syllable:
6.             print (phoneme)
```

Kode SAMPA menggunakan *database diphone* Bahasa Indonesia yang dikembangkan oleh bapak Arry Akhmad Arman, seperti yang terlihat pada kode program 4.7.

Kode Program 4.8 Kode SAMPA

```
1. sampa_codes = {
2.     'p': 'p',
3.     'b': 'b',
4.     't': 't',
5.     'd': 'd',
6.     'k': 'k',
7.     'g': 'g',
8.     'c': 'tS',
9.     'j': 'dZ',
10.    'f': 'f',
11.    's': 's',
12.    'z': 'z',
13.    'h': 'h',
14.    'm': 'm',
15.    'n': 'n',
16.    'ng': 'N',
17.    'r': 'r',
18.    'l': 'l',
19.    'w': 'w',
20.    'y': 'j',
```

```

21.    'ny' : 'nY' ,
22.    'a' : 'V' ,
23.    'e\'': '@' ,
24.    'e' : 'e' ,
25.    'i' : 'I' ,
26.    'o' : 'Q' ,
27.    'u' : 'U' ,
28.    'ai' : 'aI' ,
29.    'oi' : 'OI' ,
30.    'au' : 'aU'
31.}

```

4.2 Hasil dan Analisis Pengujian

Pengujian yang dilakukan terdiri dari pengujian standar kecepatan durasi, pengujian perbandingan durasi fonem, pengujian perbandingan durasi kalimat, DMOS dan WER.

4.2.1 Pengujian Standar Kecepatan Durasi

Pengujian kecepatan durasi dilakukan untuk mencari rentang nilai kecepatan durasi yang masih dapat diterima oleh pendengar yang dinyatakan dalam rentang persentase. Pengujian ini dilakukan terhadap satu kalimat yang digunakan sebagai standar untuk penentuan kualitas kecepatan durasi ucapan yang masih dapat diterima. Kecepatan durasi yang diujikan berupa audio rekaman suara penutur Bahasa Melayu Pontianak dengan bunyi kalimat "Lalulah die balek jadi pahlawan Belande". Audio rekaman kalimat ini dipercepat 10% hingga 50% dan diperlambat 10% hingga 50%. Pengujian ini menentukan sampai batas berapa persen pendengar menerima kecepatan normal orang berbicara, apabila kalimat tersebut dipercepat maupun diperlambat.

4.2.1.1 Hasil Pengujian Standar Kecepatan durasi

Hasil pengujian kecepatan durasi dari 30 responden dapat dilihat pada tabel 4.2.

Tabel 4.2 Hasil Uji Kecepatan durasi

Respon den	Inis ial	Nomor									
		1	2	3	4	5	6	7	8	9	10
Kecepatan Ucapan		30%	-50%	10%	-20%	40%	-10%	50%	-30%	20%	-40%
1	BI	1	0	1	1	0	1	0	0	1	1
2	GD	1	0	1	1	0	1	0	0	1	0
3	SHM	1	0	1	1	0	1	0	0	1	0

Respon den	Inis ial	Nomor									
		1	2	3	4	5	6	7	8	9	10
Kecepatan Ucapan		30%	-50%	10%	-20%	40%	-10%	50%	-30%	20%	-40%
4	IH	1	0	1	1	0	1	0	1	1	0
5	EAV	0	0	1	1	0	1	0	0	1	0
6	MAR	0	0	1	1	0	1	0	1	1	1
7	ST	0	0	1	1	0	1	0	1	1	0
8	SJA	0	0	1	1	0	1	0	1	1	1
9	TR	0	0	1	1	0	1	0	1	1	0
10	FAD	0	0	1	0	0	1	0	0	0	0
11	FA	0	0	1	0	0	1	0	0	1	0
12	MN	1	0	1	1	0	1	0	1	0	0
13	MA	1	0	1	1	0	1	0	1	1	1
14	ADP	1	0	1	1	0	1	0	1	1	1
15	FAA	1	0	1	1	0	1	0	1	1	1
16	IMP	0	1	1	1	0	1	0	1	1	1
17	MY	1	0	1	1	0	1	0	1	1	1
18	CL	1	0	1	1	0	1	0	0	1	0
19	WMF	1	0	1	1	0	1	0	0	1	0
20	MRF	0	0	1	1	0	1	0	0	1	0
21	AP	0	0	1	1	0	1	0	0	1	0
