

**SISTEM *TEXT-TO-SPEECH* BAHASA INDONESIA
BERBASIS HIDDEN MARKOV MODEL**

Laporan Tugas Akhir

Disusun sebagai syarat kelulusan tingkat sarjana

Oleh

PATRICK LUMBAN TOBING

NIM : 13510013



**PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TEKNIK ELEKTRO & INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG
JUNI 2014**

**SISTEM *TEXT-TO-SPEECH* BAHASA INDONESIA
BERBASIS HIDDEN MARKOV MODEL**

Laporan Tugas Akhir

Oleh

PATRICK LUMBAN TOBING

NIM : 13510013

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung

Telah disetujui dan disahkan sebagai Laporan Tugas Akhir
di Bandung, pada tanggal 20 Juni 2014

Pembimbing,

Dr. Dessi Puji Lestari, S.T., M.Eng

NIP. 197912012012122005

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Pengerjaan dan penulisan Tugas Akhir ini dilakukan tanpa menggunakan bantuan yang tidak dibenarkan.
2. Segala bentuk kutipan dan acuan terhadap tulisan orang lain yang digunakan di dalam penyusunan Laporan Tugas Akhir ini telah dituliskan dengan baik dan benar.
3. Tugas Akhir ini belum pernah diajukan pada program pendidikan di perguruan tinggi mana pun.

Jika terbukti melanggar hal-hal di atas, saya bersedia dikenakan sanksi sesuai dengan Peraturan Akademik dan Mahasiswa Institut Teknologi Bandung bagian Penegakan Norma Akademik dan Mahasiswa khususnya Pasal 2.1 dan Pasal 2.2.

Bandung, 20 Juni 2014

Patrick Lumban Tobing

NIM 13510013

ABSTRAK

Sistem *text-to-speech* adalah sebuah sistem yang menerima masukan berupa teks dan dapat memberikan keluaran berupa gelombang suara sebagai representasi pengucapan dari teks tersebut sesuai bahasa yang diimplementasi oleh sistem. Pada Tugas Akhir ini, dikembangkan sebuah sistem *text-to-speech* (TTS) untuk bahasa Indonesia. Salah satu pendekatan yang terkenal dalam pengembangan sistem TTS adalah dengan *hidden Markov model* (HMM). Melalui pendekatan HMM ini, sistem TTS yang dikembangkan akan memiliki ukuran yang ramping, suara yang stabil serta nantinya dapat dengan mudah mengimplementasikan berbagai variasi karakteristik suara.

Dalam pengembangannya, sistem TTS bahasa Indonesia berbasis *Hidden Markov Model* (HMM) sudah pernah dikembangkan oleh (Sakti, 2008; Vania, 2011). Pada Tugas Akhir ini, sistem yang dikembangkan memiliki tiga konfigurasi, yaitu terkait penggunaan daftar konteks. Dua di antaranya mengikuti (Sakti, 2008) dan (Vania, 2011). Satu konfigurasi lainnya merupakan kombinasi dari keduanya. Selain memvariasikan konteks, Tugas Akhir ini juga melihat pengaruh dari penggunaan *question set* yang berbeda dalam pembentukan *decision tree* untuk *clustering model* HMM.

Sistem TTS yang dikembangkan pada Tugas Akhir ini dievaluasi dengan menggunakan metode *Mean Opinion Score* (MOS) dan *Semantically Unpredictable Sentences* (SUS). Evaluasi MOS memberikan nilai tertinggi 3,84 untuk konfigurasi konteks (Vania, 2011). Evaluasi MOS ini juga menunjukkan bahwa perbaikan *question set* dapat meningkatkan kualitas ucapan, dari nilai MOS 3,76 menjadi 3,79 untuk konfigurasi konteks (Sakti, 2008). Sedangkan konfigurasi konteks ketiga memberikan nilai MOS terendah. Pada evaluasi SUS, penggunaan set konteks yang berbeda juga menunjukkan perbedaan akurasi, di mana konfigurasi konteks (Sakti, 2008) memberikan akurasi kalimat 87,33% dan akurasi kata 96,19%. Sedangkan konfigurasi konteks (Vania, 2011) memberikan nilai lebih tinggi dengan akurasi 92,67% untuk kalimat dan 98,25% untuk kata.

Kata kunci: sistem *text-to-speech*, *hidden Markov model*, bahasa Indonesia, variasi konteks, variasi *question decision tree*, kualitas ucapan, akurasi ucapan, MOS, SUS

KATA PENGANTAR

Puji syukur kepada Tuhan yang Maha Esa karena dengan berkat dan anugerahnya pengerjaan Tugas Akhir ini dapat diselesaikan.

Penulis juga mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Dr. Dessi Puji Lestari, S.T., M.Eng, selaku pembimbing Tugas Akhir
2. Dr.-Ing. Sakriani Sakti, yang memberikan berbagai panduan yang sangat berharga terkait Tugas Akhir ini
3. Dr. Eng. Ayu Purwarianti S. T., M. T., selaku Ketua Program Studi IF yang juga memberi dukungan terkait Tugas Akhir
4. Dr. Ir. Gusti Ayu Putri Saptawati, M. Comm, selaku dosen wali selama tiga tahun penulis di jurusan IF
5. Keluarga, yang selalu memberi dukungan dan doa
6. Teman-teman di kosan, yang selalu menjadi penebar tawa dan memberi banyak pelajaran serta inspirasi bagi penulis
7. Teman-teman di jurusan, yang mendukung dalam proses akademis
8. Semua orang yang telah membantu proses Tugas Akhir ini, tetapi tidak disadari oleh penulis

Pekerjaan Tugas Akhir ini diharapkan dapat menjadi dasar referensi bagi penelitian di area ini dalam bahasa Indonesia ke depannya. Penulis menyadari bahwa tak ada hal yang tidak memiliki kelemahan dan cacat di dunia ini. Oleh karena itu, penulis sangat menerima berbagai kritik dan saran terhadap pekerjaan Tugas Akhir ini.

Bandung, Juni 2014

Penulis

DAFTAR ISI

ABSTRAK	iv
KATA PENGANTAR.....	v
DAFTAR ISI.....	vi
DAFTAR LAMPIRAN	ix
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xii
BAB I. PENDAHULUAN.....	1
I.1. Latar Belakang.....	1
I.2. Rumusan Masalah.....	2
I.3. Tujuan	3
I.4. Batasan Masalah	3
I.5. Metodologi.....	3
I.6. Sistematika Pembahasan.....	4
BAB II. STUDI LITERATUR	6
II.1. <i>Hidden Markov Model</i> (Masuko, 2002)	6
II.1.1. Definisi.....	6
II.1.2. Kalkulasi Probabilitas <i>Observation</i>	8
II.1.3. Kalkulasi Sekuens <i>State Optimal</i>	9
II.1.4. Estimasi Parameter Model	9
II.2. Sistem <i>Text-to-Speech</i> berbasis <i>Hidden Markov Model</i>	9
II.2.1. Pemroses teks.....	11
II.2.2. Pemroses sinyal ucapan	12

II.2.3.	Proses <i>training</i> dan <i>synthesis</i>	14
II.2.4.	Kakas pendukung sistem TTS berbasis HMM	18
II.3.	Tata Bahasa Indonesia	19
II.3.1.	Daftar Fonem Bahasa Indonesia	20
II.3.2.	Aturan Pemenggalan Suku Kata	21
II.3.3.	Hal-hal yang Perlu Diperhatikan pada Teks Bahasa Indonesia	22
II.4.	Evaluasi Sistem	23
II.5.	Penelitian terkait sistem TTS untuk bahasa Indonesia	24
BAB III. ANALISIS DAN PERANCANGAN SISTEM TEXT-TO-SPEECH BERBASIS HMM UNTUK BAHASA INDONESIA		27
III.1.	Akuisisi data ucapan.....	29
III.2.	Segmentasi durasi.....	29
III.3.	Pemrosesan teks	30
III.4.	Ekstraksi parameter ucapan.....	34
III.5.	<i>Training</i> HMM	35
III.6.	Sintesis ucapan	36
III.7.	Evaluasi sistem	36
BAB IV. IMPLEMENTASI SISTEM PENSINTESIS UCAPAN BAHASA INDONESIA BERBASIS HIDDEN MARKOV MODEL		38
IV.1.	Spesifikasi data ucapan	41
IV.2.	Implementasi proses segmentasi ucapan	41
IV.3.	Implementasi pemrosesan teks	44
IV.3.1.	<i>Phoneset</i> untuk bahasa Indonesia	46
IV.3.2.	Aturan pemisahan suku kata untuk bahasa Indonesia	48
IV.3.3.	Aturan <i>letter-to-sound</i> untuk bahasa Indonesia	48

IV.3.4.	Pelabelan konteks terhadap fonem	50
IV.4.	Implementasi proses ekstraksi fitur ucapan	54
IV.5.	Implementasi proses <i>training</i> HMM	56
IV.6.	Implementasi proses <i>synthesis</i> ucapan	63
BAB V. EVALUASI SISTEM SINTESIS UCAPAN		65
V.1.	Hasil tes <i>Mean Opinion Score</i> (MOS)	66
V.2.	Hasil tes <i>Semantically Unpredictable Sentence</i> (SUS)	68
BAB VI. PENUTUP		74
VI.1.	Kesimpulan.....	74
VI.2.	Saran.....	74
DAFTAR REFERENSI		75

DAFTAR LAMPIRAN

LAMPIRAN A. Question Set 1 untuk Konteks (Sakti, 2008)	77
A.I. Bagian set pertanyaan terkait konteks durasi	77
A.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.	83
LAMPIRAN B. Question Set 2 untuk Konteks (Sakti, 2008)	87
B.I. Bagian set pertanyaan terkait konteks fonetis.....	87
B.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.	87
LAMPIRAN C. Question Set 3 untuk Konteks (Vania, 2008).....	96
C.I. Bagian set pertanyaan terkait konteks fonetis.....	96
C.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.	96
LAMPIRAN D. Question Set 4 untuk Kombinasi Konteks (Sakti, 2008) dan (Vania, 2011).....	103
D.I. Bagian set pertanyaan terkait konteks fonetis	103
D.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.	103
LAMPIRAN E. Daftar Kalimat yang Digunakan pada Evaluasi MOS	104
LAMPIRAN F. Daftar Kalimat yang Digunakan pada Evaluasi SUS	105

DAFTAR GAMBAR

Gambar II-1 Struktur HMM <i>3-state left-to-right</i> (Masuko, 2002).....	7
Gambar II-2 Diagram Sistem HMM-based <i>Text-to-Speech</i> (Zen, 2007)	10
Gambar II-3 Proses Pembentukan Ucapan pada Manusia (Zen, 2010)	13
Gambar II-4 Model Komputasi Pembentukan Sinyal Ucapan (Masuko, 2002) ...	14
Gambar II-5 Proses Training pada HTS (Zen, 2010).....	15
Gambar II-6 Skema Proses <i>Synthesis</i> pada Sistem HMM-based <i>Speech Synthesis</i> (Zen, 2007).....	17
Gambar II-7 Daftar konteks pada penelitian (Sakti, 2008)	25
Gambar II-8 Daftar konteks pada penelitian (Vania, 2011).....	25
Gambar III-1 Arsitektur sistem TTS berbasis HMM.....	27
Gambar III-2 Tahapan rancangan pembuatan sistem.....	28
Gambar III-3 Tahapan pemrosesan teks	33
Gambar III-4 Proses <i>training</i> HMM	35
Gambar IV-1 Tahapan proses implementasi sistem TTS	38
Gambar IV-2 Penggunaan kakas dalam implementasi dari sistem TTS.....	40
Gambar IV-3 Gambar atas adalah hasil dari segmentasi otomatis, gambar bawah hasil dari perbaikannya	42
Gambar IV-4 Label monofon dan durasinya untuk kalimat "hai selamat pagi, apa kabar"	43
Gambar IV-5 Tahapan implementasi proses segmentasi durasi	43
Gambar IV-6 Implementasi pemrosesan teks	46
Gambar IV-7 Tahapan pemisahan suku kata pada bahasa Indonesia	48

Gambar IV-8 Keluaran label dengan set konteks 1 untuk kalimat "hai selamat pagi, apa kabar"	51
Gambar IV-9 Keluaran label dengan set konteks 2 untuk kalimat "hai selamat pagi, apa kabar"	52
Gambar IV-10 Keluaran label dengan set konteks 3 untuk kalimat "hai selamat pagi, apa kabar"	54
Gambar IV-11 Tahapan implementasi proses ekstraksi fitur ucapan	56
Gambar IV-12 Label <i>full-context</i> dan durasinya untuk kalimat "hai selamat pagi, apa kabar" dengan set konteks 1	56
Gambar IV-13 Penggabungan dua fitur ucapan.....	57
Gambar IV-14 <i>Input-output</i> dari proses <i>training</i> model akustik HMM	57
Gambar IV-15 Tahapan dari proses <i>training</i> model pada HTS	58
Gambar IV-16 Tahapan proses sintesis ucapan dari teks.....	63
Gambar V-1 Hasil tes MOS dari evaluasi pertama dengan set konteks 1 terhadap tiga variasi koefisien <i>mel-cepstral</i>	66
Gambar V-2 Hasil tes MOS dari evaluasi kedua untuk ketiga konteks dengan <i>update question</i> untuk decision tree	67

DAFTAR TABEL

Tabel II-1 Daftar Fonem pada Bahasa Indonesia (Alwi, 2003).....	20
Tabel II-2 Daftar Suku Kata Baku pada Bahasa Indonesia (Alwi, 2003).....	22
Tabel III-1 Variasi daftar konteks yang digunakan.....	31
Tabel IV-1 Daftar kakas yang digunakan pada implementasi sistem	39
Tabel IV-2 Daftar fonem vokal pada bahasa Indonesia.....	46
Tabel IV-3 Daftar fonem konsonan pada bahasa Indonesia	47
Tabel IV-4 Daftar huruf beserta pasangan fonemnya pada bahasa Indonesia.....	49
Tabel IV-5 Set konteks 1 (Sakti, 2008) yang digunakan pada sistem TTS	50
Tabel IV-6 Set konteks 2 (Vania, 2011) yang digunakan pada sistem TTS.....	51
Tabel IV-7 Set konteks 3, kombinasi dari set pertama dan kedua.....	53
Tabel V-1 Hasil tes SUS dari evaluasi pertama dengan set konteks 1	68
Tabel V-2 Daftar kalimat yang salah tebak pada evaluasi SUS pertama.....	69
Tabel V-3 Hasil tes SUS dari evaluasi kedua untuk ketiga konteks dengan <i>update question</i> untuk decision tree.....	71
Tabel V-4 Daftar kalimat yang salah tebak pada evaluasi SUS kedua.....	72

BAB I.

PENDAHULUAN

I.1. Latar Belakang

Dalam masa ini, manusia sudah tidak asing lagi dengan komputer. Interaksi manusia dengan komputer terjadi hampir setiap hari bahkan setiap saat. Proses interaksi tersebut terjadi dengan adanya masukan dari manusia serta keluaran dari komputer. Keluaran dari komputer sendiri dimaksudkan agar manusia menerima hasil yang diinginkan dari proses yang dilakukan terhadap masukan yang diberikan manusia. Bentuk-bentuk dari keluaran komputer tersebut ada beberapa jenis, di antaranya adalah tulisan, gambar, suara, atau video.

Dari keempat macam jenis keluaran tersebut, suara, dalam hal ini secara khusus ucapan, memiliki karakteristik yang cukup detail, dengan membutuhkan *resource* yang tidak begitu banyak. Tulisan mungkin membutuhkan *resource* lebih sedikit dari ucapan, tetapi karakteristik dari ucapan sebagian besar lebih detail, dengan adanya penyertaan intonasi dan penekanan dalam pengucapan. Gambar pada umumnya memiliki karakteristik yang lebih kaya dari tulisan, tetapi tidak semua orang memiliki interpretasi yang sama terhadap gambar, terlebih lagi apabila gambar tersebut mengarah ke kategori abstraksi. Di lain hal, video memiliki banyak variasi, dengan gabungan suara dan gambar bergerak, tetapi suara atau ucapan memiliki kebutuhan sumber daya (*resource*) yang lebih sedikit.

Penggunaan ucapan sebagai representasi informasi juga bermanfaat bagi orang-orang yang memiliki keterbatasan visual. Hal ini khususnya dapat diimplementasikan dalam hal perubahan representasi teks menjadi ucapan. Di samping kelebihan ucapan yang dapat memberikan informasi mengenai emosi, seperti dari intonasi, dsb., bagi para penyandang keterbatasan visual, teknologi yang mengimplementasikan hal ini akan menghilangkan batasan dalam proses interaksi, baik dengan komputer maupun manusia.

Salah satu pendekatan dalam pembuatan sistem pensintesis ucapan, khususnya dari teks (*text-to-speech/TTS*), adalah dengan menggunakan teknik konkatenasi (Campbell, 1995). Implementasi teknik ini sudah pernah dilakukan terhadap bahasa Indonesia, khususnya dengan konkatenasi unit difon (Arman, 2001). Dengan teknik konkatenasi, sistem pensintesis ucapan dapat menghasilkan ucapan sintesis yang berkualitas tinggi. Namun, hal itu dibatasi oleh perlunya jumlah data yang sangat besar untuk dapat mencakup variasi ucapan yang ada. Permasalahan ini diatasi dengan menggunakan pendekatan yang terbaru saat ini, yaitu *Hidden Markov Model* (HMM). Melalui pendekatan ini, parameter ucapan dibangkitkan secara statistik melalui HMM yang dibentuk (Tokuda, 1995). Berbeda dengan teknik konkatenasi, dengan pendekatan HMM ini, sistem pensintesis ucapan dapat mengadaptasi berbagai variasi ucapan, seperti emosi, dialek, dsb. tanpa membutuhkan data yang sangat besar (Zen, 2009).

Implementasi HMM sudah dilakukan di berbagai bahasa, seperti Jepang (Yoshimura, 1999), Inggris (Tokuda, 2000), Portugal (Maia, 2006), Indonesia (Sakti, 2008; Vania, 2011), dll. Untuk bahasa Indonesia khususnya, kedua penelitian (Sakti, 2008; Vania, 2011) memiliki perbedaan pada variasi konteks pada pelabelan teks. Pada penelitian yang kedua memang terdapat adanya pengembangan, seperti konteks suku kata. Tetapi, beberapa konteks dari penelitian pertama, seperti jumlah fonem pada kata, posisi fonem pada kata, dsb. tidak diikutsertakan. Pada Tugas Akhir ini, mengacu kepada (Sakti, 2008; Vania, 2011) tersebut, dilakukan penggabungan variasi fitur dari keduanya. Penggabungan variasi fitur tersebut kemudian dievaluasi performansinya. Sesuai hasil dari evaluasi itu, akan diketahui kelayakan penggabungan kedua variasi fitur, tentunya terhadap sistem pensintesis ucapan berbasis HMM untuk bahasa Indonesia.

I.2. Rumusan Masalah

Yang menjadi rumusan masalah pada tugas akhir ini adalah bagaimana menjalankan proses untuk pembangunan sistem *text-to-speech* berbasis *Hidden Markov Model* (HMM) untuk bahasa Indonesia. Dengan melakukan penggabungan

implementasi terhadap dua penelitian yang terkait, maka diperlukan adanya evaluasi untuk mengukur performansi dari ucapan sintetis yang dihasilkan.

I.3. Tujuan

Berikut adalah tujuan dari pelaksanaan Tugas Akhir ini:

- Membangun sistem *text-to-speech* untuk bahasa Indonesia berbasis *Hidden Markov Model*, dengan mengimplementasikan aturan-aturan bahasa Indonesia untuk pemrosesan linguistik dan kakas-kakas yang tersedia untuk proses komputasi
- Melakukan evaluasi dari sistem dengan melakukan evaluasi secara subjektif melalui beberapa pendengar, terhadap kealamian dan akurasi dari ucapan sintetis yang dihasilkan. Kedua aspek, kealamian dan akurasi, dilakukan dengan dua metode evaluasi yang berbeda, yang disesuaikan dengan sifat dari kedua aspek penilaian tersebut.

I.4. Batasan Masalah

Batasan dari Tugas Akhir ini adalah domain bahasa dari sistem pensintesis ucapan yang dikembangkan hanya bahasa Indonesia yang baku, tanpa istilah asing, dan berupa kalimat pernyataan.