22	DK	0	0	1	1	0	1	0	1	1	1
23	ADM	0	0	1	1	0	1	0	1	0	1
24	RZ	0	1	1	1	0	1	0	1	1	1
25	CLO	1	0	1	1	0	1	0	1	1	0
26	DO	1	1	1	1	0	1	0	1	1	1
27	JS	0	1	1	1	0	1	0	0	0	1
28	HNN	0	0	1	1	0	1	0	1	0	0
29	BL	0	0	1	1	0	1	0	1	1	0
30	H	1	0	1	1	0	1	0	1	1	1
Rata-rata		47%	13%	100%	93%	0%	100%	0%	63%	83%	47%

Pada tabel 4.2 diperoleh hasil pengujian kecepatan ucapan menurut 30 responden. Untuk kecepatan ucapan yang diperlambat 30% pada nomor satu, rata-rata persentase responden yang dapat menerima kecepatan ucapan sebesar 47%. Nomor dua ucapan dipercepat -50%, responden yang dapat menerima ucapan dengan kecepatan ucapan tersebut sebesar 13%. Nomor tiga ucapan diperlambat 10%, jumlah responden yang dapat menerima apabila orang berbicara dengan kecepatan tersebut sebanyak 100%. Kecepatan ucapan dipercepat -20% untuk nomor empat, responden yang dapat menerima sebanyak 93%. Nomor lima kecepatan ucapan diperlambat 40%, hasil rata-rata responden sebesar 0%.

Kecepatan ucapan dipercepat -10% untuk nomor enam, jumlah responden yang dapat menerima 100%. Nomor tujuh kalimat diperlambat 50%, hasil rata-rata pendapat responden 0%. Nomor delapan kalimat dipercepat -30%, hasil

pendapat responden 63%. Nomor sembilan kalimat diperlambat 20%, jumlah responden yang dapat menerima 83%. Ketika kalimat dipercepat -40% pada nomor 10, jumlah responden yang dapat menerima apabila orang berbicara dengan kecepatan tersebut sebesar 47%.

4.2.1.2 Analisis Hasil Pengujian Kecepatan durasi

Dari hasil pengujian kecepatan durasi untuk kualitas kecepatan ucapan yang diperoleh dari pendapat 30 responden dapat dirangkum pada tabel diurutkan dari hasil yang tertinggi hingga terendah.

Tabel 4.3 Analisis Hasil Uji

Range Hasil Uji Responden	Kecepatan Ucapan (%)	Rata-rata Hasil Uji Responden (%)
80% - 100% [sangat baik]	-10%	100%
	10%	100%
	-20%	93%
	20%	83%
70% - 79% [baik]	-	-
60% - 69% [cukup baik]	-30%	63%
50% - 59% [tidak baik]	-	-
0% - 49% [sangat tidak baik]	-40%	47%
	30%	47%
	-50%	13%
	50%	0%
	40%	0%

Tabel 4.3 merupakan analisis hasil uji standar kecepatan durasi. Pada rentang 80% hingga 100% diperoleh rata-rata hasil uji responden yaitu 83%, 93%, dan 100% pada kecepatan ucapan 20%, -20%, 10% dan -10%. Rentang 70% hingga 79% tidak diperoleh hasil uji responden. Rentang 60% hingga 69% diperoleh rata-rata hasil uji responden yaitu 63% pada kecepatan ucapan -30%. Rentang 50% hingga 59% tidak diperoleh hasil uji responden. Rentang 0% hingga 49% diperoleh rata-rata hasil uji responden yaitu 0%, 13% dan 47% pada kecepatan ucapan 40%, 50%, -50% 30% dan -40%. Maka dari hasil analisis pengujian kecepatan durasi ucapan, responden dapat menerima apabila orang berbicara dengan kategori sangat baik pada kecepatan ucapan 20%, -20%, 10%,

dan -10%.