I.5. Metodologi

Pada Tugas Akhir ini, digunakan metodologi sebagai berikut:

a. Studi literatur

Studi literatur dilakukan terhadap berbagai materi dan penelitian yang telah dilakukan yang berkaitan dengan sistem *text-to-speech* berbasis *Hidden Markov Model*. Kajian juga dilakukan terhadap konsep dan aturan tata bahasa yang ada pada bahasa Indonesia. Studi literatur ini secara praktis dilaksanakan secara berkelanjutan sepanjang pengerjaan Tugas Akhir dari awal sampai akhir. Tahap ini dilakukan untuk menjadi dasar pedoman bagi proses pengerjaan secara keseluruhan.

b. Analisis solusi

Analisis dilakukan dengan mengaitkan teori-teori serta konsep yang dipelajari pada tahap studi literatur dengan tujuan yang ingin dicapai pada Tugas Akhir ini. Tahap ini dilakukan untuk menetapkan langkah-langkah yang perlu dikerjakan dalam membuat dan menyelesaikan sistem ini.

c. Desain sistem

Pada tahap ini, dilakukan perancangan sistem secara keseluruhan. Rancangan sistem dibuat berdasarkan analisis yang sudah dilakukan sebelumnya. Desain sistem yang dibuat pada tahap ini menjadi dasar pedoman untuk realisasi sistem.

d. Implementasi sistem

Dengan memiliki dasar rancangan sistem yang lengkap, tahap selanjutnya adalah implementasi sistem untuk merealisasikan sistem secara keseluruhan. Implementasi sistem ini diharapkan menghasilkan sistem pensintesis suara untuk bahasa Indonesia berbasis HMM sesuai dengan tujuan yang diinginkan.

e. Evaluasi sistem

Evaluasi terhadap hasil implementasi sistem dilakukan untuk menilai kualitas sintesis suara yang dihasilkan. Proses evaluasi dilakukan dengan dua metode yang berbeda, dengan masing-masing menilai untuk kealamian ucapan sintetis dan akurasi dari ucapan.

I.6. Sistematika Pembahasan

Penulisan Tugas Akhir ini dibagi ke dalam lima bab sebagai berikut. Bab pertama merupakan pendahuluan dari Tugas Akhir ini. Pendahuluan ini berisikan latar belakang dan rumusan masalah yang mendasari dibuatnya Tugas Akhir ini, serta tujuan, batasan masalah dan metodologi yang ditetapkan dalam pengerjaan Tugas Akhir.

Berbagai landasan teori yang mendukung pengerjaan Tugas Akhir ini dijabarkan pada bab kedua. Teori-teori ini terdiri dari teori mengenai *hidden Markov model* serta bagaimana pengaplikasiannya ke dalam *text-to-speech*, teori mengenai tata bahasa Indonesia, serta teknik-teknik evaluasi yang secara khusus digunakan untuk penelitian di bidang sintesis ucapan.

Bab tiga menjelaskan mengenai analisis dan solusi perancangan dari sistem *text-to-speech* ucapan bahasa Indonesia berbasis *hidden Markov model* (HMM). Perancangan itu terbagi ke dalam dua bagian utama, yaitu modul sistem penganalisis teks dan modul untuk *training* HMM serta sintesis ucapan. Bab ini ditutup dengan solusi untuk metode evaluasi yang digunakan.

Implementasi dari sistem yang sudah dirancang pada bab tiga tersebut dijabarkan prosesnya pada bab empat. Pada bab empat ini, dipaparkan langkah-langkah yang dilakukan dalam setiap proses yang mendasari pembangunan sistem pensintesis ucapan ini. Berbagai kondisi implementasi dan eksperimen juga diberikan dan dijabarkan untuk memberi gambaran lingkungan implementasi.

Sebagai penutup, dijabarkan pada bab lima, metode evaluasi sistem yang digunakan. Evaluasi dilakukan dengan tujuan untuk menilai kealamian, dan akurasi dari pengucapan suara sintetis yang dihasilkan sistem. Masing-masing aspek ini dinilai berdasar dua metode berbeda yang dipaparkan pada bab lima ini.

BAB II.

STUDI LITERATUR

Untuk dapat membangun sistem *text-to-speech* berbasis *Hidden Markov Model* (HMM) untuk bahasa Indonesia, perlu terlebih dahulu dilakukan pembelajaran terhadap teori-teori yang mendukung sistem ini. Pertama, perlu adanya pemahaman mengenai konsep matematis yang mendasari pendekatan HMM. Kemudian, proses mengenai penerapan konsep HMM tersebut ke dalam sistem TTS juga perlu dipahami, beserta kakas pendukung yang dapat digunakan dalam pengembangannya. Untuk dapat menerapkannya ke dalam bahasa Indonesia, perlu juga dipahami mengenai aturan tata bahasa Indonesia yang terkait di dalam pembuatan sistem TTS. Selain itu, metode untuk melakukan evaluasi terhadap sistem TTS yang sudah dibentuk juga perlu dipelajari. Terakhir, beberapa penelitian terkait, khususnya dalam bahasa Indonesia dijabarkan lebih lanjut.

II.1. *Hidden Markov Model* (Masuko, 2002)

Hidden Markov Model (HMM) adalah salah satu model rangkaian waktu (*time series*) statistikal yang digunakan pada berbagai bidang. Model ini terutama dipakai di bidang pengenalan suara (*speech recognition*), untuk mengenali rangkaian angka, huruf, kata atau kalimat dari suatu ucapan (Yamagishi, 2006). Selain sistem pengenalan suara, HMM juga dapat diimplementasikan ke dalam proses sintesis suara (*speech synthesis*), terutama pada sistem *text-to-speech synthesizer*. Pada subbab ini, dijelaskan konsep dasar dari *Hidden Markov Model*.

II.1.1. Definisi

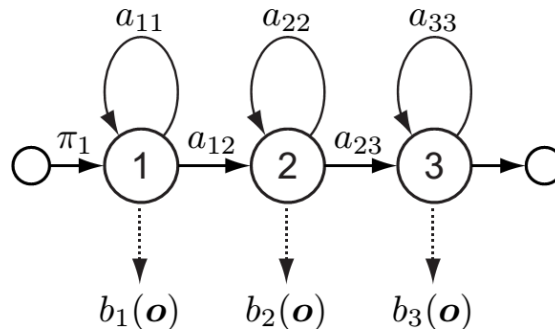
Hidden Markov Model (HMM) merupakan sebuah *finite state machine* yang membangkitkan sekuens *observation* dalam waktu diskrit. Pada setiap unit waktu, HMM mengalami perubahan *state*, yang merupakan proses Markov, sesuai dengan probabilitas *state transition* a_{ij} , di mana perpindahan akan terjadi dari *state* i ke *state*

j. Setiap *state* pada unit waktu ini akan membangkitkan data *observation* \mathbf{o} , sesuai dengan probabilitas *output* dari *state* tersebut (Yamagishi, 2006).

Sebuah HMM dengan yang terdiri dari N -*state* didefinisikan dengan probabilitas *state transition* $\mathbf{A} = \{a_{ij}\}_{i,j=1}^N$, distribusi probabilitas *output* $\mathbf{B} = \{b_i(\mathbf{o})\}_{i=1}^N$, dan probabilitas *initial state* $\mathbf{\Pi} = \{\pi_i\}_{i=1}^N$ (Masuko, 2002). Ketiga parameter tersebut dapat dinotasikan ke dalam bentuk berikut:

$$\lambda = (\mathbf{A}, \mathbf{B}, \mathbf{\Pi}). \quad (1)$$

Gambaran dari struktur HMM dapat dilihat pada gambar di bawah ini



Gambar II-1 Struktur HMM 3-*state left-to-right* (Masuko, 2002)

Struktur model seperti gambar di atas, menggambarkan HMM dengan indeks dari *state* yang bertambah atau sama seiring waktu bertambah. Pada umumnya, HMM *left-to-right* digunakan untuk memodelkan sekuens dari parameter suara. Ketiga parameter yang memodelkan HMM juga diilustrasikan pada gambar di atas, yaitu probabilitas *state transition* a_{ij} , probabilitas *output* $b_i(\mathbf{o})$, dan probabilitas *initial state* π_i .

Distribusi probabilitas *output* $b_j(\mathbf{o}_i)$ dapat bersifat diskrit maupun kontinu, tergantung dari observasi yang diinginkan. Untuk distribusi HMM yang kontinu, distribusi probabilitas *output* dimodelkan dengan *mixture* dari distribusi *multivariate Gaussian*, seperti di bawah ini,

$$b_j(\mathbf{o}) = \sum_{m=1}^M w_{jm} \mathcal{N}(\mathbf{o} | \boldsymbol{\mu}_{jm}, \mathbf{U}_{jm}), \quad (2)$$

dengan M adalah jumlah dari *mixture component*, sedangkan w_{jm} , $\boldsymbol{\mu}_{jm}$, \mathbf{U}_{jm} , masing-masing adalah *weight*, *mean vector*, dan *covariance matrix* dari *mixture component* m pada *state* j (Masuko, 2002). Distribusi *Gaussian* $\mathcal{N}(\boldsymbol{o}|\boldsymbol{\mu}_{jm}, \mathbf{U}_{jm})$ sendiri didefinisikan sebagai berikut

$$\mathcal{N}(\boldsymbol{o}|\boldsymbol{\mu}_{jm}, \mathbf{U}_{jm}) = \frac{1}{\sqrt{(2\pi)^d |\mathbf{U}|}} \exp\left(-\frac{1}{2}(\boldsymbol{o} - \boldsymbol{\mu}_{jm})^\top \mathbf{U}_{jm}^{-1}(\boldsymbol{o} - \boldsymbol{\mu}_{jm})\right) \quad (3)$$

dengan d adalah dimensi dari \boldsymbol{o} . Pada perumusan ini, *mixture component* adalah komponen yang berhubungan dengan data observasi yang bersangkutan. Misal, fundamental frekuensi dari suatu ucapan terdiri dari dua kategori nilai yang mungkin, nilai kontinu atau nilai diskrit (*white noise*), maka data observasinya terdiri dari 2 *mixture component*. *Weight* pada pengertian ini menandakan bobot dari suatu *mixture component* relatif terhadap komponen lainnya. *Mean vector* adalah kumpulan vektor dari data observasi yang memiliki nilai rata-rata dari keseluruhan data observasi yang ada. Sedangkan *covariance matrix* adalah matriks yang berisikan data *covariance* antar tiap vektor dari data observasi.

II.1.2. Kalkulasi Probabilitas *Observation*

Apabila sekuens dari *state* sudah didapat sebagai $\mathbf{Q} = (q_1, q_2, \dots, q_T)$, kemungkinan (*likelihood*) untuk membangkitkan sekuens *observation* $\mathbf{O} = (\boldsymbol{o}_1, \boldsymbol{o}_2, \dots, \boldsymbol{o}_T)$ dihitung dengan mengalikan probabilitas *state transition* dan probabilitas *output* untuk setiap *state*, sebagai berikut,

$$P(\mathbf{O}, \mathbf{Q}|\lambda) = \prod_{t=1}^T a_{q_{t-1}q_t} b_{q_t}(\boldsymbol{o}_t), \quad (4)$$

dengan $a_{q_{oj}}$ merupakan π_j . Total *likelihood* untuk membangkitkan \mathbf{O} dari HMM dengan parameter λ dihitung dengan melakukan marginalisasi dari \mathbf{O} terhadap sekuens *state* \mathbf{Q} , yaitu dengan menjumlahkan $P(\mathbf{O}, \mathbf{Q}|\lambda)$ terhadap semua kemungkinan sekuens *state*, seperti di bawah ini,

$$P(\mathbf{O}|\lambda) = \sum_{\text{all } \mathbf{Q}} \prod_{t=1}^T a_{q_{t-1}q_t} b_{q_t}(\boldsymbol{o}_t). \quad (5)$$

Likelihood dari persamaan (5) dihitung dengan menggunakan algoritma *forward-backward* (Masuko, 2002).

II.1.3. Kalkulasi Sekuens *State* Optimal

Selain penentuan kemungkinan (*likelihood*) dari pembangkitan data *observation* \mathbf{O} , penentuan satu sekuens *state* yang terbaik $\mathbf{Q}^* = (q_1^*, q_2^*, \dots, q_T^*)$, dengan diberikan sekuens *observation* $\mathbf{O} = (o_1, o_2, \dots, o_T)$, juga dapat digunakan untuk bermacam penerapan. Salah satu penerapan yang memanfaatkan ini adalah, sistem pengenalan suara (*speech recognition*). Pada umumnya, sistem pengenalan suara menggunakan *likelihood* dari sekuens *state* yang paling mungkin $P^* = P(\mathbf{O}, \mathbf{Q}^* | \lambda)$, dibanding *likelihood* total $P(\mathbf{O} | \lambda)$. Sekuens *state* terbaik \mathbf{Q}^* dapat diperoleh dengan menggunakan algoritma *Viterbi* (Masuko, 2002).

II.1.4. Estimasi Parameter Model

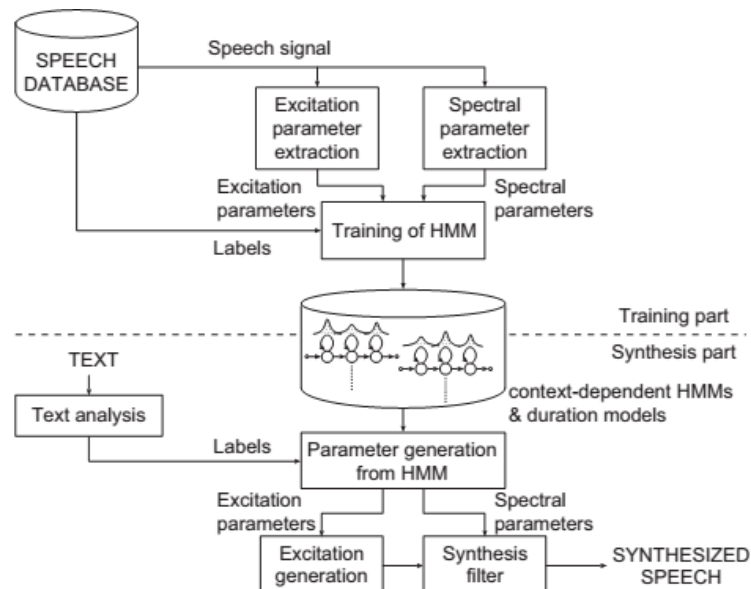
Dengan mengetahui perumusan untuk kalkulasi *likelihood* total dari sebuah sekuens *observation* \mathbf{O} , seperti pada persamaan (5), set parameter dari model HMM, yaitu λ , dapat dihitung dengan menggunakan metode *maximum likelihood* (ML). Metode ini akan mencari λ , yang memaksimalkan *likelihood* dari $P(\mathbf{O} | \lambda)$ pada persamaan (5), jika diberikan sebuah sekuens *observation* \mathbf{O} (Masuko, 2002).

Permasalahan ini merupakan permasalahan optimasi non-linear berdimensi tinggi (Masuko, 2002). Oleh karena itu, akan terdapat banyak angka optimum lokal (*local maxima*), sehingga sulit untuk mendapatkan λ yang secara global memaksimalkan $P(\mathbf{O} | \lambda)$. Walaupun demikian, dengan menggunakan perkiraan awal yang cukup baik terhadap λ , estimasi yang baik terhadap set parameter model tersebut juga dapat diperoleh secara lokal dengan algoritma *Expectation-Maximization* (EM) (yang juga sering dirujuk sebagai algoritma *Baum-Welch*) (Masuko, 2002).

II.2. Sistem *Text-to-Speech* berbasis *Hidden Markov Model*

Subbab ini menjabarkan gambaran dari sistem pensintesis ucapan dari teks (*text-to-speech synthesizer*/TTS) yang menerapkan konsep *Hidden Markov Model*

(HMM), beserta kakas-kakas pendukung yang biasanya digunakan pada pembuatannya. Skema dari sistem TTS berbasis HMM dapat dilihat pada gambar di bawah ini. Skema ini merujuk kepada kakas HTS (*HMM-based speech synthesis system*), yang dikembangkan di Nagoya Institute of Technology, Jepang (Zen, 2007), yang juga digunakan pada Tugas Akhir ini.



Gambar II-2 Diagram Sistem HMM-based *Text-to-Speech* (Zen, 2007)

Sistem *Text-to-Speech* (TTS) berbasis HMM memiliki tiga komponen utama pembangun sistem, yaitu komponen pemroses teks (berhubungan dengan Pemrosesan Bahasa Alami), komponen model akustik (berhubungan dengan Pembelajaran Mesin), dan komponen pemroses sinyal ucapan (berhubungan dengan Pemrosesan Sinyal). Pemroses teks berfungsi untuk memroses teks sesuai aturan bahasa yang digunakan dan melabelinya dengan fitur-fitur yang diperlukan, terutama untuk melihat konteks. Model akustik adalah model pembelajaran yang dibentuk dari hasil *training* terhadap data suara dan teks. Melalui model ini, dengan teks masukan tertentu, pola suara yang cocok akan dilihat dan dikeluarkan parameter suaranya. Pada sistem ini, model pembelajaran yang digunakan adalah *Hidden Markov Model*. Komponen pemroses gelombang suara berfungsi untuk menyusun parameter suara yang didapat dari model akustik ke dalam sinyal ucapan utuh yang dapat didengar dan dimengerti dengan baik. Komponen ini juga dapat

digunakan untuk memecah sinyal ucapan yang utuh ke dalam parameter-parameter pembentuknya untuk digunakan pada proses *training*.

Ketiga komponen utama di atas menyusun dua tahap dari proses TTS berbasis HMM. Tahap pertama, seperti sudah disebutkan beberapa kali sebelumnya, adalah tahap *training*, dan yang kedua adalah tahap *synthesis*. Tahap *training* adalah tahap pembentukan model akustik HMM. Pada tahap ini dilakukan pembelajaran terhadap pola data *training*, yang berupa data ucapan dan teksnya. Data ucapan diproses dengan komponen pemroses sinyal ucapan, untuk didapat parameter-parameter ucapannya. Data teks diproses dengan komponen pemroses teks, untuk didapat rangkaian fonem pembentuk ucapan, beserta konteks yang dimiliki setiap fonemnya. Hasil pemrosesan dari pemroses teks dipasangkan dengan parameter ucapannya yang sesuai. Kemudian, proses *training* dilakukan untuk mendapatkan model yang sesuai antara setiap karakteristik ucapan pada data *training* dengan label konteksnya. Tahap yang kedua, yaitu tahap *synthesis*, menggunakan model akustik yang sudah dibentuk pada tahap *training* untuk mengubah suatu teks masukan yang lain, ke dalam parameter-parameter ucapan yang sesuai, untuk kemudian parameter tersebut disusun menjadi gelombang ucapan yang dapat dimengerti. Berikut dijelaskan lebih lanjut mengenai komponen pemroses teks dan pemroses sinyal ucapan, proses *training* dan *synthesis* yang dilakukan, serta beberapa kakas pendukung yang biasanya digunakan dalam pembuatan sistem TTS berbasis HMM.

II.2.1. Pemroses teks

Komponen ini berfungsi untuk memroses teks ke dalam spesifikasi linguistik yang dapat digunakan sistem TTS berbasis HMM, baik untuk proses pembelajaran (*training*) maupun sintesis (*synthesis*). Untuk mewakili banyaknya variasi pengucapan yang mungkin, spesifikasi linguistik ini harus terdiri dari konteks yang dapat merepresentasikan setiap variasi bunyi tersebut. Setiap spesifikasi linguistik yang ditentukan nantinya dimuat ke dalam label konteks pada masukan dari sistem TTS. Spesifikasi linguistik yang dibutuhkan pada sistem TTS ini biasanya terdiri dari dua kategori konteks (Yamagishi, 2008), yaitu:

a. Konteks fonetik

Konteks fonetik adalah rangkaian fonem yang mengelilingi suatu fonem pada saat pengucapan fonem tersebut. Fonem pada dasarnya merupakan satuan bunyi terkecil yang dapat membedakan arti. Setiap kalimat yang diucapkan terdiri dari susunan kata. Setiap kata tersebut terdiri dari susunan fonem. Perbedaan rangkaian fonem yang mengelilingi suatu fonem pada saat pengucapan dapat mempengaruhi pengucapan fonem tersebut. Oleh karena itu, dibutuhkan pengetahuan akan fonem-fonem yang mengelilingi suatu fonem pada saat pengucapan.

b. Konteks prosodi

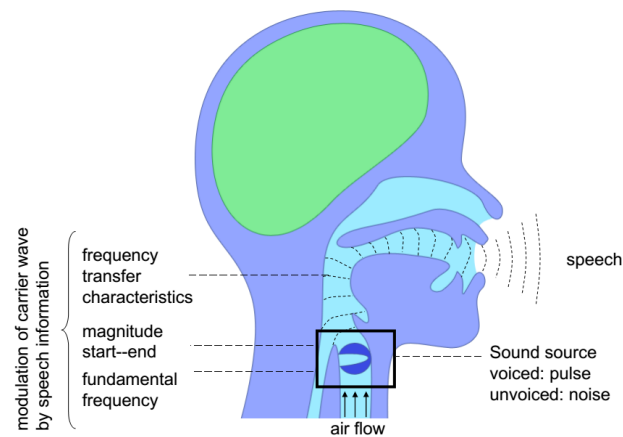
Konteks prosodi adalah konteks yang terdiri dari berbagai cara pengucapan, seperti penekanan atau intonasi, berbagai kondisi pemosisian fonem, suku kata, maupun kata pada saat pengucapan, atau kondisi-kondisi lainnya yang dapat menjadi variasi bagi bunyi dari suatu fonem. Khusus untuk penelitian dalam bahasa Indonesia, (Sakti, 2008) menggunakan konteks pemosisian fonem pada kata, jumlah fonem pada kata, pemosisian kata pada kalimat, serta jumlah kata pada kalimat. Sedangkan (Vania, 2011), lebih menekankan penggunaan konteks suku kata, seperti pemosisian fonem pada suku kata, penekanan pada suku kata, pemosisian suku kata pada kalimat, dsb.

II.2.2. Pemroses sinyal ucapan

Komponen pemroses sinyal ucapan adalah komponen yang tentunya berkaitan dengan pemrosesan sinyal. Secara umum, komponen ini memiliki dua peran penting. Pertama, melakukan analisis terhadap sinyal ucapan, lalu membangkitkan parameter eksitasi (*excitation parameter*), parameter spektrum (*spectral parameter*) dan parameter dari durasi untuk setiap segmen pengucapan. Ketiga parameter ini digunakan oleh proses *training* pada sistem *HMM-based speech synthesis*. Kedua, melakukan sintesis (*synthesis*) terhadap parameter eksitasi,

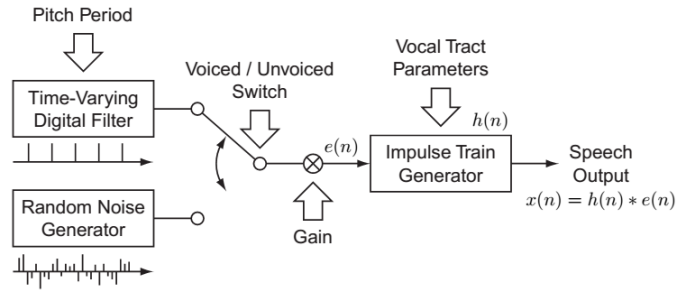
parameter spektrum dan durasi dari setiap segmen untuk menghasilkan gelombang ucapan yang dapat dimengerti.

Parameter eksitasi dapat berupa frekuensi fundamental (*fundamental frequency*), yang dihasilkan apabila pita suara manusia bergetar saat berbicara, atau dapat berupa *white noise*, yang dihasilkan apabila pita suara manusia tidak bergetar. Parameter spektrum merupakan parameter yang mendefinisikan karakteristik dari gelombang suara. Sumber bunyi yang dihasilkan oleh pita suara manusia akan disalurkan melalui rongga suara (*vocal tract*), dan mengalami modulasi, sehingga terjadi perubahan karakteristik sesuai keadaan rongga suara saat itu. Setiap bunyi suara yang dihasilkan manusia memiliki karakteristik suara atau parameter spektrum yang berbeda-beda. Parameter spektrum ini dimodelkan dalam bentuk *mel-cepstral frequency* (Masuko, 2002). Di bawah ini digambarkan proses pembentukan ucapan pada manusia.



Gambar II-3 Proses Pembentukan Ucapan pada Manusia (Zen, 2010)

Proses pembentukan ucapan pada manusia tersebut dapat diimitasi dengan proses komputasi. Berikut adalah skema proses komputasi pembentukan sinyal ucapan.



Gambar II-4 Model Komputasi Pembentukan Sinyal Ucapan (Masuko, 2002)

Dari skema di atas sinyal suara $x(n)$ dikalkulasi dengan perumusan *convolution* terhadap nilai eksitasi, parameter eksitasi, $e(n)$ dan respons impuls, parameter spektrum, $h(n)$, sebagai berikut (Masuko, 2002). Parameter durasi sendiri menentukan lamanya kelompok sinyal yang menentukan suatu satuan bunyi dihasilkan.

(6)

$$x(n) = h(n) * e(n)$$

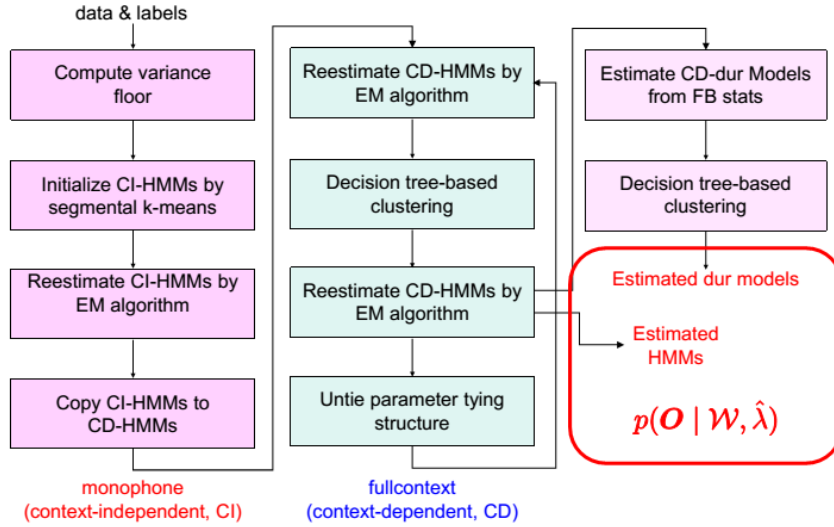
II.2.3. Proses *training* dan *synthesis*

Pada subbab ini dijelaskan mengenai tahapan dalam pembentukan model HMM (*training*) dan proses sintesis ucapan berdasarkan model yang dibentuk (*synthesis*). Pada proses *training*, dengan menggunakan konteks label yang dihasilkan dari pemroses teks serta data parameter ucapan yang dihasilkan dari pemroses gelombang ucapan, *Hidden Markov Model* (HMM) akan dibentuk dengan menggunakan perumusan seperti di bawah ini (Zen, 2007),

(7)

$$\hat{\lambda} = \arg \max_{\lambda} \{p(\mathbf{O} | \mathcal{W}, \lambda)\},$$

di mana λ adalah set parameter dari model HMM, dijelaskan pada subbab II.2.1, \mathbf{O} adalah kumpulan data *training* ucapan, dan \mathcal{W} adalah kumpulan transkripsi (sekuens kata beserta konteks label) yang sesuai dengan \mathbf{O} . Set parameter λ didapat dengan menggunakan metode *maximum likelihood* (ML), yang memaksimalkan *likelihood* dari kemunculan \mathbf{O} , dijelaskan pada subbab II.2.1. Parameter terbaik yang didapat itulah yang membentuk HMM untuk model akustik yang sesuai dengan data *training*. Untuk proses *training* ini sendiri, langkah-langkah yang dilakukan pada kakas HTS dapat dilihat di bawah ini.



Gambar II-5 Proses Training pada HTS (Zen, 2010)

Setelah model didapat dari proses *training*, proses selanjutnya adalah proses *synthesis*. Pada proses ini, dengan menggunakan estimasi model $\hat{\lambda}$ tersebut, parameter suara \mathbf{o} dapat dibangkitkan, jika diberikan sekuens kata w , seperti dirumuskan di bawah ini (Zen, 2007).

$$\hat{\mathbf{o}} = \arg \max_{\mathbf{o}} \left\{ p(\mathbf{o} | w, \hat{\lambda}) \right\}. \quad (8)$$

Proses *synthesis* secara matematis melakukan maksimalisasi terhadap persamaan (2). Pertama-tama, sekuens kata yang diberikan diubah ke dalam sekuens label yang konteks-dependen dengan menggunakan komponen pemroses teks, kemudian sekuens HMM yang lengkap dibentuk dengan melakukan konkatenasi terhadap beberapa konteks-dependen HMM yang sesuai dengan sekuens label tersebut. Kedua, algoritma pembangkitan parameter ucapan (*speech parameter generation*) akan membangkitkan sekuens parameter ucapan, yang terdiri oleh parameter spektrum (*spectral parameter*) dan parameter eksitasi (*excitation parameter*) dari sekuens HMM yang dibentuk pada tahap pertama tadi.

Asumsikan distribusi *output* dari setiap *state* HMM adalah *single stream* dan *single multi-variate Gaussian distribution*, sebagai berikut (Zen, 2007)

$$b_j(\mathbf{o}_t) = \mathcal{N}(\mathbf{o}_t ; \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j), \quad (9)$$

di mana \mathbf{o}_t adalah vektor *state-output* pada *frame* t , dan $b_j(\cdot)$, $\boldsymbol{\mu}_j$, serta $\boldsymbol{\Sigma}_j$ masing-masing merupakan distribusi *state-output* ke- j , vektor *mean*, dan *covariance matrix*. Di dalam sistem HMM-based *speech synthesis*, persamaan (7) dapat di-estimasi sebagai berikut (Zen, 2007)

$$\hat{\mathbf{o}} = \arg \max_{\mathbf{o}} \left\{ p(\mathbf{o} \mid w, \hat{\boldsymbol{\lambda}}) \right\} \quad (10)$$

$$= \arg \max_{\mathbf{o}} \left\{ \sum_{\mathbf{q}} p(\mathbf{o}, \mathbf{q} \mid w, \hat{\boldsymbol{\lambda}}) \right\} \quad (11)$$

$$\approx \arg \max_{\mathbf{o}} \max_{\mathbf{q}} \left\{ p(\mathbf{o}, \mathbf{q} \mid w, \hat{\boldsymbol{\lambda}}) \right\} \quad (12)$$

$$= \arg \max_{\mathbf{o}} \max_{\mathbf{q}} \left\{ P(\mathbf{q} \mid w, \hat{\boldsymbol{\lambda}}) \cdot p(\mathbf{o} \mid \mathbf{q}, \hat{\boldsymbol{\lambda}}) \right\} \quad (13)$$

$$\approx \arg \max_{\mathbf{o}} \left\{ p(\mathbf{o} \mid \hat{\mathbf{q}}, \hat{\boldsymbol{\lambda}}) \right\} \quad (14)$$

$$= \arg \max_{\mathbf{o}} \left\{ \mathcal{N}(\mathbf{o}; \boldsymbol{\mu}_{\hat{\mathbf{q}}}, \boldsymbol{\Sigma}_{\hat{\mathbf{q}}}) \right\}, \quad (15)$$

di mana $\mathbf{o} = [\mathbf{o}_1^\top, \dots, \mathbf{o}_T^\top]^\top$ adalah vektor *state-output* yang akan dibangkitkan, $\mathbf{q} = \{q_1, \dots, q_T\}$ adalah sekuens *state*, dan $\boldsymbol{\mu}_{\mathbf{q}} = [\boldsymbol{\mu}_{q_1}^\top, \dots, \boldsymbol{\mu}_{q_T}^\top]^\top$ adalah vektor *mean* untuk \mathbf{q} . Pada perumusan ini, $\boldsymbol{\Sigma}_{\mathbf{q}} = \text{diag}[\boldsymbol{\Sigma}_{q_1}, \dots, \boldsymbol{\Sigma}_{q_T}]$ adalah *covariance matrix* untuk \mathbf{q} dan T adalah jumlah total *frame* pada \mathbf{o} . Sekuens *state*, $\hat{\mathbf{q}}$, ditentukan untuk memaksimalkan probabilitas dari *state-duration*-nya (Zen, 2007), yaitu

$$\hat{\mathbf{q}} = \arg \max_{\mathbf{q}} \left\{ P(\mathbf{q} \mid w, \hat{\boldsymbol{\lambda}}) \right\}. \quad (16)$$

Dengan menggunakan hanya perumusan di atas, estimasi parameter suara $\hat{\mathbf{o}}$ akan bersifat *piece-wise stationary*, di mana pada setiap segmen waktu pada tiap *state* hanya merujuk kepada vektor *mean* dari *state* tersebut (Zen, 2007). Hal ini tentunya akan menghasilkan ucapan yang berkualitas kurang baik karena pada kenyataannya parameter suara memiliki variasi yang lebih halus. Untuk menghasilkan rangkaian parameter suara yang lebih realistis, algoritma pembangkitan parameter suara ini juga menggunakan keterkaitan antara fitur statis dan dinamis (*dynamic and static features*) dalam perumusan maksimalisasi probabilitas (Zen, 2007). Apabila vektor *state-output*, \mathbf{o}_t , mengandung fitur statis berdimensi M , \mathbf{c}_t , dan fitur dinamis pada *first-order (delta)*, $\Delta \mathbf{c}_t$, sebagai berikut,

$$\mathbf{o}_t = [\mathbf{c}_t^\top, \Delta \mathbf{c}_t^\top]^\top, \quad (17)$$

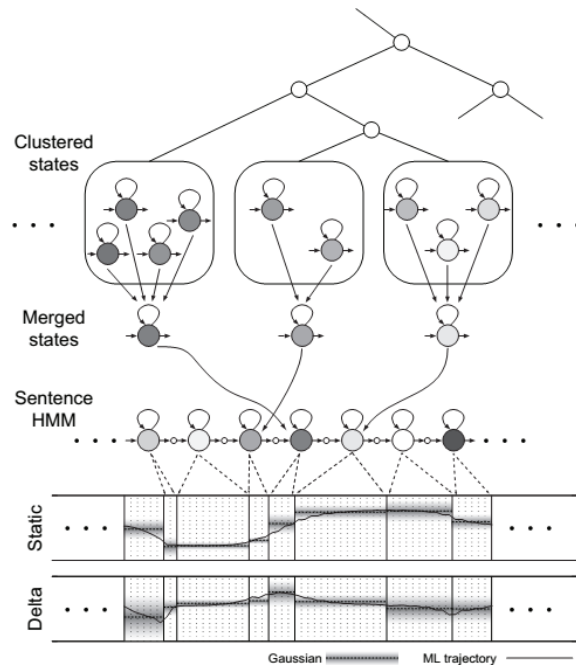
dan fitur dinamisnya dikalkulasi sebagai,

$$\Delta \mathbf{c}_t = \mathbf{c}_t - \mathbf{c}_{t-1}, \quad (18)$$

keterhubungan antara \mathbf{o}_t dan \mathbf{c}_t dapat disusun dalam matriks sebagai berikut (Zen, 2007),

$$\begin{bmatrix} \vdots \\ \mathbf{c}_{t-1} \\ \Delta \mathbf{c}_{t-1} \\ \mathbf{c}_t \\ \Delta \mathbf{c}_t \\ \mathbf{c}_{t+1} \\ \Delta \mathbf{c}_{t+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \dots & \vdots & \vdots & \vdots & \vdots & \dots \\ \dots & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots \\ \dots & -\mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots \\ \dots & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \dots \\ \dots & \mathbf{0} & -\mathbf{I} & \mathbf{I} & \mathbf{0} & \dots \\ \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \dots \\ \dots & \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{I} & \dots \\ \dots & \vdots & \vdots & \vdots & \vdots & \dots \end{bmatrix} \begin{bmatrix} \vdots \\ \mathbf{c}_{t-2} \\ \mathbf{c}_{t-1} \\ \mathbf{c}_t \\ \mathbf{c}_{t+1} \\ \vdots \end{bmatrix} \quad (19)$$

Dari perumusan di atas, dapat digambarkan proses *synthesis*, dengan menggunakan model akustik yang telah terbentuk, pada sistem HMM-based *speech synthesis* sebagai berikut,



Gambar II-6 Skema Proses *Synthesis* pada Sistem HMM-based *Speech Synthesis* (Zen, 2007)

II.2.4. Kakas pendukung sistem TTS berbasis HMM

Untuk dapat mengembangkan sistem *text-to-speech* (TTS) berbasis HMM, penggunaan berbagai kakas pendukung yang mengimplementasikan setiap komponen penyusun sistem sangatlah diperlukan. Kakas-kakas pendukung sistem ini dapat dikategorikan ke dalam beberapa komponen yang menyusun sistem ini sendiri, yaitu:

a. Kakas pemroses teks

Pada dasarnya, pengembang sistem TTS berbasis HMM dapat menggunakan komponen pemroses teks buatan sendiri ataupun menggunakan kakas sudah tersedia untuk bahasa yang diinginkan. Salah satu kakas pemrosesan teks yang dapat secara generik digunakan untuk pengembangan sistem TTS adalah Festival (Zen, 2007). Walaupun secara *default* kakas ini diperuntukkan untuk bahasa Inggris, dengan sedikit penyesuaian terhadap tata bahasa Indonesia, kakas ini dapat digunakan sebagai pemrosesan teks pada sistem TTS berbasis HMM untuk bahasa Indonesia.

b. Kakas pemroses sinyal

Salah satu metode pemrosesan sinyal ucapan yang banyak digunakan pada sistem TTS berbasis HMM adalah STRAIGHT (Kawahara, 1999). Selain itu, terdapat juga kakas SPTK (*Signal Processing Toolkit*) yang dapat digunakan untuk melakukan pemrosesan sinyal (Tokuda, 2011). Kakas *hts_engine* (Tokuda, 2011) juga biasanya digunakan dalam proses sintesis sinyal untuk sistem TTS berbasis HMM.

c. Kakas pembuat model HMM

HTS (*HMM-based speech synthesis system*) adalah sebuah kakas yang dikembangkan secara khusus untuk mendukung pembuatan model akustik pada sistem *text-to-speech* (TTS) yang berbasis HMM (Zen, 2007). Kakas HTS ini dikembangkan sebagai ekstensi dari kakas HTK (*Hidden Markov*

Model toolkit), yang biasanya digunakan untuk pengenalan ucapan (*speech recognition*) (Young, 2009)

d. Kakas tambahan

Untuk dapat memenuhi semua fitur yang dibutuhkan, diperlukan juga cara yang efektif dalam menentukan parameter durasi dari setiap segmen pengucapan. Durasi ini menentukan lamanya setiap kelompok sinyal dari satuan bunyi terkecil yang ada. Salah satu tekniknya adalah dengan menggunakan segmentasi otomatis menggunakan *Hidden Markov Model* (HMM) juga, yang merupakan bagian dari proses pengenalan ucapan (*speech recognition*) (Masuko, 2002). Untuk dapat menjalankan proses tersebut, salah satu kakas yang dapat digunakan adalah HTK (*Hidden Markov Model toolkit*) (Young, 2009).

II.3. Tata Bahasa Indonesia

Indonesia merupakan negara dengan populasi terbesar keempat di dunia. Bahasa Indonesia adalah bahasa nasional dari negara Indonesia. Oleh karenanya, bahasa Indonesia termasuk ke dalam kategori bahasa yang diucapkan oleh banyak orang. Indonesia digunakan sebagai bahasa utama bagi sekitar 23 juta orang, dan sebagai bahasa kedua bagi kurang lebih 140 juta orang (Alwi, 2003). Bagi penduduk negara Indonesia yang terletak di daerah perkotaan, bahasa Indonesia menjadi sarana komunikasi lisan yang utama. Di lain pihak, bagi penduduk yang terletak di daerah pedesaan, bahasa Indonesia yang baku tidak digunakan sesering bahasa asli di daerah masing-masing.

Sebagai bahasa yang dituturkan oleh banyak orang, bahasa Indonesia memiliki aturan-aturan yang menjadi pedoman bagi bahasa Indonesia yang baku. Pedoman ini diperlukan untuk menyetarakan pemahaman penuturan antar berbagai kalangan penutur. Aturan tata bahasa ini juga sangat dibutuhkan untuk pembuatan sistem yang secara langsung memroses penuturan dalam bahasa Indonesia. Salah satu sistem yang membutuhkan aturan tata bahasa Indonesia adalah sistem

pensintesis ucapan dari teks (*text-to-speech system*/TTS) yang dikembangkan pada Tugas Akhir ini.

Sistem TTS merupakan sistem komputasi yang didasari oleh pemrograman dengan bahasa komputer. Dalam hal ini berarti komputer yang bersangkutan tidak dapat mengerti kalimat atau ucapan dalam bahasa Indonesia tanpa adanya aturan yang diberikan. Oleh karena itu, untuk mengatasi keterbatasan ini, diperlukan penggunaan aturan tata bahasa Indonesia yang terkait dengan proses sintesis ucapan. Dalam proses sintesis ucapan, secara garis besar proses dilakukan dengan mengubah teks masukan menjadi keluaran ucapan. Untuk menjalankan proses ini, komputer perlu memahami teks ke dalam satuan kata, suku kata, sampai ke fonem. Pada subbab berikutnya, dijabarkan daftar fonem pada bahasa Indonesia, aturan pemenggalan suku kata yang baku, dan batasan-batasan dalam penentuan urutan fonem pada suatu kata.