4.2.2 Pengujian Perbandingan Nilai Durasi

Pengujian perbandingan durasi dilakukan dua kali perbandingan. Pengujian ini bertujuan untuk membandingkan durasi penutur dan durasi klatt.

1. Perbandingan durasi klatt dan durasi penutur berdasarkan durasi fonem terhadap 10 kalimat uji.
2. Perbandingan durasi klatt dan durasi penutur berdasarkan durasi kalimat terhadap 100 kalimat uji.

Contoh menghitung perbandingan durasi baik berdasarkan durasi fonem maupun durasi kalimat sebagai berikut. Contohnya fonem "a" untuk durasi Klatt 120ms dan durasi penutur 95ms.

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

$$\text{Persentase perbandingan} = \frac{120 \text{ ms} - 95 \text{ ms}}{95 \text{ ms}} \times 100$$

$$\text{Persentase perbandingan} = \frac{25 \text{ ms}}{95 \text{ ms}} \times 100$$

$$\text{Persentase perbandingan} = 0,26 \times 100 = 26 \%$$

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

Jadi hasil perbandingan durasi yang terdapat pada contoh sebesar 26% yaitu durasi klatt lebih lambat dari durasi penutur.

4.2.2.1 Hasil Pengujian Perbandingan Nilai Durasi Berdasarkan Fonem

Pengujian perbandingan durasi berdasarkan fonem terhadap 10 kalimat uji, terlebih dahulu dilakukan perbandingan nilai durasi berdasarkan kalimat. Hasil perbandingan durasi klatt dan durasi penutur berdasarkan kalimat terhadap 10 kalimat uji dapat dilihat pada tabel 4.4.

Tabel 4.4 Hasil Perbandingan Durasi Terhadap 10 Kalimat Uji

No	Kalimat	Jumlah Durasi Penutur(ms)	Jumlah Durasi Klatt(ms)	Perbedaan (%)
1	Aok ingat benar aku.	1260	1212	4%
2	Manelah negare kau tuh nak maju kalok tebiat pemerintahe tak tentu rudu macam itu.	5555	5308	-4%
3	Pade suatu hari kamek pegi ke kubu.	2434	2317	-5%
4	Bende ini ni ba tanggung jawab kitak semue.	3172	2731	-14%
5	Mang jago mang gugel ni.	1391	1408	1%
6	Aku sengaja turun lamak-lamak.	1925	1877	-2%
7	Lalu aku pon keluar dari pesawat.	1699	2148	26%
8	Jeket lelong kote baru ni kebetolan pulak warne-e jingge.	3203	3533	10%
9	Tiket pon udah dibeli.	1373	1436	5%
10	Sehe karang kau kenak pangkong pakek kayuk same orang belande.	3063	3738	22%

Tabel 4.4 merupakan hasil perbandingan durasi berdasarkan kalimat terhadap 10 kalimat uji. Dari 10 kalimat uji, persentase kecepatan durasi paling lambat sebesar 26% pada kalimat ke tujuh. Persentase kecepatan durasi paling cepat terdapat pada kalimat ke enam sebesar -14%. Kalimat uji tersebut dirincikan lagi perbandingan nilai durasi berdasarkan nilai durasi fonem dalam kalimat. Hasil pengujian perbandingan nilai durasi berdasarkan durasi fonem dapat dilihat pada lampiran I.

4.2.2.2 Analisis Hasil Pengujian Perbandingan Nilai Durasi Fonem

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

1. Hasil perbandingan nilai durasi berdasarkan durasi kalimat

Tabel 4.5 Hasil Analisis 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%	4	40%

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

Tabel 4.5 diperoleh hasil kalimat yang masuk dalam kecepatan durasi yang dapat diterima oleh pendengar berdasarkan hasil uji standar kecepatan durasi pada tabel 4.3 dalam *range* kecepatan durasi ucapan 20% hingga -20% pada kategori sangat baik adalah sebanyak 80% kalimat dari 10 kalimat uji.

2. Hasil perbandingan durasi berdasarkan durasi fonem

Hasil persentase perbandingan nilai durasi fonem yang terdapat pada lampiran I diurutkan dari nilai terkecil hingga terbesar. Hasil dapat dilihat pada tabel 4.6.

Tabel 4.6 Analisis Hasil Perbandingan Durasi Fonem

<i>Range</i> Kecepatan 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%

Tabel 4.6 diperoleh hasil fonem yang masuk dalam kecepatan durasi yang dapat diterima oleh pendengar berdasarkan hasil uji standar kecepatan durasi ucapan pada tabel 4.3 dalam *range* kecepatan durasi ucapan 20% hingga -20% pada kategori sangat baik adalah sebanyak 95 fonem (vokal maupun konsonan) dengan persentase 30% dari 320 fonem terhadap 10 kalimat uji. Dari hasil tersebut, model Klatt belum dapat memprediksi durasi masing-masing fonem (vokal maupun konsonan) berdasarkan standar kecepatan durasi yang dapat

diterima oleh pendengar. Perbandingan jumlah fonem vokal dan konsonan yang mempengaruhi hasil perbandingan durasi fonem yang dapat diterima dan tidak dapat diterima oleh pendengar terhadap 320 fonem pada 10 kalimat uji dapat dilihat pada tabel 4.7.

Tabel 4.7 Perbandingan Jumlah Fonem

Fonem (diterima)		Fonem (tidak dapat diterima)	
Vokal (%)	Konsonan (%)	Vokal (%)	Konsonan (%)
40,8%	59,2%	47,7%	52,3%

Tabel 4.7 diperoleh persentase jumlah fonem vokal dan konsonan dengan durasi fonem yang dapat diterima oleh pendengar yaitu untuk fonem vokal sebesar 40,8% dan untuk fonem konsonan sebesar 59,2%. Sedangkan persentase jumlah fonem dengan durasi fonem yang tidak dapat diterima yaitu untuk fonem vokal sebesar 47,7% dan untuk fonem konsonan sebesar 52,3%.

Pada hasil perbandingan tersebut diperoleh kecepatan durasi klatt yang lebih lambat atau lebih cepat dari durasi penutur. Hal ini dikarenakan nilai durasi fonem untuk model Klatt baik vokal maupun konsonan bervariasi. Hal ini dikarenakan nilai tiap-tiap fonem diperoleh dari nilai konstanta fonem, yang mana nilai konstanta fonem ini ditentukan berdasarkan posisi fonem pada suatu kata (tipe suku kata).

4.2.2.3 Hasil Pengujian Perbandingan Nilai Durasi Kalimat

Tahap menghitung perbandingan durasi kalimat sama dengan menghitung perbandingan durasi fonem. Hasil pengujian perbandingan nilai durasi berdasarkan kalimat terhadap 100 kalimat uji dapat dilihat pada lampiran J. Rangkuman hasil pengujian perbandingan durasi kalimat dapat dilihat pada tabel 4.8.