II.3.1. Daftar Fonem Bahasa Indonesia

Fonem merupakan satuan bunyi terkecil pada suatu bahasa yang dapat membedakan arti. Bahasa Indonesia sendiri memiliki total 34 fonem, yang terdiri dari 9 fonem vokal dan 25 fonem konsonan. Berbeda dengan huruf vokal dan konsonan pada bahasa Indonesia, yang masing-masing berjumlah 5 dan 21, jumlah fonem vokal dan konsonan ini sedikit lebih banyak dengan adanya beberapa diftong (gabungan dua huruf vokal) dan gugus konsonan (gabungan dua huruf konsonan) (Alwi, 2003). Tabel di bawah ini menjabarkan daftar fonem yang ada pada bahasa Indonesia.

Jenis Fonem	Huruf Biasa	Diftong	Gugus konsonan	Total
Vokal	a,i,u,e,ə,o	ay,aw,oy	-	9
Konsonan	b,c,d,f,g,h,j,k,l,m,n,p,q,r,s,t,v,w,x,y,z	-	kh,ng,ny,sy	25

Tabel II-1 Daftar Fonem pada Bahasa Indonesia (Alwi, 2003)

Seperti yang sudah disebutkan sebelumnya, dari tabel di atas dapat dilihat bahwa masing-masing fonem vokal dan konsonan memiliki jumlah 9 dan 25. Fonem konsonan sendiri terdiri dari 21 huruf konsonan ditambah empat gugus konsonan, yaitu kh, ng, ny, dan sy. Sedangkan fonem vokal terdiri dari 3 diftong, yaitu /ay/, /aw/, dan /oy/, serta 6 bunyi huruf vokal, dengan huruf e memiliki dua variasi bunyi, yaitu /e/, seperti pada “lebar” dan /ə/, seperti pada “selamat”. Dari perincian daftar fonem tersebut, dapat dilihat bahwa penggunaan fonem pada bahasa Indonesia terbilang sederhana, dengan sebagian besar mengikuti pelafalan hurufnya masing-masing. Hanya, pengecualian biasanya dilakukan terhadap huruf “e” yang memiliki dua variasi bunyi, serta beberapa fonem dengan huruf ganda, seperti diftong atau gugus konsonan.

II.3.2. Aturan Pemenggalan Suku Kata

Suku kata adalah struktur yang terdiri dari urutan fonem, yang merupakan bagian penyusun suatu kata. Setiap suku kata pasti dan hanya boleh memiliki satu fonem vokal (termasuk diftong). Tabel di bawah ini menjabarkan pola suku kata baku yang ada pada bahasa Indonesia, beserta contoh penggunaannya (Alwi, 2003).

No.	Suku Kata	Contoh
1.	V	<i>a-nak, i-rit</i>
2.	VK	<i>an-da, um-pan</i>
3.	KV	<i>se-bab, man-di</i>
4.	KVK	<i>lan-tai, ma-kan</i>
5.	KKV	<i>sas-tra, pri-ma</i>
6.	KKVK	<i>frik-si, kon-trak</i>
7.	VKK	<i>ons, eks</i>

8.	KVKK	<i>pers, kon-teks</i>
9.	KKVKK	<i>kom-pleks</i>
10.	KKKV	<i>in-stru-men</i>
11.	KKKVK	<i>struk-tur</i>

Tabel II-2 Daftar Suku Kata Baku pada Bahasa Indonesia (Alwi, 2003)

Untuk memroses suatu kata ke dalam suku-suku kata pembentuknya, terdapat pula aturan langkah-langkah yang ditetapkan (Basuki, 2000). Berikut adalah langkah-langkah pemisahan kata menjadi suku-suku katanya:

1. Apabila di tengah kata, terdapat dua huruf vokal (selain diftong) yang berurutan, pisahkan kedua huruf vokal tersebut.
2. Apabila di tengah kata, terdapat huruf konsonan yang tepat diapit dua huruf vokal, pemisahan dilakukan di antara huruf vokal pertama dan konsonan tersebut.
3. Apabila di tengah kata, terdapat dua atau lebih huruf konsonan yang berurutan, lakukan pemisahan setelah huruf konsonan yang pertama.

II.3.3. Hal-hal yang Perlu Diperhatikan pada Teks Bahasa Indonesia

Seperti halnya bahasa-bahasa yang lain (Black, 2003), bahasa Indonesia pun memiliki beberapa hal yang perlu diperhatikan saat membangun sistem pemroses teks. Beberapa hal tersebut adalah:

- Angka

Penulisan angka dapat bervariasi tujuannya. Angka dapat dituliskan sebagai angka biasa, nomor telepon, rentang nilai, angka romawi, atau jumlah uang.
- Singkatan

Singkatan dapat berupa dua macam. Pertama dapat berupa akronim, yaitu singkatan yang dapat dibaca seperti layaknya kata biasa, misal ABRI. Jenis

yang satu lagi adalah singkatan yang tidak bisa dibaca seperti kata biasa, melainkan harus dibaca per huruf, misal TNI.

- Tanggal dan waktu

Penulisan angka dan huruf yang merujuk kepada tanggal maupun waktu juga harus diperhatikan. Konteks kata dan kalimat yang bersangkutan menentukan pemrosesan pada hal ini.

- Karakter khusus

Berbagai karakter-karakter selain alfabet dan angka dapat diikutsertakan ke dalam teks bahasa Indonesia. Salah satu contohnya adalah operator aritmatika, seperti “+” (tambah) atau “/” (bagi) yang dapat menimbulkan keambiguan. Karakter seperti titik juga perlu diperhatikan. Titik tidak hanya digunakan untuk mengakhiri kalimat, tetapi juga dapat digunakan untuk menuliskan singkatan, seperti gelar, misal Dr. (Doktor), atau Rp. (Rupiah).

II.4. Evaluasi Sistem

Evaluasi sistem pensintesis ucapan tentunya dilakukan dengan melakukan penilaian kualitas dari suara ucapan hasil sintesis. Penilaian kualitas ucapan ini biasanya dilakukan dengan dua metode, yang pertama adalah evaluasi objektif, dan kedua adalah evaluasi subjektif. Berikut dijelaskan mengenai kedua metode tersebut:

- Evaluasi Objektif

Evaluasi objektif dilakukan dengan mengukur perbedaan jarak antara nilai parameter *cepstral* pada ucapan sintetis dan ucapan asli. Parameter *cepstral* adalah parameter yang menentukan karakteristik frekuensi pada rongga suara. Apabila jarak antara nilai di ucapan sintetis dan ucapan asli semakin dekat, maka ucapan hasil keluaran sistem tersebut dikatakan semakin baik kualitasnya (Meyer, 2006).

- Evaluasi Subjektif

Evaluasi subjektif merupakan evaluasi yang dilakukan untuk menilai kualitas suara berdasarkan pendengaran manusia. Penilaian subjektivitas ini

sendiri dilakukan untuk menentukan dua hal, yaitu kealamian dari ucapan dan kemiripan ucapan buatan dengan yang asli. Untuk menilai kealamian dari ucapan, digunakan uji *Mean Opinion Score* (MOS) (Yamagishi, 2008). Pada uji MOS ini, pendengar diberi tugas untuk menilai kealamian ucapan yang didengar. Interval nilai yang diberikan adalah dari 1-5, dengan 1 merupakan sangat tidak alami, dan 5 adalah sangat alami. Selain uji MOS, biasanya dijalankan juga uji *Comparison Category Rating* (CCR) untuk menilai kemiripan ucapan buatan dengan yang asli. Beberapa pendengar yang sudah ditentukan tadi akan dihadapkan kepada kumpulan pasangan ucapan buatan dan asli, dan menilai kemiripannya. Interval nilai yang ditentukan juga 1-5, dengan 1 merupakan sangat tidak mirip, dan 5 adalah sangat mirip. Jenis evaluasi yang terakhir adalah *Semantically Unpredictable Sentence* (SUS) *test* yang dilakukan untuk menilai akurasi pengucapan fonem/kata dalam kalimat berdasar persepsi pendengar akan kata yang diucapkan (Sakti, 2010).

II.5. Penelitian terkait sistem TTS untuk bahasa Indonesia

Pengembangan sistem *text-to-speech* (TTS) pertama untuk bahasa Indonesia dilakukan dengan menggunakan pendekatan konkatenasi difon (Arman, 2001). Melalui pendekatan ini, unit-unit ucapan disimpan di dalam basis data, untuk kemudian diambil dan digabung-gabungkan sesuai teks masukannya. Kualitas ucapan yang dihasilkan dengan menggunakan metode ini memang sangat baik, khususnya di level sinyalnya. Namun, untuk mendapatkan variasi dari karakteristik ucapan dibutuhkan jumlah data yang sangat besar. Ditambah pula terdapatnya diskontinuitas pada ucapan, apabila ucapan yang diinginkan tidak terdapat pada basis data (Tokuda, 2000).

Untuk mengatasi kelemahan dari pendekatan berbasis konkatenasi tersebut, pendekatan terbaru dengan menggunakan *Hidden Markov Model* (HMM) dikembangkan (Yoshimura, 1999). Kualitas yang dihasilkan dari pendekatan ini memang tidak setinggi pada pendekatan konkatenasi. Namun, sinyal yang dihasilkan lebih stabil dan halus karena parameter yang digunakan didapat secara

statistikal. Selain itu, untuk menambahkan berbagai variasi karakteristik dari ucapan dapat dilakukan dengan lebih mudah tanpa memerlukan data yang sangat besar (Tokuda, 2000).

Sistem TTS yang menggunakan pendekatan HMM sudah pernah dikembangkan sebelumnya (Sakti, 2008; Vania, 2011). Sakti (2008) memulai penelitian sistem TTS berbasis HMM untuk bahasa Indonesia yang pertama kali. Penelitian ini kemudian dikembangkan oleh (Vania, 2011). Secara garis besar, perbedaan antara keduanya terletak pada penggunaan konteks untuk pelabelan teks. Di bawah ini adalah cuplikan daftar konteks yang digunakan pada masing-masing dari kedua penelitian tersebut.

<ul style="list-style-type: none"> • Phoneme level: <ul style="list-style-type: none"> - {second preceding, preceding, current, succeeding, second succeeding} phoneme; - position of current phoneme in the current word (forward and backward); • Word level: <ul style="list-style-type: none"> - number of phonemes in {preceding, current, succeeding} word; - position of current word in the current utterance (forward and backward); • Utterance level: <ul style="list-style-type: none"> - number of words in the utterance; - utterance types: declarative, interrogative or imperative sentence.

Gambar II-7 Daftar konteks pada penelitian (Sakti, 2008)

Phoneme: - {second preceding, preceding, current, succeeding, second succeeding} phoneme - Position of current phoneme in the current syllable
Syllable: - Number of phonemes at {preceding, current, succeeding} syllable - Stress of {preceding, current, succeeding} syllable - Position of current syllable in current word - Number of syllables {from previous to next} stressed syllable - Vowel within current syllable
Word: - Number of syllables in {preceding, current, succeeding} word - Position of current word in the current sentences
Utterance: - Number of syllable in the utterance

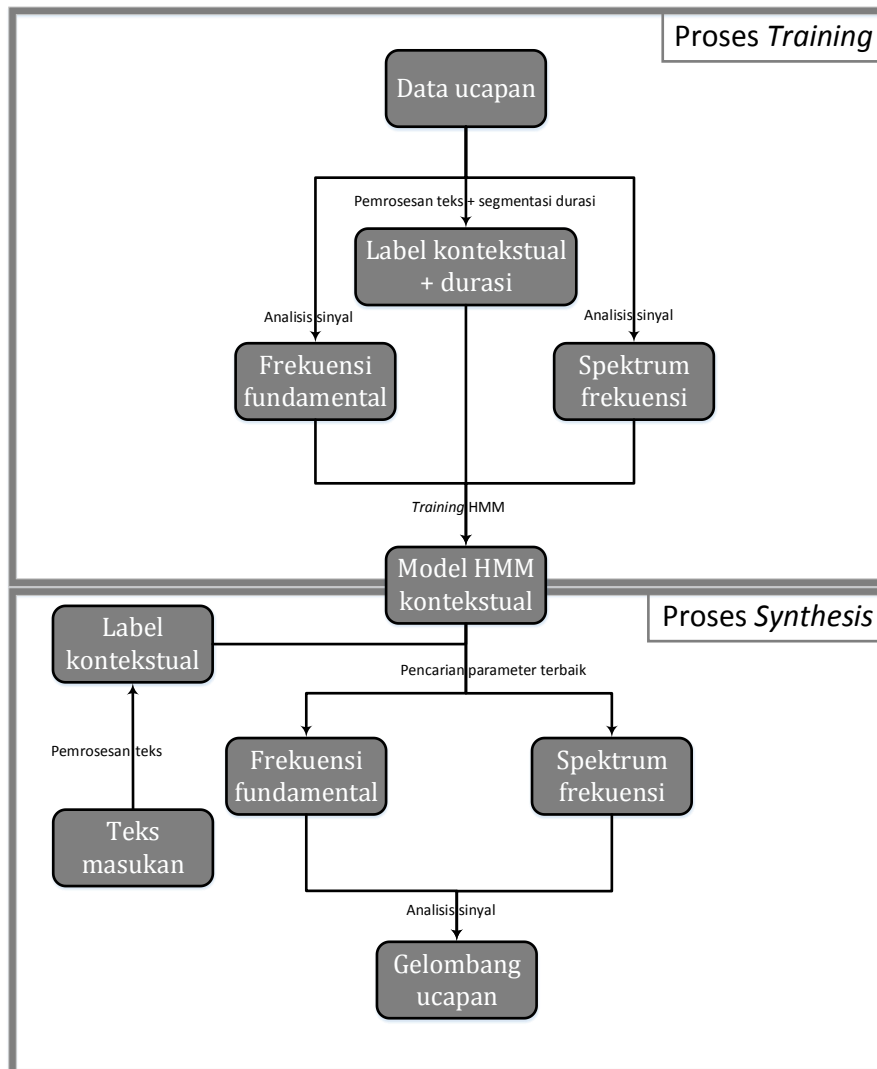
Gambar II-8 Daftar konteks pada penelitian (Vania, 2011)

Sakti (2008) menggunakan konteks yang terdiri dari tingkat fonem, kata dan kalimat. Vania (2011) menambahkan penggunaan tingkat suku kata pada pelabelan konteksnya. Menurut (Vania, 2011) penambahan konteks suku kata, terutama penekanan suku kata, memberikan hasil yang lebih baik. Tetapi, dapat dilihat juga bahwa, beberapa konteks dari (Sakti, 2008) belum terdapat pada (Vania, 2011), seperti posisi fonem pada kata, jumlah fonem pada kata, jumlah kata pada kalimat, dan pemosisian dari belakang (*backward*) untuk beberapa konteks. Penggabungan keseluruhan dari konteks tersebut patut dicoba untuk dilihat apakah terdapat peningkatan kualitas dari sistem. Hal ini menjadi salah satu target dari Tugas Akhir ini.

BAB III.

ANALISIS DAN PERANCANGAN SISTEM TEXT-TO-SPEECH BERBASIS HMM UNTUK BAHASA INDONESIA

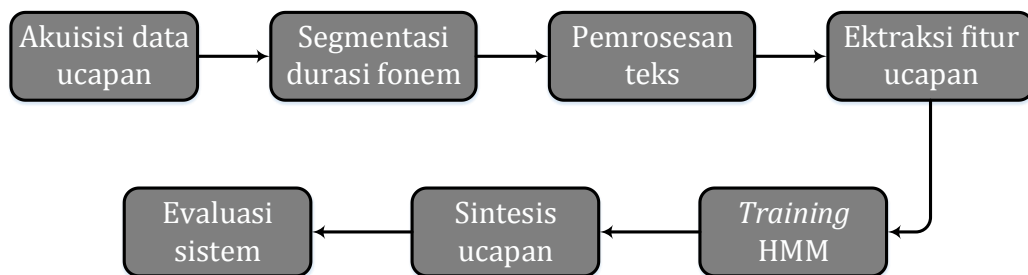
Pada bab ini dijelaskan mengenai analisis dan perancangan dari sistem TTS berbasis HMM untuk bahasa Indonesia. Berdasar kepada dasar teori mengenai sistem *text-to-speech* (TTS) berbasis *Hidden Markov Model* (HMM), di bawah ini merupakan gambaran dari sistem yang TTS yang dikembangkan pada Tugas Akhir ini.



Gambar III-1 Arsitektur sistem TTS berbasis HMM

Secara singkat, sistem TTS berbasis HMM terdiri dari dua proses utama, yaitu proses *training* dan proses *synthesis*. Proses *training* mencakup pemrosesan sinyal ucapan menjadi parameter ucapan yang dibutuhkan, seperti frekuensi fundamental dan spektrum frekuensi. Proses *training* juga membutuhkan adanya komponen pemroses teks untuk membuat label kontekstual, serta langkah segmentasi durasi fonem untuk memberikan informasi dari setiap segmen pengucapan. Label kontekstual di sini maksudnya adalah label yang berisikan segmen-segmen pengucapan, seperti rangkaian fonem, disertai dengan konteks yang dimilikinya, seperti posisi fonem tersebut pada kalimat, fonem di sekitarnya, dsb. Ketiga fitur ini dipakai dalam pembuatan model akustik HMM. Model akustik yang terbentuk kemudian digunakan dalam proses *synthesis*. Teks masukan yang baru diubah ke dalam label kontekstual menggunakan pemroses teks. Label kontekstual tersebut digunakan untuk mencari parameter ucapan, yaitu frekuensi fundamental dan spektrum frekuensi, yang sesuai dari HMM yang ada.

Untuk dapat mengembangkan sistem TTS berbasis HMM seperti gambaran di atas, khususnya untuk bahasa Indonesia, diperlukan beberapa tahap pengembangan sistem. Di bawah ini adalah tahapan yang perlu dilakukan dalam Tugas Akhir ini untuk membuat sistem TTS berbasis HMM untuk bahasa Indonesia secara menyeluruh.



Gambar III-2 Tahapan rancangan pembuatan sistem

Proses pembuatan sistem ini diawali oleh proses akuisisi data ucapan untuk mendapatkan sinyal ucapan beserta teks kalimat yang menyertainya. Dari sinyal ucapan tersebut, perlu juga dilakukan segmentasi terhadap durasi dari setiap fonem pada kalimat yang ada. Komponen pemroses teks juga perlu dikembangkan untuk

mengubah teks kalimat menjadi konteks label yang sesuai. Kemudian, tahap analisis sinyal dilakukan untuk mengekstraksi parameter-parameter ucapan yang sesuai dari sinyal ucapan, seperti frekuensi fundamental dan frekuensi spektral. Dengan menggunakan seluruh fitur yang didapat dari ketiga tahapan di atas, dilakukan proses *training* untuk mendapatkan model akustik HMM yang sesuai. Model HMM tersebut kemudian digunakan untuk mensintesis ucapan sintetis sesuai teks masukan. Terakhir, evaluasi dari sistem tersebut dilakukan untuk menentukan kinerjanya. Setiap proses-proses tersebut dijelaskan lebih lanjut pada subbab-subbab berikut.

III.1. Akuisisi data ucapan

Untuk dapat membuat sistem TTS, khususnya yang berbasis HMM, tentunya diperlukan adanya data yang dijadikan referensi dalam proses sintesis ucapan nantinya. Sebagai referensi di sini maksudnya adalah, sistem memiliki pengetahuan mengenai bagaimana harus mensintesis ucapan dari teks masukan sesuai data awal yang sudah diproses terlebih dahulu. Proses ini tentunya dengan melakukan metode pembelajaran sehingga menghasilkan suatu model yang dapat dipahami oleh sistem.

Data yang dibutuhkan ini terdiri dari kumpulan teks kalimat dalam bahasa Indonesia beserta data suara pengucapannya. Kumpulan kalimat ini harus memiliki minimal sifat seimbang secara fonem (*phonetically balanced*). Apabila sifat ini sudah terpenuhi, maka penambahan data kalimat diperlukan untuk menambah variasi konteks pengucapan yang mungkin. Data suara pengucapannya sendiri setidaknya direkam dengan menggunakan alat rekam yang memiliki kualitas baik, dengan kondisi kebisingan yang rendah. Ini untuk meminimalisir *noise/gangguan* terhadap sinyal, sehingga sinyal ucapan yang digunakan nantinya memang sesuai dengan pengucapan kalimatnya.

III.2. Segmentasi durasi

Setelah mendapatkan data suara ucapan beserta teksnya, maka langkah berikutnya adalah mengekstraksi fitur-fitur yang diperlukan, diawali dengan fitur durasi.