Tabel 4.8 Rangkuman Hasil Pengujian Perbandingan Durasi

Keterangan	Persentase Tinggi	Persentase Rendah	Jumlah Kalimat
< 0%	-1%	-23%	50
0%	0%	0%	3
> 0%	26%	1%	47
Total Jumlah Kalimat			100

Tabel 4.8 merupakan rangkuman hasil pengujian perbandingan durasi klatt

dan durasi penutur terhadap 100 kalimat uji. Durasi Klatt lebih lambat dari durasi penutur sebanyak 47 kalimat. Perbandingan nilai durasi tertinggi sebesar 26% pada kalimat yang berbunyi "lalu aku pon keluar dari pesawat" yang mana jumlah total durasi fonem penutur sebesar 1699ms dan durasi klatt sebesar 2148ms. Sedangkan perbandingan nilai durasi terendah sebesar 1% pada kalimat "payah dah kalok dah ngumongkan kantor pusat" yang memiliki jumlah durasi penutur sebesar 2657ms dan durasi klatt 2677ms. Untuk perbandingan nilai durasi Klatt sama atau 0% dengan nilai durasi penutur sebanyak 3 kalimat, salah satu kalimatnya berbunyi "Sepok aku pon mulai keluar" dengan jumlah durasi penutur sebesar 1741ms dan jumlah durasi Klatt 1747ms.

Sedangkan untuk hasil perbandingan durasi Klatt lebih cepat dari durasi Penutur sebanyak 50 kalimat. Perbandingan nilai durasi tertinggi sebesar -1% pada kalimat yang berbunyi "cuman satu kate yak yang ade tang dalam otak aku" memiliki durasi penutur sebesar 2986ms dan durasi klatt 2961ms. Persentase perbandingan terendah sebesar -23% pada kalimat yang berbunyi "warne lengkap bendere tu jingge tue , puteh biru , biru" memiliki jumlah durasi penutur 4343ms dan durasi klatt 3336ms.

4.2.2.4 Analisis Hasil Pengujian Perbandingan Nilai Durasi Kalimat

Pengujian perbandingan durasi dianalisis terhadap rentang hasil perbandingan durasi yang digunakan sebagai standar pengujian kejelasan kalimat.

Tabel 4.9 Hasil Analisa Pengujian Perbandingan Nilai Durasi

<i>Range</i> Kecepatan Durasi	Jumlah Kalimat	Persentase Kalimat
41% - 50%	-	-
31% - 40%	-	-
21% - 30%	3	3%
11% - 20%	9	9%
1% - 10%	35	35%
0%	3	3%
(-1%) - (-10%)	30	30%
(-11%) - (-20%)	17	17%
(-21%) - (-30%)	3	3%
(-31%) - (-40%)	-	-
(-41%) - (-50%)	-	-
Total Jumlah	100	100%

Pada tabel 4.9 pada rentang kurang dari 10% hingga lebih dari -10%

adalah standar kualitas ucapan yang telah ditentukan untuk kejelasan kalimat uji WER. Hasil analisis pengujian perbandingan durasi penutur dan durasi klatt terhadap durasi kalimat dengan jumlah kalimat uji 100 kalimat menyatakan 50% perbandingan durasi kalimat dibawah (-) atau $< 0\%$. Maka pada pengujian ini durasi klatt pada kalimat uji 100 kalimat lebih cepat dari pada durasi penutur. Sedangkan untuk jumlah kalimat yang dapat diterima oleh pendengar berdasarkan hasil uji kecepatan durasi untuk kecepatan durasi pada tabel 4.4 dalam *range* kecepatan durasi ucapan -20% hingga 20% pada kategori sangat baik adalah sebanyak 94% kalimat dari 100 kalimat uji. Hasil tersebut menunjukkan bahwa model durasi Klatt dapat memprediksi durasi berdasarkan total durasi fonem pada tiap kalimat dengan kecepatan durasi ucapan yang dapat diterima oleh pendengar.

4.2.3 Pengujian DMOS

Pengujian DMOS (Degradation Mean Opinion Score) bertujuan untuk menguji tingkat kemiripan yang dihasilkan sintesa ucapan terhadap panjang pendeknya durasi suatu kalimat pada durasi klatt dan durasi penutur.

4.2.2.1 Hasil Pengujian DMOS

Contoh perhitungan DMOS pada responden nomor 1 adalah sebagai berikut.

$$DMOS = \frac{\sum_{i=0}^n x(i) \cdot k}{N}$$

$$DMOS = \frac{x_1 + x_2 + \dots + x_{10}}{10}$$

$$DMOS = \frac{4+3+3+3+4+4+2+3+2+4}{10}$$

$$DMOS = \frac{32}{10} = 3.2$$

Hasil DMOS yang diperoleh dari responden nomor 1 adalah 3.2. Nilai DMOS ini akan dikonversi menjadi skala ITU-800.1. Contoh hasil konversi DMOS pada responden nomor 1 adalah sebagai berikut.

$$DMOS_{ITU-800.1} = \frac{5}{4} DMOS$$

$$DMOS_{ITU-P.800.1} = \frac{5}{4} \times 3.2$$

$$DMOS_{ITU-P.800.1} = \frac{5}{4} \times 3.2 = 4$$

Hasil DMOS dari responden nomor 1 adalah 4 (baik). Konversi dapat dilakukan setelah DMOS pada satu tabel dirata-ratakan atau dikonversi terlebih dahulu kemudian dirata-ratakan untuk mendapatkan hasil akhir nilai DMOS. Hasil pengujian DMOS terhadap 10 kalimat uji dapat dilihat pada tabel 4.10.

Tabel 4.10 Hasil Pengujian DMOS

Responden	Inisial	Nomor Kalimat									
		1	2	3	4	5	6	7	8	9	10
1	IMP	4	3	3	3	4	4	2	3	2	4
2	ADP	4	3	3	3	4	4	4	3	3	3
3	YUP	4	3	3	3	4	4	2	3	2	4
4	FAA	4	3	3	3	4	4	4	3	3	3
5	SHM	4	2	3	2	3	3	3	2	2	2
6	IMY	4	3	3	3	4	4	4	3	2	2
7	YFF	4	3	3	3	4	4	4	3	2	2
8	TK	4	2	3	3	4	4	3	4	2	4
9	LL	4	2	3	3	4	4	3	4	2	4
10	GD	3	2	3	2	4	3	3	2	2	2
11	S	3	2	3	2	4	3	3	3	2	4
12	EAV	4	2	3	3	4	3	2	3	2	3
13	MAR	3	2	3	2	4	3	2	3	2	4
14	FA	3	2	2	3	4	4	3	3	2	4
15	A	4	3	3	2	3	3	3	3	2	3
16	MN	3	2	3	2	4	4	3	3	3	4
17	SJA	4	3	4	3	4	4	3	2	2	4
18	TR	3	3	4	2	4	3	3	2	3	4
19	AP	4	3	3	2	4	3	4	3	3	4
20	ADM	3	2	3	3	3	4	3	3	2	3
21	DK	4	3	2	3	4	3	2	3	3	2
22	IH	4	3	3	3	4	3	2	3	2	3
23	CL	4	2	3	3	4	3	3	4	2	4
24	R	4	2	3	2	4	3	3	2	2	4
25	NN	4	3	3	3	4	4	3	4	2	4
26	M	4	3	3	3	4	4	2	3	2	2
27	SP	4	2	3	2	4	4	3	2	2	2
28	YH	4	3	3	3	4	4	2	3	2	2
29	KN	4	2	3	2	4	4	3	3	2	2
30	AG	4	3	4	3	4	4	3	2	2	2
Rata-rata DMOS		3.7 7	2.5 3	3.03	2.63	3.9 0	3.6 0	2.9 0	2.90	2.2 0	3.13

Responden	Inisial	Nomor Kalimat									
		1	2	3	4	5	6	7	8	9	10
Total Rata-rata		3.06									

Nilai rata-rata DMOS tiap kalimat uji bervariasi, untuk kalimat nomor satu nilai DMOS sebesar 3.77, kalimat dua 2.53, kalimat tiga 3.03, kalimat empat 2.63, kalimat lima 3.90, kalimat enam 3.60, kalimat tujuh 2.90, kalimat delapan 2.90, kalimat sembilan 2.20, dan kalimat ke sepuluh sebesar 3.13. Nilai DMOS tertinggi terdapat pada kalimat uji nomor 5 dengan nilai sebesar 3.90. Sedangkan untuk nilai DMOS terendah terdapat pada kalimat uji nomor 9 dengan nilai sebesar 2.20.