Durasi diperlukan oleh sistem untuk menentukan panjang-pendeknya suatu kumpulan sinyal yang mewakili satuan bunyi dibangkitkan. Dalam sistem ini, satuan bunyi yang dimaksud adalah fonem. Oleh karena itu, diperlukan adanya pengetahuan akan durasi dari setiap fonem yang ada pada setiap kalimat dari data ucapan yang dimiliki.

Proses segmentasi ini sendiri dapat dilakukan dengan dua cara, yaitu secara otomatis atau manual. Untuk melakukan proses segmentasi secara otomatis, diperlukan suatu proses yang dapat mempelajari rangkaian fonem beserta karakteristik ucapannya. Melalui pengetahuan tersebut, proses ini dapat melakukan pengaturan otomatis (*forced alignment*) untuk menentukan durasi dari setiap kata dan fonem yang ada. Proses ini pada dasarnya merupakan bagian dari sistem pengenalan ucapan (*speech recognition*). Cara manual sendiri tidak menggunakan proses komputasi untuk menentukan durasi dari segmen pengucapan, melainkan diperlukannya usaha manusia untuk mengatur setiap label dari segmen pengucapan agar sesuai dengan durasi dari sinyalnya. Cara ini memang tidak memungkinkan untuk dilakukan pada data yang sangat besar. Namun, metode ini dapat menjadi efektif apabila dilakukan untuk memperbaiki segmentasi otomatis yang dihasilkan oleh proses komputasi.

III.3. Pemrosesan teks

Setelah mendapatkan durasi dari setiap segmen pengucapan, yaitu durasi dari fonem, berikutnya diperlukan adanya komponen yang dapat mengekstraksi konteks yang dimiliki setiap fonem tersebut di dalam kalimat. Komponen ini dapat dinamakan pemrosesan teks karena menyangkut bagaimana sebuah teks masukan diolah menjadi suatu data yang lain yang lebih bermanfaat. Konteks, dalam hal ini merupakan variasi dari suatu fonem pada pengucapannya. Variasi tersebut dapat disebabkan oleh berbagai hal, seperti posisi fonem pada saat pengucapan, fonem sesudah dan sebelum yang dapat mempengaruhi artikulasinya, penekanan pada pengucapan, dsb. Berbagai kemungkinan konteks tersebut dapat menghasilkan variasi karakteristik pengucapan dari suatu fonem. Dalam Tugas Akhir ini,

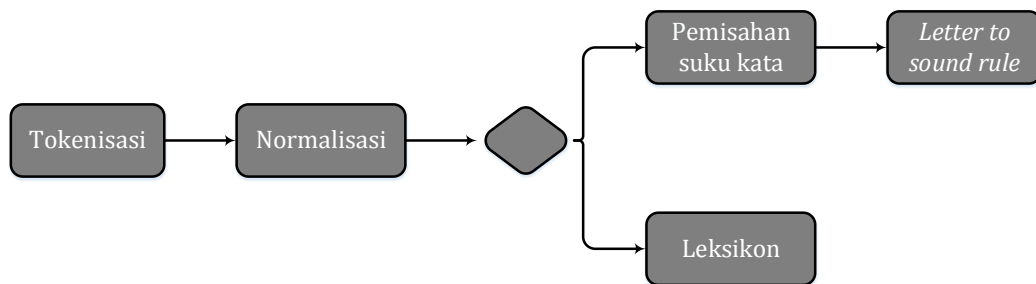
digunakan tiga set konteks, yang mengacu kepada (Sakti, 2008; Vania, 2011) dan penggabungannya. Berikut adalah daftar konteks yang digunakan.

Tabel III-1 Variasi daftar konteks yang digunakan

Set Konteks 1	Set Konteks 2	Set Konteks 3
Tingkat Fonem	Tingkat Fonem	Tingkat Fonem
Dua fonem sebelum	Dua fonem sebelum	Dua fonem sebelum
Fonem sebelum	Fonem sebelum	Fonem sebelum
Fonem saat ini	Fonem saat ini	Fonem saat ini
Fonem setelah	Fonem setelah	Fonem setelah
Dua fonem setelah	Dua fonem setelah	Dua fonem setelah
Posisi fonem pada kata (dari depan)	Posisi fonem pada suku kata	Posisi fonem pada kata (dari depan)
Posisi fonem pada kata (dari belakang)	-	Posisi fonem pada kata (dari belakang)
-	-	Posisi fonem pada suku kata (dari depan)
-	-	Posisi fonem pada suku kata (dari belakang)
Tingkat Suku Kata	Tingkat Suku Kata	Tingkat Suku Kata
-	Jumlah fonem pada suku kata sebelum	Jumlah fonem pada suku kata sebelum
-	Penekanan pada suku kata sebelum	Penekanan pada suku kata sebelum
-	Jumlah fonem pada suku kata saat ini	Jumlah fonem pada suku kata saat ini
-	Penekanan pada suku kata saat ini	Penekanan pada suku kata saat ini
-	Posisi suku kata pada kata	Posisi suku kata pada kata (dari depan)
-	Jarak dengan suku kata sebelumnya yang memiliki penekanan	Posisi suku kata pada kata (dari belakang)

-	Jarak dengan suku kata setelahnya yang memiliki penekanan	Jarak dengan suku kata sebelumnya yang memiliki penekanan
-	Huruf hidup pada suku kata ini	Jarak dengan suku kata setelahnya yang memiliki penekanan
-	Jumlah fonem pada suku kata sesudah	Huruf hidup pada suku kata ini
-	Penekanan pada suku kata sesudah	Jumlah fonem pada suku kata sesudah
-	-	Penekanan pada suku kata sesudah
Tingkat Kata	Tingkat Kata	Tingkat Kata
Jumlah fonem pada kata sebelum	Jumlah suku kata pada kata sebelum	Jumlah fonem pada kata sebelum
Jumlah fonem pada kata saat ini	Jumlah suku kata pada kata saat ini	Jumlah suku kata pada kata sebelum
Posisi kata pada kalimat (dari depan)	Posisi kata pada kalimat	Jumlah fonem pada kata saat ini
Posisi kata pada kalimat (dari belakang)	Jumlah suku kata pada kata sesudah	Jumlah suku kata pada kata saat ini
Jumlah fonem pada kata sesudah	-	Posisi kata pada kalimat (dari depan)
-	-	Posisi kata pada kalimat (dari belakang)
-	-	Jumlah fonem pada kata sesudah
-	-	Jumlah suku kata pada kata sesudah
Tingkat Kalimat	Tingkat Kalimat	Tingkat Kalimat
Jumlah kata total	Jumlah suku kata total	Jumlah suku kata total
-	-	Jumlah kata total

Untuk dapat menghasilkan informasi konteks yang sesuai untuk setiap segmen pengucapan, diperlukan beberapa tahapan dalam komponen pemrosesan teks. Berikut adalah diagram dari tahapan tersebut.



Gambar III-3 Tahapan pemrosesan teks

Berdasarkan diagram tersebut, berikut penjelasan per tahapnya lebih lanjut:

a. Tokenisasi

Proses pemrosesan teks diawali dengan tokenisasi untuk memudahkan pemrosesan lebih lanjut. Tokenisasi ini dimaksudkan untuk memisahkan rangkaian karakter pada kalimat menjadi kumpulan token-token yang dipisahkan berdasarkan spasi.

b. Normalisasi

Langkah berikutnya adalah melakukan proses normalisasi untuk mengubah karakter-karakter selain alfabet ke dalam rangkaian karakter alfabet. Perubahan dilakukan terhadap angka dan karakter khusus. Karakter khusus, seperti + (tambah), % (persen), dsb. dibaca sesuai pengucapannya dalam bahasa Indonesia.

c. Leksikon

Setelah dilakukan normalisasi, token-token tersebut akan diubah ke dalam rangkaian fonem pembentuknya. Perubahan tersebut bisa dilakukan dengan dua cara, salah satunya adalah dengan menggunakan leksikon/kamus. Leksikon ini terdiri dari kata beserta rangkaian fonemnya. Selain kata biasa, terdapat juga beberapa kata “khusus”, yaitu singkatan dan kata yang mengandung huruf ‘e’. Ini ditujukan agar sistem dapat mengatasi perubahan token yang berupa singkatan dan dualitas dari huruf ‘e’ yang memiliki dua jenis bunyi. Sayangnya, pada Tugas Akhir ini, leksikon ini masih belum ada, sehingga singkatan dan dualitas bunyi dari huruf ‘e’ belum dapat ditangani.

d. Pemisahan suku kata

Selain dengan menggunakan leksikon, cara kedua untuk melakukan pengubahan dari kata menjadi rangkaian fonemnya adalah dengan menggunakan aturan. Aturan tersebut pada bahasa Indonesia terdiri dari dua tahap, yaitu pemisahan suku kata dan penggunaan aturan huruf ke bunyi (*letter to sound/LTS rule*). Pertama-tama pemisahan suku kata dilakukan dengan menggunakan aturan yang terdapat pada bahasa Indonesia. Kemudian, setiap suku kata sebelum suku kata terakhir diberikan tanda penekanan pengucapan.

e. *Letter to sound rule*

Setelah kumpulan suku kata sudah dapat dibentuk, dilakukan pengubahan dari huruf-hurufnya ke dalam bunyi yang sesuai. Aturan LTS dalam bahasa Indonesia sangat sederhana karena bunyi/fonemnya serupa dengan penulisan hurufnya.

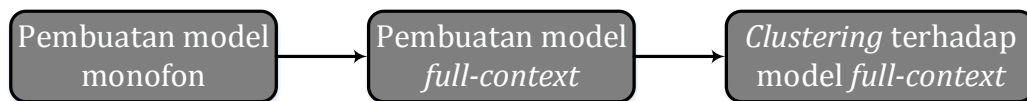
III.4. Ekstraksi parameter ucapan

Sinyal ucapan terdiri dari dua parameter penting, yaitu frekuensi fundamental dan frekuensi spektral. Frekuensi fundamental menentukan tinggi rendahnya nada pengucapan (*pitch*). Frekuensi spektral menentukan warna suara (*timbre*). Keduanya diekstrak dari sinyal ucapan murni menggunakan teknik analisis sinyal, seperti *Fast Fourier Transform* (FFT). Dengan menggunakan metode FFT, sebuah sinyal dapat dipecah menjadi sinyal-sinyal yang lebih sederhana.

Pada Tugas Akhir ini, proses analisis sinyal dilakukan dengan menggunakan kakas pendukung yang sudah ada. Beberapa hal yang perlu diperhatikan adalah pengaturan konfigurasi terhadap kondisi ekstraksi. Pertama, konfigurasi dari batas atas dan batas bawah dari ekstraksi frekuensi fundamental (dalam Hz). Kedua, *sampling rate* dari sinyal serta panjang dari analisis FFT yang disesuaikan. Untuk suara wanita, jarak frekuensi dasar, berkisar antara 120-400 Hz, sedangkan untuk pria antara 60-350 Hz. Untuk *sampling rate* 16 kHz, panjang FFT yang dianjurkan adalah 512, sedangkan untuk 44 kHz dianjurkan panjangnya 2048.

III.5. Training HMM

Setelah semua fitur, yaitu label kontekstual, parameter ucapan, seperti frekuensi fundamental, frekuensi spektral dan durasinya, berhasil didapatkan, langkah berikutnya adalah membuat model akustik HMM yang sesuai dengan data-data tersebut. Pembuatan model HMM ini dinamakan dengan proses *training*. Dalam proses *training* ini terdapat beberapa tahap, seperti dapat dilihat pada diagram berikut.



Gambar III-4 Proses *training* HMM

Pertama-tama, model monofon HMM dibentuk, diambil dari label monofon (label tanpa konteks). Proses inisialisasi terhadap model monofon dilakukan dengan pendekatan algoritma *viterbi*, kemudian dilakukan re-estimasi untuk memperbaiki nilainya dengan pendekatan algoritma *baum-welch*. Setelah model HMM didapatkan untuk setiap fonem yang ada, dilakukan pembuatan model *full-context* untuk setiap konteks yang ada pada seluruh data label konteks. Proses re-estimasi dilakukan terhadap model awal dari model monofon dengan algoritma *baum-welch*. Untuk mengatasi berbagai kemungkinan konteks lain yang tidak terdapat pada data, maka perlu dilakukan *clustering* terhadap model *full-context*. Metode *clustering* ini dilakukan dengan menggunakan pendekatan *decision tree*. Pertanyaan-pertanyaan pada setiap simpulnya ditentukan berdasarkan konteks-konteks yang digunakan. Konteks tersebut dapat terdiri dari cara atau letak pengucapan dari fonem bahasa Indonesia, serta variasi dari posisi fonem, suku kata atau kata pada kalimat yang mungkin pada bahasa Indonesia. Pembuatan *decision tree* ini akan mengelompokkan model-model kontekstual yang memiliki kemiripan ke dalam satu model *cluster* HMM sehingga dapat menangani berbagai kemungkinan variasi konteks yang tidak terdapat pada data.

III.6. Sintesis ucapan

Pada akhirnya, setelah model akustik HMM dibentuk, proses sintesis ucapan dapat dilakukan. Proses sintesis ucapan ini dilakukan dengan mencari parameter terbaik dari model yang dibentuk, sesuai dengan fitur masukan yang diberikan. Fitur masukan yang diberikan adalah label kontekstual sesuai dengan teks masukan. Sedangkan model yang diacu adalah model yang dihasilkan dari proses *training*, yang terdiri dari *decision tree*, serta distribusi probabilitas (*probability distribution function/pdf*) dari setiap *cluster* HMM yang ada.

Label kontekstual masukan memiliki nilai-nilai yang sesuai dengan pertanyaan-pertanyaan dari *decision tree* sesuai konteks pada bahasa Indonesia. Penelusuran dilakukan terhadap setiap label konteks tersebut untuk setiap fonem, sampai ditemukan simpul daunnya. Simpul daun tersebut merupakan *cluster* HMM yang sesuai dengan label konteks yang ditelusuri pohonnya. Sesuai dengan *cluster* yang didapat, maka parameter ucapan dapat dibangkitkan dengan melihat ke dalam distribusi probabilitas yang dimiliki oleh *cluster* HMM tersebut. Parameter yang berupa frekuensi fundamental, frekuensi spektral dan durasinya kemudian dimasukkan ke dalam pemroses sinyal untuk diolah dan dibentuk menjadi gelombang ucapan yang dapat dimengerti.

III.7. Evaluasi sistem

Untuk menilai kinerja dari sistem *text-to-speech* (TTS) berbasis HMM untuk bahasa Indonesia ini, diperlukan tahapan evaluasi. Evaluasi dilakukan terhadap suara sintetis untuk menilai beberapa aspek, yaitu kualitas dan kealamian dari ucapan, serta kejelasan dan akurasi dari pengucapan. Kualitas dan kealamian menunjukkan seberapa persisnya suara sintetis tersebut dengan pengucapan manusia sebenarnya. Kejelasan dan akurasi mengukur ketepatan pengucapan dari suara sintetis terhadap teks masukan.

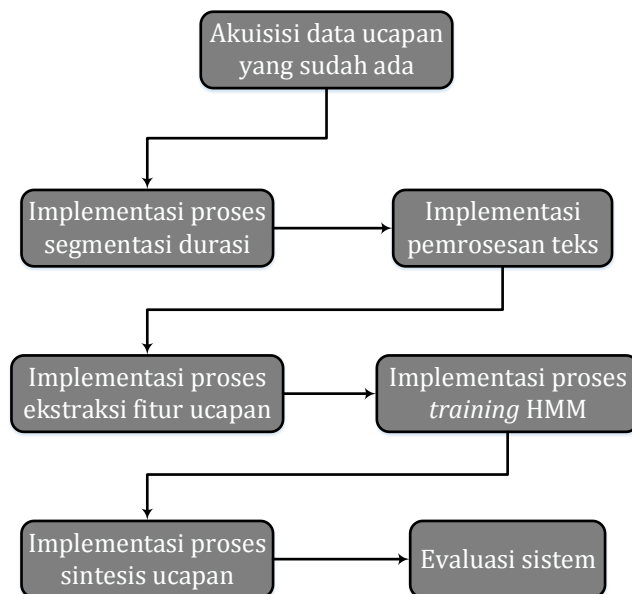
Kualitas dan kealamian dievaluasi dengan menggunakan metode *Mean Opinion Score* (MOS). Pada metode ini, beberapa ucapan sintetis yang sudah dihasilkan dari sistem diperdengarkan kepada sejumlah evaluator. Masing-masing

evaluator tersebut akan menilai kualitas dan kealamian dari setiap ucapan dalam 5 skala nilai, 1 (sangat buruk) sampai 5 (sangat baik). Sedangkan kejelasan dan akurasi dievaluasi dengan metode *Semantically Unpredictable Sentences* (SUS). Pada metode SUS, evaluator akan mendengarkan sejumlah kalimat yang kata-katanya disusun secara acak dalam bahasa Indonesia. Pengacakan kata-kata ini dilakukan untuk menghindari penafsiran dari setiap evaluator sesuai konteks kalimatnya. Walaupun demikian, terdapat beberapa struktur kalimat yang harus diikuti dalam pengacakan katanya. Terdapat 5 struktur kalimat dalam bahasa Indonesia yang digunakan dalam metode SUS ini, yaitu kalimat *Transitive*, *Intransitive*, *Imperative*, *Interrogative*, dan *Relative*. Pengukuran nilai dari metode SUS ini dilakukan terhadap ketepatan kata dan kalimat secara keseluruhan.

BAB IV.

IMPLEMENTASI SISTEM PENSINTESIS UCAPAN BAHASA INDONESIA BERBASIS HIDDEN MARKOV MODEL

Setelah melakukan analisis dan perancangan terhadap sistem TTS yang dikembangkan pada Tugas Akhir ini, pada bab ini dijelaskan mengenai proses implementasi dari sistem tersebut untuk menghasilkan produk yang nyata sampai pada tahap evaluasi. Sebagai gambaran awal, tahapan implementasi dapat dilihat pada gambar berikut.



Gambar IV-1 Tahapan proses implementasi sistem TTS

Pertama-tama, ditentukan penggunaan data ucapan pada sistem TTS yang dikembangkan. Data ucapan menggunakan data yang sudah ada, dengan kata lain tidak merekam sendiri. Kemudian proses implementasi sistem mulai dilakukan. Implementasi diawali dengan proses segmentasi dari durasi fonem, implementasi pemrosesan teks, proses ekstraksi fitur ucapan, proses *training* HMM dan proses sintesis ucapan. Setelah seluruh proses implementasi dilakukan, tahapan diakhiri

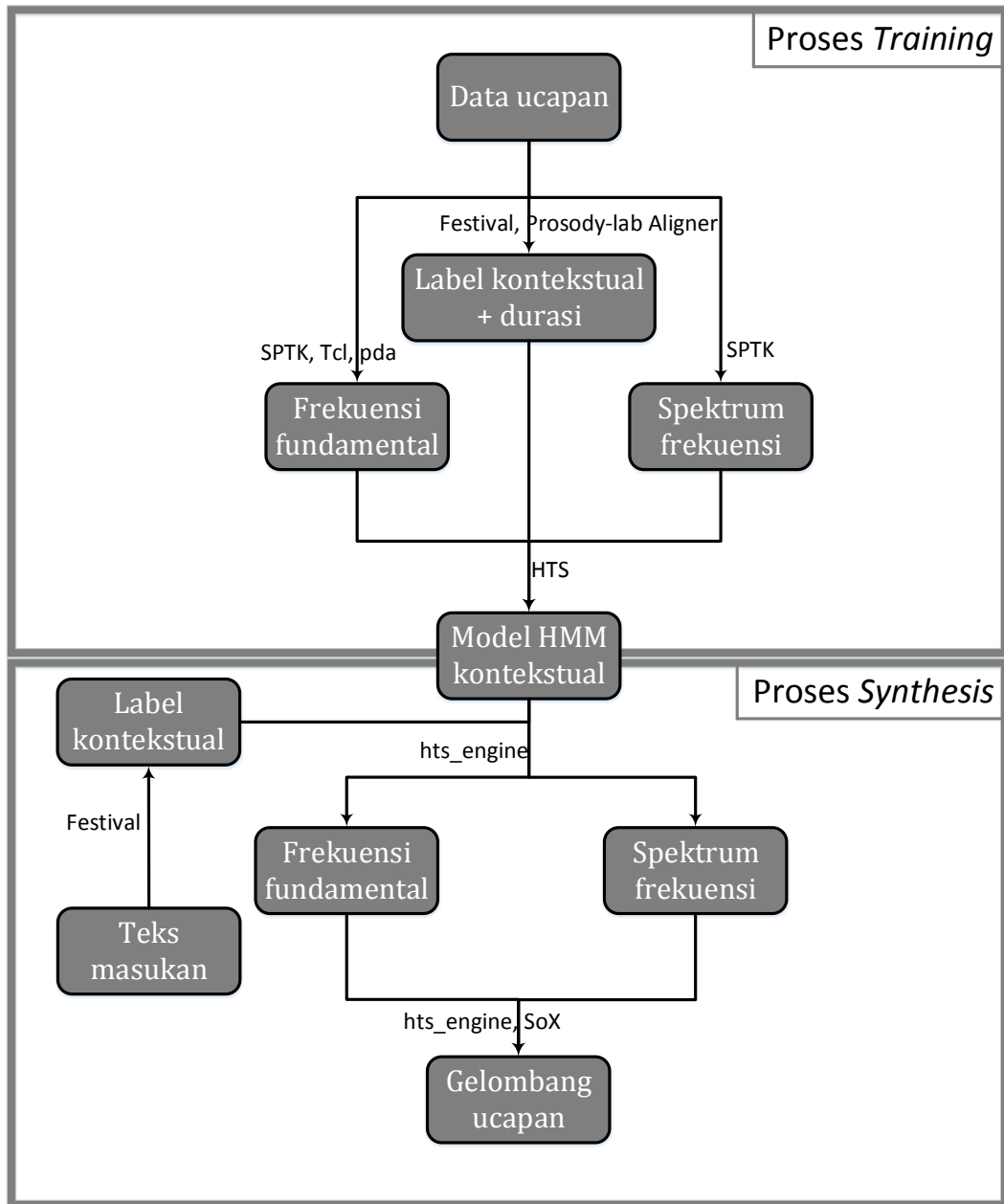
dengan evaluasi terhadap sistem, secara khusus terhadap ucapan sintetis yang dihasilkan.