4.2.2.2 Analisis Hasil Pengujian DMOS

Kualitas ucapan yang diuji pada pengujian DMOS adalah tingkat kemiripan durasi yang diucapkan oleh sintesa ucapan durasi Klatt terhadap sintesa ucapan durasi penutur. Kemiripan durasi Klatt terhadap durasi penutur dapat direpresentasikan dengan lama waktu pengucapan tiap kata dan kalimat. Pada penilaian yang dilakukan oleh 30 orang responden terhadap pengujian perbandingan kecepatan durasi klatt terhadap durasi penutur dengan nilai total rata-rata 3.06. Total nilai rata-rata tersebut dikonversi menjadi skala ITU-800.1 sebagai berikut.

$$DMOS_{ITU-P.800.1} = \frac{5}{4} DMOS$$

$$DMOS_{ITU-P.800.1} = \frac{5}{4} \times 3.06$$

$$DMOS_{ITU-P.800.1} = \frac{5}{4} \times 3.06 = 3.83$$

Hasil DMOS setelah dikonversi sebesar 3.83. Berdasarkan nilai DMOS yang telah dikonversi, maka kecepatan durasi Klatt yang diucapkan hasil sintesa mendekati durasi penutur berdasarkan tabel kriteria penilaian hasil rata-rata DMOS pada tabel 3.21.

4.2.4 Pengujian WER

Pengujian WER berupa pengujian dikte atau menulis kembali apa yang telah diperdengarkan. Kalimat pengujian WER diambil dari hasil analisis

pengujian perbandingan durasi berdasarkan durasi kalimat terhadap 100 kalimat uji. Kalimat yang digunakan sebanyak 10 kalimat yang dipilih secara acak terdapat pada range kecepatan durasi yaitu 20% hingga -20%. Kalimat uji WER dapat dilihat pada tabel 4.11.

Tabel 4.11 Kalimat Uji WER

No	Kalimat Uji	Kecepatan Durasi Kalimat (%)
1	Naek gunong tak suah tapi punye tas gunong.	0%
2	Sebelum aku sampai tang belande ni pon aku tak tau gak.	1%
3	Hamper nak limak belas ribu.	5%
4	Lalu tak sengaja aku ketemu ngan bapak-bapak yang kayak-e satu pesawat ngan aku.	7%
5	Ade jalan husus untok sepeda dibuatkan.	10%
6	Cuman satu kate yak yang ade tang dalam otak aku.	-1%
7	tepoto sanak tepoto sinek.	-4%
8	Ade duak momen dimane belande berubah jadi jingge.	-5%
9	Karne biar cemane pon die haros pergi ke sepanyol.	-6%
10	Aku pun tau , gugel yang ngabarkan aku.	-9%

4.2.3.1 Hasil Pengujian WER

Perhitungan WER menggunakan rumus 3.5 dan akurasi WER menggunakan rumus 3.6. Contoh perhitungan WER pada responden nomor 1 untuk urutan kalimat 1 adalah sebagai berikut.