Dalam setiap tahapan implementasi pada sistem TTS ini dibutuhkan beberapa kakas terkait. Berikut adalah daftar kakas-kakas yang digunakan beserta tahapan penggunaannya.

Tabel IV-1 Daftar kakas yang digunakan pada implementasi sistem

Nama Kakas	Website	Digunakan Pada
HTS (HMM-based speech synthesis system)	http://hts.sp.nitech.ac.jp/	proses <i>training</i> HMM
<i>Prosody-lab Aligner</i>	http://prosodylab.org/tools/aligner	proses segmentasi durasi
HTK (<i>Hidden markov model toolkit</i>)	http://htk.eng.cam.ac.uk/	instalasi HTS dan instalasi <i>Prosodylab-Aligner</i>
<i>Festival</i>	http://www.cstr.ed.ac.uk/projects/festival/	pemrosesan teks
<i>Edinburgh speech tools</i>	http://www.cstr.ed.ac.uk/projects/speech_tools/	instalasi <i>Festival</i>
SPTK	http://sp-tk.sourceforge.net/	ekstraksi fitur ucapan
Tcl, SoX, pda	dari <i>Linux Operating System</i>	ekstraksi fitur ucapan dan proses sintesis ucapan (SoX)
<i>hts_engine</i>	http://hts-engine.sourceforge.net/	proses sintesis ucapan

Dengan mengacu kepada daftar kakas di atas, untuk memudahkan penggambaran dari arsitektur implementasi sistem, maka pengalokasian tiap kakas untuk tiap proses dapat dilihat pada gambar berikut.



Gambar IV-2 Penggunaan kakas dalam implementasi dari sistem TTS

Secara ringkas, gambar di atas menjelaskan bagaimana mengimplementasikan setiap proses pada pembuatan sistem TTS di Tugas Akhir ini. Proses secara keseluruhan dibagi menjadi dua bagian utama, yaitu proses *training* dan proses *synthesis*. Pada proses *training*, data ucapan diekstrak fitur-fiturnya melalui proses pemrosesan teks dan segmentasi durasi, serta analisis sinyal untuk ekstraksi fitur ucapan. Pada pemrosesan teks digunakan kakas *Festival*. Pada

segmentasi durasi digunakan kakas *Prosody-lab Aligner*. Pada ekstraksi fitur ucapan digunakan kakas SPTK, Tcl dan pda. Fitur-fitur tersebut digunakan pada proses *training* HMM untuk menghasilkan model akustik HMM kontekstual yang merepresentasikan pemodelan fitur teks dengan ucapan. Pada proses *training* ini digunakan kakas HTS. Setelah model HMM didapat, berikutnya masuk ke dalam proses *synthesis*. Kembali dengan menggunakan kakas *Festival*, teks masukan diubah menjadi label kontekstualnya. Label ini bersama dengan model HMM dimasukkan ke dalam kakas *hts_engine* untuk dicari fitur ucapannya, yaitu frekuensi fundamental dan spektrum frekuensi yang sesuai. Fitur-fitur ucapan ini pada akhirnya diubah ke dalam gelombang ucapan oleh kakas *hts_engine* tersebut dan juga SoX. Berikut akan dijelaskan lebih lanjut mengenai penggunaan kakas-kakas tersebut di dalam setiap prosesnya, dengan diawali penjelasan terhadap data ucapan yang digunakan. Evaluasi dari sistem sendiri dijelaskan terpisah pada bab berikutnya.

IV.1. Spesifikasi data ucapan

Pada Tugas Akhir ini akuisisi data ucapan tidak dilakukan dengan perekaman dari nol, melainkan menggunakan data ucapan yang sudah ada. Data ucapan ini terdiri dari kalimat-kalimat dalam bahasa Indonesia. Kalimat diucapkan oleh seorang pembicara wanita.

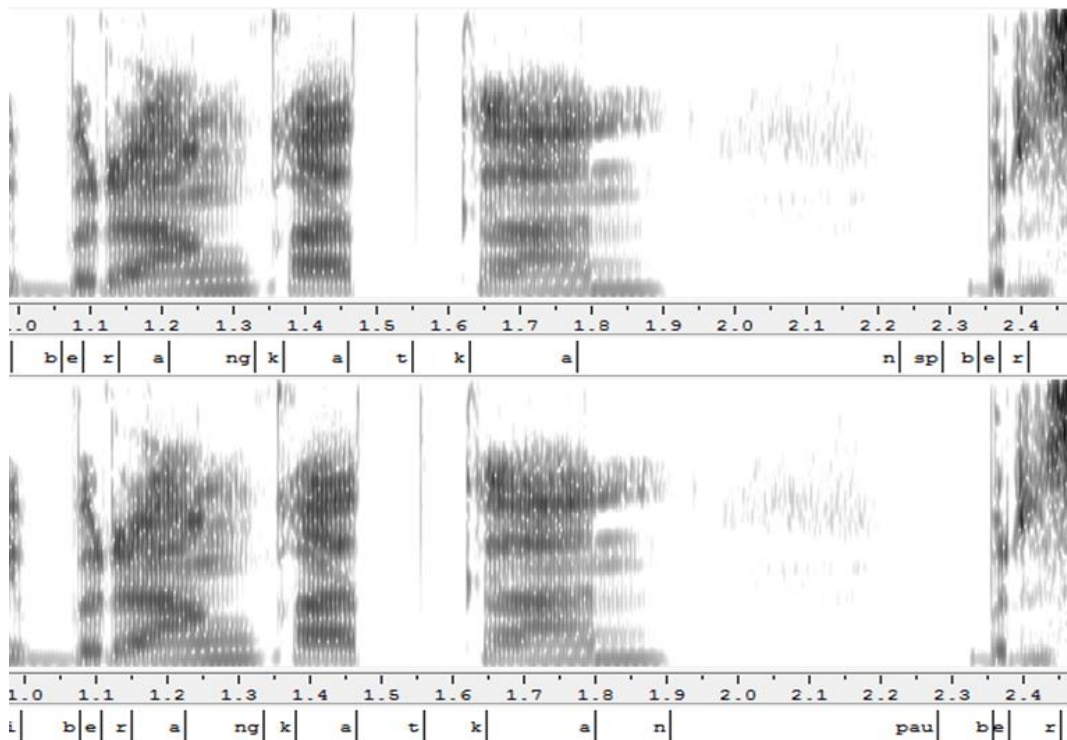
Total kalimat dalam data ucapan adalah 343 buah. Data ucapan ini sudah seimbang secara fonetik (*phonetically balanced*). *Sampling frequency* untuk sinyal ucapannya sendiri adalah 16000 Hz. Set kalimat yang ada pada data ini seluruhnya merupakan kalimat pernyataan.

IV.2. Implementasi proses segmentasi ucapan

Untuk dapat membuat model akustik terhadap setiap satuan bunyi (dalam hal ini fonem) dari data ucapan, diperlukan informasi mengenai selang waktu untuk setiap fonem pada setiap pengucapan yang ada. Selang waktu atau durasi tersebut digunakan untuk menandai bahwa fitur yang digunakan dalam pembuatan model fonem tersebut menggunakan fitur dalam selang waktu yang telah ditentukan itu.

Penentuan durasi harus dilakukan untuk setiap satuan bunyi dari setiap kalimat pengucapan yang ada.

Proses segmentasi akan sangat melelahkan apabila dilakukan secara manual dari nol, terlebih lagi apabila dilakukan untuk data yang jumlahnya tidak sedikit. Untuk mengefisiensikan proses, segmentasi dapat dilakukan secara otomatis. Teknik yang dilakukan bernama *forced alignment*. Teknik ini pada dasarnya mengimplementasikan proses pengenalan ucapan (*speech recognition*). Hanya saja, dalam proses pengenalan ucapannya, data yang dimasukkan adalah data yang sama dengan data yang digunakan dalam pembuatan modelnya. Proses ini pada dasarnya memanfaatkan kaskas HTK. Namun, untuk memudahkan pengimplementasiannya, digunakan kaskas *Prosody-lab Aligner* yang sudah mengintegrasikan setiap langkah-langkah dalam proses *forced alignment* pada HTK.



Gambar IV-3 Gambar atas adalah hasil dari segmentasi otomatis, gambar bawah hasil dari perbaikannya

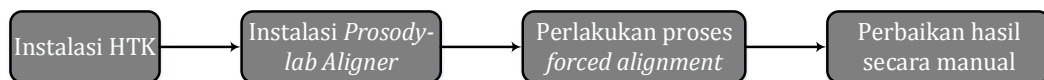
Dikarenakan data yang digunakan untuk membuat model akustik pada proses *forced alignment* tidak terlalu banyak, maka hasil segmentasi yang didapat masih terdapat banyak kesalahan. Oleh karena itu, diperlukan perbaikan secara manual terhadap hasil dari segmentasi otomatis tersebut. Gambar di atas merupakan perbandingan contoh hasil segmentasi otomatis yang masih memiliki beberapa kesalahan, terutama di label “pau” dengan perbaikan manualnya.

Setelah melakukan segmentasi otomatis dan perbaikan manualnya, maka didapat label fonem beserta durasinya untuk setiap pengucapan. Berikut adalah contoh label monofon untuk kalimat “hai selamat pagi, apa kabar”.

	0	3340872	pau
3340872		4143586	h
4143586		6885249	ay
6885249		8863768	s
8863768		9208596	e
9208596		9592994	l
9592994		10604866	a
10604866		11339745	m
11339745		12057665	a
12057665		12775585	t
12775585		13561340	p
13561340		14658005	a
14658005		15472024	g
15472024		17026576	i
17026576		18598086	pau
18598086		19479940	a
19479940		20390059	p
20390059		21240050	a
21240050		22249868	k
22249868		23188251	a
23188251		24126635	b
24126635		25986444	a
25986444		27054844	r
27054844		28300000	pau

Gambar IV-4 Label monofon dan durasinya untuk kalimat "hai selamat pagi, apa kabar"

Berikut adalah tahapan dalam proses segmentasi durasi pada Tugas Akhir ini.



Gambar IV-5 Tahapan implementasi proses segmentasi durasi

IV.3. Implementasi pemrosesan teks

Festival adalah sebuah kakas yang menyediakan framework untuk membangun sebuah sistem pensintesis suara, baik terhadap bahasa yang sudah tersedia, maupun dengan bahasa baru yang belum terdapat pada framework yang disediakan. Sistem Festival sendiri dibangun dalam bahasa C++ dan menggunakan Edinburgh Speech Tools untuk pemrosesan sinyal. Kakas Festival menggunakan Scheme interpreter (SIOD – Scheme in One Defun), yang berbasis bahasa fungsional LISP, sebagai bahasa untuk kontrol perintah (command control) dari pengguna.

Untuk menjalankan proses instalasi kakas *Festival* ini, diperlukan beberapa persyaratan, seperti:

- *Festival Speech Synthesis System Source* (dalam tugas akhir ini digunakan versi 2.1)
- Data leksikon (untuk ucapan dalam bahasa *default* pada *Festival*)
- Data ucapan (untuk ucapan dalam bahasa *default* pada *Festival*)
- *Edinburgh Speech Tools* (dalam tugas akhir ini juga digunakan versi 2.1)
- Sistem operasi UNIX (dalam tugas akhir ini digunakan Ubuntu v10.4)
- *Compiler C++* (dalam tugas akhir ini digunakan versi 4.5)
- *GNU make* (dalam tugas akhir ini digunakan versi 3.81)

Proses instalasi untuk kakas *Festival* ini sendiri tergolong cukup mudah. Berikut adalah langkah-langkah instalasi dalam *linux command*:

1. \$ cd <direktori *Edinburgh Speech Tools*>
2. \$./configure
3. \$ make
4. \$ cd <direktori *Festival*>
5. \$./configure
6. \$ make

Seperti dapat dilihat dari langkah-langkah di atas bahwa, sebelum melakukan instalasi terhadap *Festival*, instalasi terhadap *Edinburgh Speech Tools* perlu dilakukan terlebih dahulu. Hal ini dikarenakan bahasa *default* pada *Festival*, yaitu

bahasa Inggris memerlukan kakas ini dalam prosesnya. Walaupun demikian, untuk sistem pensintesis ucapan bahasa Indonesia di tugas akhir ini, *speech tools* tersebut tidak akan digunakan.

Sistem penganalisis teks ini yang dikembangkan dengan kakas *Festival* ini, pada dasarnya bertujuan untuk melakukan pengubahan teks masukan pengguna menjadi data *utterance*, yang merupakan data ucapan dalam format festival, untuk diproses lebih lanjut ke dalam label transkripsi yang akan digunakan oleh HTS nantinya. Pada pembuatan sistem penganalisis teks untuk bahasa Indonesia pada Tugas Akhir ini, digunakan sistem penganalisis teks bahasa Inggris yang tersedia pada kakas sebagai dasar referensi.

Proses analisis teks pada *Festival* sendiri terbagi ke dalam beberapa langkah-langkah utama, yaitu:

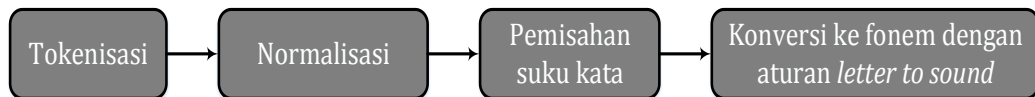
1. Penentuan data fonem (*phoneset*)
2. Proses tokenisasi dan normalisasi (pemisahan token-token menjadi kata)
3. Proses *part-of-speech tagging* (pelabelan jenis kata)
4. Proses pelabelan suku kata dan fonem terhadap kata (melihat leksikon dan aturan *letter-to-sound*)
5. Proses *phrasing* (pemisahan frase)
6. Proses pelabelan konteks pada setiap fonem

Langkah-langkah di atas adalah langkah-langkah yang digunakan pada sistem sintesis ucapan bahasa Inggris yang tersedia pada kakas *Festival*. Untuk mengadaptasikannya ke dalam bahasa Indonesia, perlu dilakukan beberapa perubahan, seperti:

1. Pembuatan data fonem (*phoneset*) untuk bahasa Indonesia
2. Pembuatan proses pemisahan suku kata untuk bahasa Indonesia
3. Pembuatan aturan *letter-to-sound* untuk bahasa Indonesia
4. Pembuatan daftar konteks untuk bahasa Indonesia

Beberapa proses lainnya, seperti tokenisasi, *pos-tagging*, dan *phrasing* pada tugas akhir ini masih mengikuti dengan yang disediakan *Festival*. Hal ini dikarenakan,

tokenisasi memiliki proses yang serupa dengan bahasa Indonesia kecuali dalam hal karakter khusus, seperti angka, tanda baca, dsb. Selain itu, fitur *pos-tagging* dan frase masih belum digunakan pada pembuatan tugas akhir ini (digunakan untuk *future work*). Berikut adalah diagram dari langkah pemrosesan teks pada Tugas Akhir ini. Sistem tidak menggunakan kamus kata (leksikon).



Gambar IV-6 Implementasi pemrosesan teks

Keempat langkah-langkah di atas yang dibuat untuk mengadaptasi *Festival* ke dalam bahasa Indonesia, dijelaskan pada sub-subbab berikut.

IV.3.1. *Phonset* untuk bahasa Indonesia

Bahasa Indonesia memiliki 32 fonem yang berbeda, terdiri dari 9 vokal, termasuk 4 diftong, dan 22 konsonan, termasuk 4 gugus konsonan. Pada sistem ini digunakan pula satu fonem *silence* (jeda), sehingga secara keseluruhan menjadi 32 fonem yang berbeda. Untuk 31 fonem yang secara khusus dimiliki oleh bahasa Indonesia tersebut dijabarkan ke dalam dua tabel berikut, yang masing-masing memberikan daftar huruf vokal dan konsonan. Berikut adalah daftar untuk fonem vokal dan konsonan yang digunakan pada Tugas Akhir ini berdasar kepada (Sakti, 2008).

Tabel IV-2 Daftar fonem vokal pada bahasa Indonesia

	Depan	Tengah	Belakang
Tinggi	/i/	-	/u/
Sedang	/e/, /ey/	-	/o/
Rendah	/a/	/ay/, /aw/	/oy/

Tabel IV-3 Daftar fonem konsonan pada bahasa Indonesia

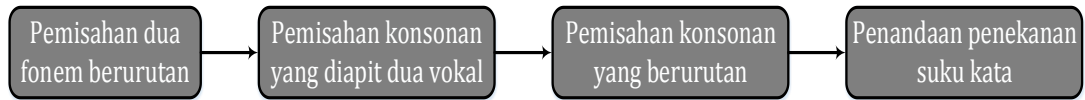
	Plosif	Afrikat	Frikatif	Sengau	Geletar	Lateral	Semivokal
Bilabial	/p/, /b/(+)	-	-	/m/	-	-	/w/
Labiodental	-	-	/f/	-	-	-	-
Alveolar	/t/, /d/(+)	-	/s/, /z/(+)	-	/r/	/l/	-
Palatal	-	/c/, /j/(+)	/sy/	/ny/	-	-	-
Velar	/k/, /g/(+)	-	/kh/	/ng/	-	-	/y/
Glottal	-	-	/h/	-	-	-	-

Dapat dilihat dari tabel fonem vokal, bahwa fonem-fonem vokal dibagi berdasarkan ketinggian posisi lidah (tinggi, sedang, rendah) dan posisi kedepanan (*frontness*) lidah pada rongga mulut (depan, tengah, belakang – dengan posisi ini bibir menjadi membulat/*rounding*). Dari 9 fonem vokal tersebut, terdapat 4 diftong di dalamnya, yaitu /ay/, /aw/, /ey/, dan /oy/.

Pada tabel fonem konsonan, masing-masing konsonan dikategorikan berdasarkan cara pengucapan (*manner of articulation*), letak pengucapan/terjadinya hambatan suara konsonan (*place of articulation*), dan pembunyian (*voiced/unvoiced*). Secara keseluruhan kategori tersebut adalah cara pengucapan yang dibagi ke dalam 7 jenis, letak pengucapan ke dalam 8 jenis, dan bergetar atau tidaknya pita suara saat pengucapan konsonan (tanda + menandakan konsonan tersebut memiliki pembunyian). Dari 22 konsonan tersebut, 4 di antaranya adalah gugus konsonan, yaitu /kh/, /ng/, /ny/, dan /sy/.

IV.3.2. Aturan pemisahan suku kata untuk bahasa Indonesia

Proses pemisahan suku kata dilakukan dengan mengacu kepada (Basuki, 2002). Berikut adalah tahapan dari pemisahan suku kata pada bahasa Indonesia.



Gambar IV-7 Tahapan pemisahan suku kata pada bahasa Indonesia

1. Apabila di tengah kata, terdapat dua huruf vokal (selain diftong) yang berurutan, pisahkan kedua huruf vokal tersebut.
2. Apabila di tengah kata, terdapat huruf konsonan yang tepat diapit dua huruf vokal, pemisahan dilakukan di antara huruf vokal pertama dan konsonan tersebut.
3. Apabila di tengah kata, terdapat dua atau lebih huruf konsonan yang berurutan, lakukan pemisahan setelah huruf konsonan yang pertama.
4. Penandaan penekanan pada suku kata sebelum suku kata terakhir.

Khusus untuk langkah terakhir, penekanan suku kata dilakukan pada suku kata sebelum suku kata terakhir mengacu kepada (Vania, 2011). Pemisahan suku kata ini bermanfaat untuk pelabelan konteks nantinya. Khususnya untuk konteks yang membutuhkan informasi tentang konteks dari fonem pada tingkat suku kata.

IV.3.3. Aturan *letter-to-sound* untuk bahasa Indonesia

Pada *Festival*, proses konversi fonem pada kata dilakukan terlebih dahulu dengan melakukan pencarian pada leksikon (daftar kamus kata dengan pasangan fonem-fonemnya). Apabila kata yang bersangkutan tidak terdapat pada leksikon, maka pelabelan fonem dilakukan dengan menggunakan aturan *letter-to-sound*. Pada Tugas Akhir ini hanya digunakan aturan *letter-to-sound* untuk melakukan pemetaan dari kata ke fonem-fonemnya. Tabel di bawah ini menjabarkan aturan yang dipakai pada implementasinya.

Tabel IV-4 Daftar huruf beserta pasangan fonemnya pada bahasa Indonesia

Huruf	Pasangan Fonemnya
ai	/ay/
au	/aw/
ei	/ey/
oi	/oy/
a	/a/
i	/i/
u	/u/
e	/e/
o	/o/
b	/b/
c	/c/
d	/d/
f	/f/
g	/g/
h	/h/
j	/j/
kh	/kh/
k	/k/
l	/l/
m	/m/
ng	/ng/
ny	/ny/
n	/n/
p	/p/
q	/k/
r	/r/
s	/s/
t	/t/
v	/f/
w	/w/
x	/ks/
y	/y/
z	/z/

IV.3.4. Pelabelan konteks terhadap fonem

Pelabelan konteks diperlukan untuk menandai setiap fonem pada suatu kalimat dengan berbagai konteks/kondisi fonem tersebut di kalimat yang bersangkutan. Konteks dari fonem itu dapat berupa konteks fonetis, yaitu fonem apa saja yang berada di sekitar fonem tersebut, atau dapat juga berupa konteks prosodi, seperti posisi fonem pada kata, penekanan suku kata, jumlah kata pada kalimat, dsb. Konteks ini berguna nantinya bagi kakas HTS untuk dapat memodelkan parameter ucapan dengan lebih akurat sesuai kondisi fonem pada saat pengucapan.

Pada Tugas Akhir ini digunakan tiga set konteks. Konteks pertama mengacu kepada (Sakti, 2008), konteks kedua mengacu kepada (Vania, 2011) dan konteks ketiga merupakan kombinasi/gabungan dari keduanya. Secara singkat, (Sakti, 2008) hanya menggunakan konteks pada tingkat fonem, kata dan kalimat. Sedangkan (Vania, 2011) sudah menekankan penggunaan konteks pada tingkat suku kata. Kombinasi dari keduanya dilakukan dengan mengikutsertakan konteks yang ada pada salah satu, tetapi tidak pada keduanya, seperti posisi fonem pada kata yang ada pada (Sakti, 2008) tetapi tidak pada (Vania, 2011), dsb. Berikut adalah ketiga set konteks yang digunakan beserta contoh keluaran label konteksnya.

Tabel IV-5 Set konteks 1 (Sakti, 2008) yang digunakan pada sistem TTS

Set Konteks 1
Tingkat Fonem
Dua fonem sebelum
Fonem sebelum
Fonem saat ini
Fonem setelah
Dua fonem setelah
Posisi fonem pada kata (dari depan)
Posisi fonem pada kata (dari belakang)
Tingkat Kata
Jumlah fonem pada kata sebelum
Jumlah fonem pada kata saat ini
Posisi kata pada kalimat (dari depan)
Posisi kata pada kalimat (dari belakang)

Jumlah fonem pada kata sesudah
Tingkat Kalimat
Jumlah kata total

$x^x\text{-pau+h=ay}@x_x/A:0/B:x@x-x/C:2/D:5$
 $x^{\text{pau-h+ay=s}}@1_2/A:0/B:2@1-5/C:7/D:5$
 $\text{pau}^{\text{h-ay+s=e}}@2_1/A:0/B:2@1-5/C:7/D:5$
 $\text{h}^{\text{ay-s+e=l}}@1_7/A:2/B:7@2-4/C:4/D:5$
 $\text{ay}^{\text{s-e+l=a}}@2_6/A:2/B:7@2-4/C:4/D:5$
 $\text{s}^{\text{e-l+a=m}}@3_5/A:2/B:7@2-4/C:4/D:5$
 $\text{e}^{\text{l-a+m=a}}@4_4/A:2/B:7@2-4/C:4/D:5$
 $\text{l}^{\text{a-m+a=t}}@5_3/A:2/B:7@2-4/C:4/D:5$
 $\text{a}^{\text{m-a+t=p}}@6_2/A:2/B:7@2-4/C:4/D:5$
 $\text{m}^{\text{a-t+p=a}}@7_1/A:2/B:7@2-4/C:4/D:5$
 $\text{a}^{\text{t-p+a=g}}@1_4/A:7/B:4@3-3/C:0/D:5$
 $\text{t}^{\text{p-a+g=i}}@2_3/A:7/B:4@3-3/C:0/D:5$
 $\text{p}^{\text{a-g+i=pau}}@3_2/A:7/B:4@3-3/C:0/D:5$
 $\text{a}^{\text{g-i+pau=a}}@4_1/A:7/B:4@3-3/C:0/D:5$
 $\text{g}^{\text{i-pau+a=p}}@x_x/A:4/B:x@x-x/C:3/D:5$
 $\text{i}^{\text{pau-a+p=a}}@1_3/A:0/B:3@4-2/C:5/D:5$
 $\text{pau}^{\text{a-p+a=k}}@2_2/A:0/B:3@4-2/C:5/D:5$
 $\text{a}^{\text{p-a+k=a}}@3_1/A:0/B:3@4-2/C:5/D:5$
 $\text{p}^{\text{a-k+a=b}}@1_5/A:3/B:5@5-1/C:0/D:5$
 $\text{a}^{\text{k-a+b=a}}@2_4/A:3/B:5@5-1/C:0/D:5$
 $\text{k}^{\text{a-b+a=r}}@3_3/A:3/B:5@5-1/C:0/D:5$
 $\text{a}^{\text{b-a+r=pau}}@4_2/A:3/B:5@5-1/C:0/D:5$
 $\text{b}^{\text{a-r+pau=x}}@5_1/A:3/B:5@5-1/C:0/D:5$
 $\text{a}^{\text{r-pau+x=x}}@x_x/A:5/B:x@x-x/C:0/D:5$

Gambar IV-8 Keluaran label dengan set konteks 1 untuk kalimat "hai selamat pagi, apa kabar"

Tabel IV-6 Set konteks 2 (Vania, 2011) yang digunakan pada sistem TTS

Set Konteks 2
Tingkat Fonem
Dua fonem sebelum
Fonem sebelum
Fonem saat ini
Fonem setelah
Dua fonem setelah
Posisi fonem pada suku kata
Tingkat Suku Kata
Jumlah fonem pada suku kata sebelum
Penekanan pada suku kata sebelum

Jumlah fonem pada suku kata saat ini
Penekanan pada suku kata saat ini
Posisi suku kata pada kata
Jarak dengan suku kata sebelumnya yang memiliki penekanan
Jarak dengan suku kata setelahnya yang memiliki penekanan
Huruf hidup pada suku kata ini
Jumlah fonem pada suku kata sesudah
Penekanan pada suku kata sesudah
Tingkat Kata
Jumlah suku kata pada kata sebelum
Jumlah suku kata pada kata saat ini
Posisi kata pada kalimat
Jumlah suku kata pada kata sesudah
Tingkat Kalimat
Jumlah suku kata total

$x^x\text{-pau+h=ay}$ @x/A:0_0/B:x_x@x!x-x^x/C:2_0/D:0/E:x@x/F:1/G:10
 $x^{\text{pau-h+ay=s}}$ @1/A:0_0/B:2_0@1!0-2^ay/C:2_0/D:0/E:1@1/F:3/G:10
 $\text{pau}^{\text{h-ay+s=e}}$ @2/A:0_0/B:2_0@1!0-2^ay/C:2_0/D:0/E:1@1/F:3/G:10
 $\text{h}^{\text{ay-s+e=l}}$ @1/A:2_0/B:2_0@1!0-1^e/C:2_1/D:1/E:3@2/F:2/G:10
 $\text{ay}^{\text{s-e+l=a}}$ @2/A:2_0/B:2_0@1!0-1^e/C:2_1/D:1/E:3@2/F:2/G:10
 $\text{s}^{\text{e-l+a=m}}$ @1/A:2_0/B:2_1@2!0-2^a/C:3_0/D:1/E:3@2/F:2/G:10
 $\text{e}^{\text{l-a+m=a}}$ @2/A:2_0/B:2_1@2!0-2^a/C:3_0/D:1/E:3@2/F:2/G:10
 $\text{l}^{\text{a-m+a=t}}$ @1/A:2_1/B:3_0@3!1-1^a/C:2_1/D:1/E:3@2/F:2/G:10
 $\text{a}^{\text{m-a+t=p}}$ @2/A:2_1/B:3_0@3!1-1^a/C:2_1/D:1/E:3@2/F:2/G:10
 $\text{m}^{\text{a-t+p=a}}$ @3/A:2_1/B:3_0@3!1-1^a/C:2_1/D:1/E:3@2/F:2/G:10
 $\text{a}^{\text{t-p+a=g}}$ @1/A:3_0/B:2_1@1!2-0^a/C:2_0/D:3/E:2@3/F:0/G:10
 $\text{t}^{\text{p-a+g=i}}$ @2/A:3_0/B:2_1@1!2-0^a/C:2_0/D:3/E:2@3/F:0/G:10
 $\text{p}^{\text{a-g+i=pau}}$ @1/A:2_1/B:2_0@2!1-0^i/C:0_0/D:3/E:2@3/F:0/G:10
 $\text{a}^{\text{g-i+pau=a}}$ @2/A:2_1/B:2_0@2!1-0^i/C:0_0/D:3/E:2@3/F:0/G:10
 $\text{g}^{\text{i-pau+a=p}}$ @x/A:2_0/B:x_x@x!x-x^x/C:1_1/D:2/E:x@x/F:2/G:10
 $\text{i}^{\text{pau-a+p=a}}$ @1/A:0_0/B:1_1@1!0-2^a/C:2_0/D:0/E:2@4/F:2/G:10
 $\text{pau}^{\text{a-p+a=k}}$ @1/A:1_1/B:2_0@2!1-1^a/C:2_1/D:0/E:2@4/F:2/G:10
 $\text{a}^{\text{p-a+k=a}}$ @2/A:1_1/B:2_0@2!1-1^a/C:2_1/D:0/E:2@4/F:2/G:10
 $\text{p}^{\text{a-k+a=b}}$ @1/A:2_0/B:2_1@1!2-0^a/C:3_0/D:2/E:2@5/F:0/G:10
 $\text{a}^{\text{k-a+b=a}}$ @2/A:2_0/B:2_1@1!2-0^a/C:3_0/D:2/E:2@5/F:0/G:10
 $\text{k}^{\text{a-b+a=r}}$ @1/A:2_1/B:3_0@2!1-0^a/C:0_0/D:2/E:2@5/F:0/G:10
 $\text{a}^{\text{b-a+r=pau}}$ @2/A:2_1/B:3_0@2!1-0^a/C:0_0/D:2/E:2@5/F:0/G:10
 $\text{b}^{\text{a-r+pau=x}}$ @3/A:2_1/B:3_0@2!1-0^a/C:0_0/D:2/E:2@5/F:0/G:10
 $\text{a}^{\text{r-pau+x=x}}$ @x/A:3_0/B:x_x@x!x-x^x/C:0_0/D:2/E:x@x/F:0/G:10

Gambar IV-9 Keluaran label dengan set konteks 2 untuk kalimat "hai selamat pagi, apa kabar"

Tabel IV-7 Set konteks 3, kombinasi dari set pertama dan kedua

Set Konteks 3
Tingkat Fonem
Dua fonem sebelum
Fonem sebelum
Fonem saat ini
Fonem setelah
Dua fonem setelah
Posisi fonem pada kata (dari depan)
Posisi fonem pada kata (dari belakang)
Posisi fonem pada suku kata (dari depan)
Posisi fonem pada suku kata (dari belakang)
Tingkat Suku Kata
Jumlah fonem pada suku kata sebelum
Penekanan pada suku kata sebelum
Jumlah fonem pada suku kata saat ini
Penekanan pada suku kata saat ini
Posisi suku kata pada kata (dari depan)
Posisi suku kata pada kata (dari belakang)
Jarak dengan suku kata sebelumnya yang memiliki penekanan
Jarak dengan suku kata setelahnya yang memiliki penekanan
Huruf hidup pada suku kata ini
Jumlah fonem pada suku kata sesudah
Penekanan pada suku kata sesudah
Tingkat Kata
Jumlah fonem pada kata sebelum
Jumlah suku kata pada kata sebelum
Jumlah fonem pada kata saat ini
Jumlah suku kata pada kata saat ini
Posisi kata pada kalimat (dari depan)
Posisi kata pada kalimat (dari belakang)
Jumlah fonem pada kata sesudah
Jumlah suku kata pada kata sesudah
Tingkat Kalimat
Jumlah suku kata total
Jumlah kata total

```

x^x-pau+h=ay@x-x_x-x/A:0_0/B:x_x#x-x&x-x!x-x^x/C:2_0/D:0_0/E:x_x|x-x/F:2_1/G:10_5
x^pau-h+ay=s@1-2_1-2/A:0_0/B:2_0#1-1&1-10!0-2^ay/C:2_0/D:0_0/E:2_1|1-5/F:7_3/G:10_5
pau^h-ay+s=e@2-1_2-1/A:0_0/B:2_0#1-1&1-10!0-2^ay/C:2_0/D:0_0/E:2_1|1-5/F:7_3/G:10_5
h^ay-s+e=l@1-2_1-7/A:2_0/B:2_0#1-3&2-9!0-1^e/C:2_1/D:2_1/E:7_3|2-4/F:4_2/G:10_5
ay^s-e+l=a@2-1_2-6/A:2_0/B:2_0#1-3&2-9!0-1^e/C:2_1/D:2_1/E:7_3|2-4/F:4_2/G:10_5
s^e-l+a=m@1-2_3-5/A:2_0/B:2_1#2-2&3-8!0-2^a/C:3_0/D:2_1/E:7_3|2-4/F:4_2/G:10_5
e^l-a+m=a@2-1_4-4/A:2_0/B:2_1#2-2&3-8!0-2^a/C:3_0/D:2_1/E:7_3|2-4/F:4_2/G:10_5
l^a-m+a=t@1-3_5-3/A:2_1/B:3_0#3-1&4-7!1-1^a/C:2_1/D:2_1/E:7_3|2-4/F:4_2/G:10_5
a^m-a+t=p@2-2_6-2/A:2_1/B:3_0#3-1&4-7!1-1^a/C:2_1/D:2_1/E:7_3|2-4/F:4_2/G:10_5
m^a-t+p=a@3-1_7-1/A:2_1/B:3_0#3-1&4-7!1-1^a/C:2_1/D:2_1/E:7_3|2-4/F:4_2/G:10_5
a^t-p+a=g@1-2_1-4/A:3_0/B:2_1#1-2&5-6!2-0^a/C:2_0/D:7_3/E:4_2|3-3/F:0_0/G:10_5
t^p-a+g=i@2-1_2-3/A:3_0/B:2_1#1-2&5-6!2-0^a/C:2_0/D:7_3/E:4_2|3-3/F:0_0/G:10_5
p^a-g+i=pau@1-2_3-2/A:2_1/B:2_0#2-1&6-5!1-0^i/C:0_0/D:7_3/E:4_2|3-3/F:0_0/G:10_5
a^g-i+pau=a@2-1_4-1/A:2_1/B:2_0#2-1&6-5!1-0^i/C:0_0/D:7_3/E:4_2|3-3/F:0_0/G:10_5
g^i-pau+a=p@x-x_x-x/A:2_0/B:x_x#x-x&x-x!x-x^x/C:1_1/D:4_2/E:x_x|x-x/F:3_2/G:10_5
i^pau-a+p=a@1-1_1-3/A:0_0/B:1_1#1-2&7-4!0-2^a/C:2_0/D:0_0/E:3_2|4-2/F:5_2/G:10_5
pau^a-p+a=k@1-2_2-2/A:1_1/B:2_0#2-1&8-3!1-1^a/C:2_1/D:0_0/E:3_2|4-2/F:5_2/G:10_5
a^p-a+k=a@2-1_3-1/A:1_1/B:2_0#2-1&8-3!1-1^a/C:2_1/D:0_0/E:3_2|4-2/F:5_2/G:10_5
p^a-k+a=b@1-2_1-5/A:2_0/B:2_1#1-2&9-2!2-0^a/C:3_0/D:3_2/E:5_2|5-1/F:0_0/G:10_5
a^k-a+b=a@2-1_2-4/A:2_0/B:2_1#1-2&9-2!2-0^a/C:3_0/D:3_2/E:5_2|5-1/F:0_0/G:10_5
k^a-b+a=r@1-3_3-3/A:2_1/B:3_0#2-1&10-1!1-0^a/C:0_0/D:3_2/E:5_2|5-1/F:0_0/G:10_5
a^b-a+r=pau@2-2_4-2/A:2_1/B:3_0#2-1&10-1!1-0^a/C:0_0/D:3_2/E:5_2|5-1/F:0_0/G:10_5
b^a-r+pau=x@3-1_5-1/A:2_1/B:3_0#2-1&10-1!1-0^a/C:0_0/D:3_2/E:5_2|5-1/F:0_0/G:10_5
a^r-pau+x=x@x-x_x-x/A:3_0/B:x_x#x-x&x-x!x-x^x/C:0_0/D:5_2/E:x_x|x-x/F:0_0/G:10_5

```

Gambar IV-10 Keluaran label dengan set konteks 3 untuk kalimat "hai selamat pagi, apa kabar"

IV.4. Implementasi proses ekstraksi fitur ucapan

Setelah mendapatkan label kontekstual, yang harus disiapkan berikutnya adalah data yang berasal dari sinyal ucapan. Data tersebut adalah data berupa parameter spektral (karakteristik resonansi rongga suara/spektrum frekuensi) dan parameter F0 (frekuensi fundamental – digunakan nilai *log*-nya). Data ucapan yang berupa sinyal ucapan harus diekstraksi parameter-parameternya. Proses ekstraksi ini dilakukan dengan menggunakan beberapa kakas, yaitu *pda*, *SPTK*, dan *Tcl*. Pertama-tama, dilakukan terlebih dahulu penentuan batas atas dan batas bawah F0 untuk memberikan jarak (*range*) terhadap kakas *Tcl* dalam proses ekstraksi nilai *log* dari F0 nantinya. Dengan menggunakan *pda*, dan sebuah *script* dalam bahasa *C shell script*, dapat diekstrak nilai F0 untuk setiap data ucapan yang ada. Berikut adalah perintah ekstraksi F0 untuk satu *file*:

- `$pda <wav_data_input> -o <output_file> -fmax <upper_F0_value> -fmin <lower_F0_value>`

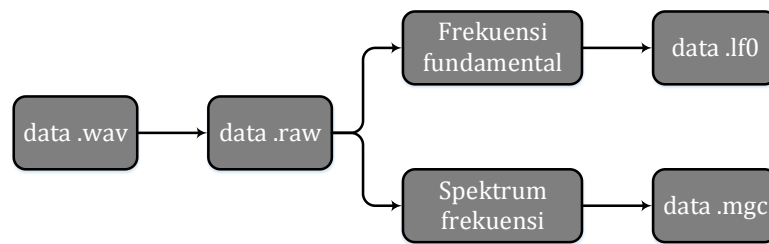
Perintah ini dijalankan dengan menggunakan kisaran awal F0 110-400 Hz dikarenakan pembicara pada data yang digunakan adalah wanita, dan ini adalah *range* suara wanita pada umumnya. Setelah semua data F0 didapatkan, dilakukan perhitungan histogram dan dilihat distribusinya. Nilai terkecil dari distribusi probabilitas frekuensi tersebut diambil sebagai *lower* F0, dan nilai terbesar sebagai *upper* F0. Pada tugas akhir ini, dengan melakukan langkah tersebut, didapat batas atas F0 adalah 400 Hz, dan batas bawahnya adalah 140Hz.