K_0 = “naek gunong tak suah tapi punye tas gunong”

K_1 = “naik gunung kan suah kamek punye tas gunong”

N_r = [“naek”, “gunong”, “tak”, “suah”, ”tapi”, ”punye”, ”tas”, ”gunong”] = 8

I = [”kan”, ”kamek”] = 2

S = [“naek” => “naik”, “gunong” => “gunung”] = 2

D = [“tak”, “tapi”] = 2

$$WER = \frac{S+D+I}{N_r} = \frac{2+2+2}{8} = \frac{6}{8} = 0.75 \text{ Akurasi WER} = 1 - 0.75 = 0.25 = 25\%$$

Hasil perhitungan akurasi WER yang didapatkan oleh responden nomor 1 untuk urutan kalimat 1 adalah 0.25 atau 25 %. Hasil pengujian WER dapat dilihat pada lampiran **K** dan dirangkum pada tabel 4.12. Hasil Akurasi WER dari 30 responden terhadap 10 kalimat uji sebesar 59%.

Tabel 4.12 Rangkuman Hasil Pengujian WER

Nomor	Kecepatan Durasi Kalimat (%)	Akurasi WER/ Kalimat
1	0%	48%
2	1%	23%
3	5%	73%
4	7%	47%
5	10%	61%
6	-1%	72%
7	-4%	61%
8	-5%	55%
9	-6%	58%
10	-9%	89%
Rata-rata Akurasi WER		59%

4.2.3.2 Analisis Hasil Pengujian WER

Pengujian WER bertujuan untuk menilai tingkat kejelasan dari sintesa ucapan durasi Klatt. Dari hasil pengujian WER, ada beberapa hal yang menjadi faktor pengganggu dalam melihat kejelasan sintesa ucapan menggunakan durasi klatt yaitu kecepatan durasi, menggunakan *diphone* bahasa Indonesia sebagai pembangkit ucapan, dan ucapan yang dibangkitkan tidak menggunakan nilai nada ucapan (pitch). Dari hasil akurasi WER 59%, maka tingkat kejelasan sintesa ucapan menurut pendapat responden terhadap 10 kalimat uji yang dibangkitkan dengan durasi Klatt adalah cukup jelas.

BAB V

PENUTUP

5.1 Kesimpulan

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

1. Berdasarkan hasil pengujian kalimat standar, pendengar dapat menerima ketika pembicara berbicara dengan kategori sangat baik pada kecepatan 20%, -20%, 10% dan -10%.
2. Berdasarkan pengujian perbandingan nilai durasi.
 - a. 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).
 - b. Hasil perbandingan durasi kalimat terhadap 100 kalimat uji diperoleh 94% kalimat yang termasuk dalam rentang kecepatan durasi ucapan pada kategori sangat baik.
3. Berdasarkan hasil pengujian *Degradation Mean Opinion Score* (DMOS), tingkat kemiripan durasi ucapan yang dihasilkan sintesa ucapan dalam mengucapkan suatu kalimat antara durasi klatt dan durasi penutur memiliki nilai DMOS sebesar 3.83 yaitu durasi klatt ucapan sintesa mendekati durasi penutur ucapan sintesa berdasarkan rentang skala 4 (3.01-4.00).
4. Berdasarkan hasil pengujian *Word Error Rate* (WER), nilai akurasi yang dihasilkan sebesar 59% menunjukkan bahwa tingkat kejelasan kalimat uji yang dibangkitkan dengan sintesa ucapan untuk model durasi Klatt adalah cukup jelas.

5. Berdasarkan hasil pengujian tersebut, model durasi Klatt yang menggunakan aturan suku kata dan tipe suku kata dapat memprediksi durasi fonem total sintesa ucapan pada Bahasa Melayu Pontianak namun masih perlu penelitian lanjut untuk memprediksi durasi masing-masing fonem.

5.1 Saran

Hal yang perlu ditambahkan dalam penelitian implementasi durasi Klatt pada Bahasa Melayu Pontianak, yaitu dengan menghubungkan nilai durasi fonem terhadap nilai *pitch* pada setiap fonem. Penambahan nilai *pitch* ini didasarkan pada, ketika orang berbicara dengan nada (*pitch*) yang tinggi biasanya durasinya cenderung lebih pendek dibandingkan ketika orang berbicara dengan nada (*pitch*) yang rendah biasanya durasinya cenderung lebih panjang.