Untuk dapat mengekstrak fitur yang sesuai, data ucapan harus terlebih dahulu dikonversi ke dalam format “.raw”. Format data awal yang tersedia adalah “.wav”. Dengan menggunakan kaskas SoX, dapat dilakukan perubahan format yang sesuai. Perintah yang digunakan untuk proses ini adalah:

```
➤ $sox -c 1 -s -2 -t wav -r <sample_frequency> <wav_data> -c 1 -s -2 -t raw  
-r <sample_frequency> > <raw_data>
```

Pada tugas akhir ini, digunakan pengaturan 16kHz *sample frequency*, sesuai dengan data ucapan yang ada.

Proses ekstraksi parameter ucapan akhirnya dapat dilakukan dengan menggunakan *script* yang sudah tersedia pada *file demo* yang disediakan oleh kaskas HTS. *Script* ini pada intinya melakukan ekstraksi parameter spektral dari sinyal ucapan, dengan spesifikasi 24-dimensi *mel-cepstral analysis* (ditambah dimensi ke-0 sebagai fitur *power*). Proses ekstraksi juga dilakukan untuk mengambil nilai *log* dari F0 untuk setiap sinyal suara. Proses ekstraksi parameter spektral dilakukan dengan menggunakan kaskas SPTK, sedangkan untuk ekstraksi F0 digunakan SPTK beserta Tcl. Untuk proses ini, beberapa pengaturan lain juga diperlukan, seperti analisis sinyal dengan 5ms *frameshift* dan 25ms *Hamming window*. Berikut adalah tahapan dari implementasi proses ekstraksi fitur ucapan.



Gambar IV-11 Tahapan implementasi proses ekstraksi fitur ucapan

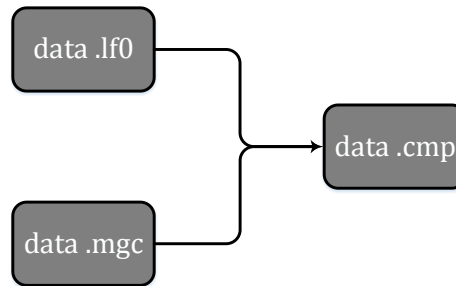
IV.5. Implementasi proses *training* HMM

Untuk membangun sistem TTS berbasis *hidden Markov model* (HMM), tentunya perlu dilakukan proses *training* terhadap data yang ada untuk didapatkan model akustik yang sesuai. Untuk melakukan proses *training* tersebut, diperlukan format data yang sesuai dengan kaskas yang digunakan, dalam hal ini adalah HTS. Format data tersebut terkait dengan tiga fitur sumber (*source feature*) yang digunakan, yaitu label kontekstual, durasi, frekuensi fundamental dan frekuensi spektral. Label kontekstual dan durasi digabungkan di dalam format teks hasil keluaran dari pemrosesan teks. Berikut adalah contoh label monofon dengan durasi yang sudah digabungkan dengan label *full-context* tanpa durasi.

	0	3340872	x^x-pau+h=ay@x_x/A:0/B:x@x-x/C:2/D:5
3340872	4143586	x^pau-h+ay=s@1_2/A:0/B:2@1-5/C:7/D:5	
4143586	6885249	pau^h-ay+s=e@2_1/A:0/B:2@1-5/C:7/D:5	
6885249	8863768	h^ay-s+e=l@1_7/A:2/B:7@2-4/C:4/D:5	
8863768	9208596	ay^s-e+l=a@2_6/A:2/B:7@2-4/C:4/D:5	
9208596	9592994	s^e-l+a=m@3_5/A:2/B:7@2-4/C:4/D:5	
9592994	10604866	e^l-a+m=a@4_4/A:2/B:7@2-4/C:4/D:5	
10604866	11339745	l^a-m+a=t@5_3/A:2/B:7@2-4/C:4/D:5	
11339745	12057665	a^m-a+t=p@6_2/A:2/B:7@2-4/C:4/D:5	
12057665	12775585	m^a-t+p=a@7_1/A:2/B:7@2-4/C:4/D:5	
12775585	13561340	a^t-p+a=g@1_4/A:7/B:4@3-3/C:0/D:5	
13561340	14658005	t^p-a+g=i@2_3/A:7/B:4@3-3/C:0/D:5	
14658005	15472024	p^a-g+i=pau@3_2/A:7/B:4@3-3/C:0/D:5	
15472024	17026576	a^g-i+pau=a@4_1/A:7/B:4@3-3/C:0/D:5	
17026576	18598086	g^i-pau+a=p@x_x/A:4/B:x@x-x/C:3/D:5	
18598086	19479940	i^pau-a+p=a@1_3/A:0/B:3@4-2/C:5/D:5	
19479940	20390059	pau^a-p+a=k@2_2/A:0/B:3@4-2/C:5/D:5	
20390059	21240050	a^p-a+k=a@3_1/A:0/B:3@4-2/C:5/D:5	
21240050	22249868	p^a-k+a=b@1_5/A:3/B:5@5-1/C:0/D:5	
22249868	23188251	a^k-a+b=a@2_4/A:3/B:5@5-1/C:0/D:5	
23188251	24126635	k^a-b+a=r@3_3/A:3/B:5@5-1/C:0/D:5	
24126635	25986444	a^b-a+r=pau@4_2/A:3/B:5@5-1/C:0/D:5	
25986444	27054844	b^a-r+pau=x@5_1/A:3/B:5@5-1/C:0/D:5	
27054844	28300000	a^r-pau+x=x@x_x/A:5/B:x@x-x/C:0/D:5	

Gambar IV-12 Label *full-context* dan durasinya untuk kalimat "hai selamat pagi, apa kabar" dengan set konteks 1

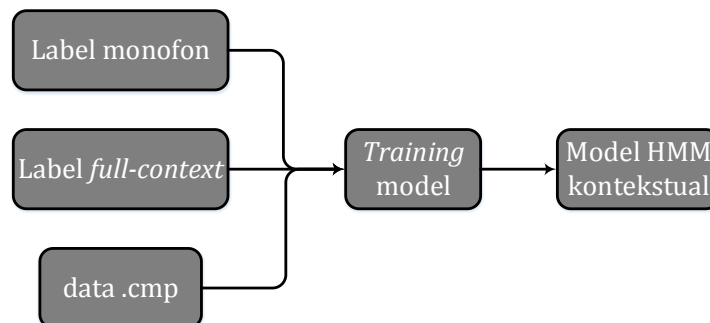
Selain itu, untuk kedua fitur lainnya yang menyangkut sinyal ucapan, yaitu frekuensi fundamental dan frekuensi spektral, digunakan data hasil ekstraksi pada proses ekstraksi fitur ucapan. Kedua fitur ini, yang masing-masing memiliki format '.lf0' dan '.mgc' digabungkan menjadi satu data dengan format '.cmp' dengan *script* yang disediakan oleh kakas HTS. Berikut adalah gambaran dari persiapan data fitur ucapannya.



Gambar IV-13 Penggabungan dua fitur ucapan

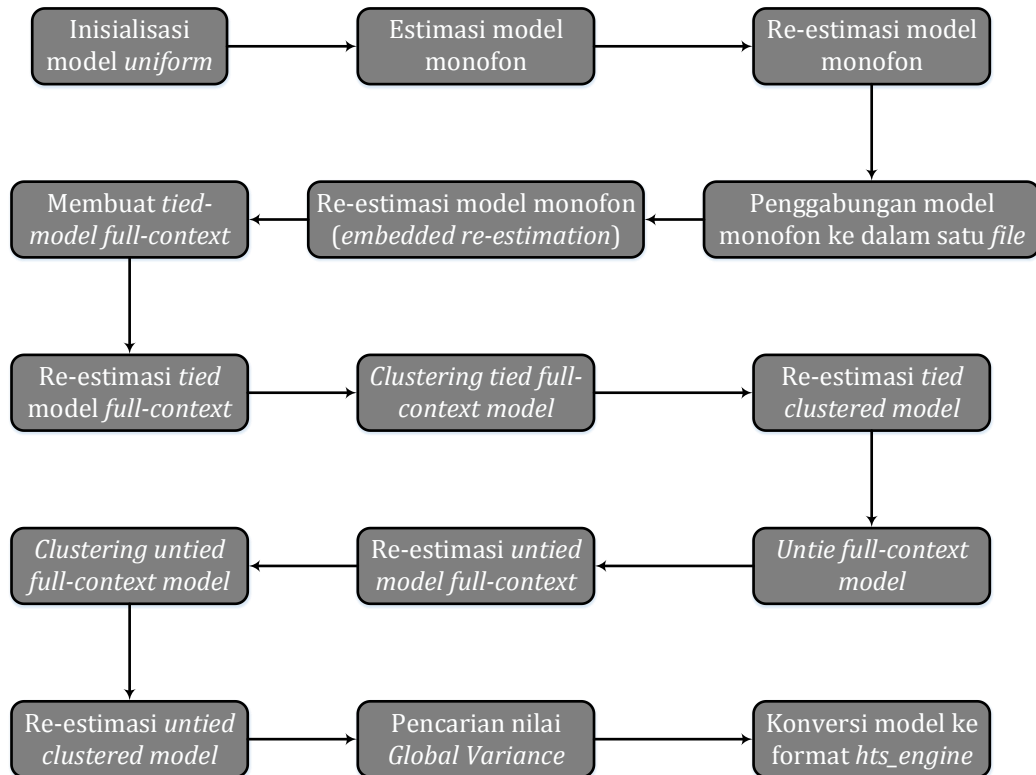
Sebagai gambaran, sistem TTS pada HTS ini dibangun dengan *script* berbahasa *Perl*. *Script* ini mengacu kepada demo TTS berbahasa Inggris yang disediakan. Untuk mengimplementasi sistem dalam bahasa Indonesia, perlu dilakukan perubahan konfigurasi terkait kondisi data ucapan, direktori penyimpanan, serta data yang digunakan pada proses *training* sendiri.

Setelah mendapatkan data-data yang diperlukan, yaitu label monofon, label *full-context*, data fitur ucapan (data '.cmp'), maka langkah berikutnya adalah melakukan proses *training* untuk membuat model akustik HMM. Berikut adalah gambaran singkat dari prosesnya.



Gambar IV-14 *Input-output* dari proses *training* model akustik HMM

Proses *training* HMM yang terdapat pada diagram tersebut sebenarnya terdiri dari banyak tahapan. Berikut adalah diagram dari tahapan proses *training* HMM.



Gambar IV-15 Tahapan dari proses *training* model pada HTS

Berikut dijelaskan mengenai perintah-perintah yang dilakukan pada setiap tahapan.

1. Inisialisasi nilai *mean* dan *variance* secara global dari data parameter ucapan untuk setiap fonem:
 - Sebagai nilai awal untuk estimasi parameter
 - Nilai *mean* dan *variance* awal dari parameter spektral dan F0 diinisialisasi secara global/*uniform* untuk setiap model fonem HMM
 - Nilai *mean* dan *variance* awal dari parameter durasi diinisialisasi secara global/*uniform* untuk setiap model fonem HMM
 - Perhitungan *mean* dan *variance* global untuk parameter spektral dan F0 dilakukan dengan perintah *HCompV* pada HTS

- Perhitungan *mean* dan *variance* global untuk parameter durasi dilakukan dengan menghitung rasio antara jumlah *frame* dengan perkalian antara jumlah fonem dan jumlah *state* HMM yang digunakan.
2. Estimasi nilai parameter untuk setiap fonem (model monofon):
 - Menggunakan algoritma Viterbi untuk estimasi nilai parameter spektral dan F0, untuk setiap fonem
 - Estimasi dilakukan terhadap model awal HMM hasil dari langkah sebelumnya, terhadap data *training* ucapan dengan menggunakan label mono yang telah diberikan durasi tiap fonemnya
 - Proses dilakukan dengan menjalankan perintah *HInit* pada HTS
 3. Re-estimasi nilai parameter untuk setiap fonem:
 - Menggunakan algoritma Baum-Welch untuk melakukan re-estimasi nilai parameter spektral, F0, dan juga durasi
 - Re-estimasi dilakukan untuk memperbaiki parameter spektral dan F0 dari model yang dihasilkan pada langkah kedua
 - Re-estimasi juga dilakukan untuk melakukan estimasi terhadap durasi tiap fonem dengan menggunakan probabilitas transisi *state* dari hasil langkah kedua sebagai model awal
 - Proses dilakukan dengan menjalankan perintah *HRest* pada HTS
 4. Menggabungkan semua model fonem HMM ke dalam satu *macro master file* (mmf):
 - Menggabungkan setiap model fonem HMM untuk parameter spektral dan F0 ke dalam satu *file*
 - Menggabungkan setiap model fonem HMM untuk parameter durasi ke dalam satu *file*
 - Menggunakan model HMM hasil re-estimasi pada langkah ke-3
 - Proses dilakukan dengan menjalankan perintah *HHed*
 5. Melakukan re-estimasi terhadap model monofon yang sudah terintegrasi:
 - Melakukan re-estimasi nilai parameter terhadap model gabungan hasil proses pada langkah ke-4 untuk parameter spektral dan F0

- Melakukan re-estimasi nilai parameter terhadap model gabungan hasil proses pada langkah ke-4 untuk parameter durasi
 - Re-estimasi dilakukan dengan algoritma Baum-Welch
 - Proses dilakukan dengan menjalankan perintah HERest
6. Membuat model fonem HMM yang *full-context* dari model yang *context-independent* hasil dari langkah sebelumnya (*tied full-context model*):
- Membuat model fonem HMM yang sesuai dengan konteks-konteks yang mungkin untuk setiap fonem tersebut
 - Setiap model HMM dengan fonem yang sama disatukan (*tied*) parameternya
 - Menggunakan perintah HHEd pada HTS
7. Melakukan re-estimasi terhadap *full-context*:
- Menggunakan algoritma Baum-Welch untuk melakukan re-estimasi terhadap *full-context* model hasil dari langkah 6
 - Estimasi dilakukan terhadap parameter spektral, F0 dan durasi tentunya
8. Melakukan *clustering* terhadap model *full-context* HMM dengan struktur data *binary decision tree*:
- *Clustering* dilakukan agar sistem dapat menemukan dengan tepat parameter yang sesuai dengan konteks dari fonem yang dicari
 - *Decision tree* dibangun dengan menggunakan pertanyaan yang berkaitan dengan fonetis, seperti cara pengucapan fonem, jenis fonem, ataupun mengenai posisi fonem. Daftar pertanyaan yang digunakan terdapat pada lampiran.
 - Penentuan pertanyaan yang digunakan pada suatu simpul dilakukan dengan mengambil pertanyaan yang memberikan probabilitas (*log likelihood*) tertinggi pada kedua anaknya.
 - Simpul yang dipilih untuk ditentukan anaknya adalah simpul yang memiliki *log likelihood* tertinggi dari kumpulan simpul yang sudah terbentuk

- Simpul daun adalah simpul yang memiliki probabilitas observasi untuk setiap parameter ucapan.
 - Masing-masing parameter spektral, F0 dan durasi memiliki *decision tree* yang berbeda
 - Pembentukan *decision tree* dilakukan dengan perintah HHed
9. Re-estimasi nilai parameter dari model HMM yang sudah di-*cluster* dengan *decision tree* pada langkah ke-8:
- Dilakukan dengan algoritma Baum-Welch
 - Ketiga parameter spektral, F0 dan durasi dire-estimasi ulang dengan menggunakan model HMM pada *decision tree* dari langkah 8
 - Perintah yang digunakan adalah HERest
10. Pemisahan (*untying*) parameter ucapan dari setiap fonem, sehingga fonem yang sama tetapi dengan konteks yang berbeda akan memiliki parameter yang berbeda
- Dilakukan terhadap model hasil re-estimasi pada langkah 9
 - Perintah dilakukan dengan HHed
11. Re-estimasi *untied* model, yaitu model yang memiliki parameter yang terpisah antar tiap fonem yang berbeda konteks:
- Dilakukan dengan menggunakan algoritma Baum-Welch
 - Perintahnya adalah HERest
12. Pembuatan struktur *decision tree* untuk membuat *cluster-cluster* parameter ucapan sesuai dengan jalur pertanyaan yang sesuai.
- Dilakukan terhadap *untied* model hasil langkah 11
 - Menggunakan perintah HHed
13. Melakukan re-estimasi terakhir terhadap model-model fonem HMM *full-context* yang telah di-*cluster* dengan *decision tree* (*untied clustered model*):
- Re-estimasi dilakukan dengan algoritma Baum-Welch
 - Dilakukan terhadap model hasil dari langkah 12
 - Perintah yang digunakan adalah HERest
14. Mencari nilai *global variance* dari parameter-parameter ucapan:

- Dilakukan untuk mengurangi efek *over-smoothing* dari proses pembangkitan parameter ucapan pada proses sintesis ucapan nantinya
- Menggunakan perintah HSMMAalign dan kaskas SPTK

15. Mengkonversi *macro master file* (mmf) yang berisi model-model fonem HMM ke dalam format yang dipahami HTS *engine*:

- Membuat *file* yang berisi distribusi probabilitas dari setiap parameter spektral, F0, durasi dan *global variance*
- Membuat *file* yang berisi struktur *decision tree* dari setiap parameter spektral, F0, durasi dan *global variance*
- Kedua jenis *file* (distribusi probabilitas dan *decision tree*) ini adalah model HMM akhir dari setiap fonem dengan konteks yang ditentukan pada proses *training* ini.

Pada Tugas Akhir ini, dibuat empat model akustik yang berbeda. Berikut adalah daftar model yang dihasilkan dan diuji kualitasnya.

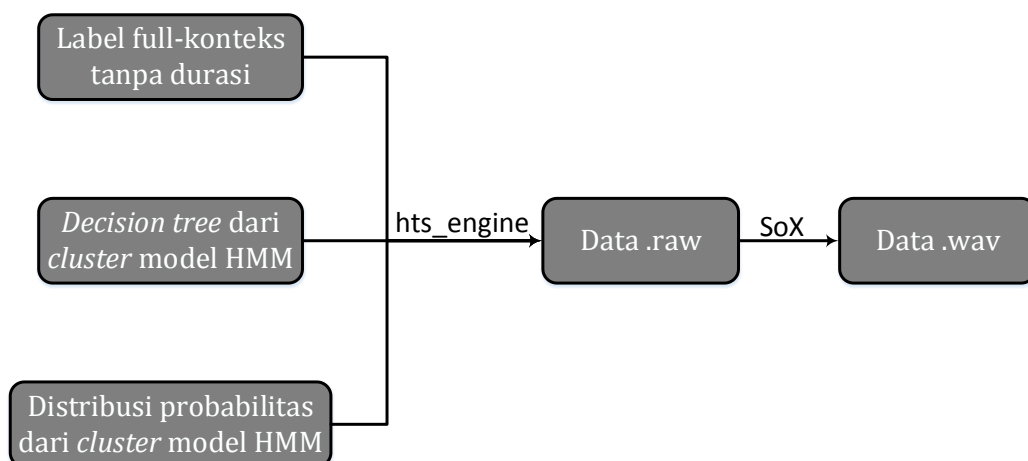
- a. Model dengan set konteks 1, (Sakti, 2008). Set pertanyaan (*question set*) untuk *decision tree* dari model ini terdapat pada lampiran A. Set pertanyaan awalnya diadaptasi dari bahasa Inggris ke bahasa Indonesia. Beberapa pertanyaan terkait fonetis sudah disesuaikan ke bahasa Indonesia. Namun, untuk pertanyaan terkait pemosisian fonem, suku kata, atau kata beserta jumlahnya masih mengikuti bahasa Inggris.
- b. Model dengan set konteks 1, (Sakti, 2008). Set pertanyaan (*question set*) untuk *decision tree* dari model ini terdapat pada lampiran B. Sama seperti set pertanyaan pada lampiran A, hanya saja pertanyaan terkait pemosisian dan jumlah dari komponen penyusun kalimat, seperti fonem, suku kata ataupun kata sudah disesuaikan dengan bahasa Indonesia. Hal ini dikarenakan dalam bahasa Indonesia, misalnya satu kata bisa memiliki sampai 20 fonem dibanding bahasa Inggris yang pada umumnya lebih sedikit, dsb.

- c. Model dengan set konteks 2, (Vania, 2011). Set pertanyaan (*question set*) untuk *decision tree* dari model ini terdapat pada lampiran C. Set pertanyaan pada model ini mengikuti dari lampiran B, dengan penyesuaian terhadap daftar konteks yang digunakan.
- d. Model dengan set konteks 3, kombinasi (Sakti, 2008) dan (Vania, 2011). Set pertanyaan (*question set*) untuk *decision tree* dari model ini terdapat pada lampiran D. Set pertanyaan pada model ini mengikuti dari lampiran C, dengan penyesuaian terhadap daftar konteks yang digunakan, yaitu kombinasi konteks antara (Sakti, 2008) dan (Vania, 2011).

IV.6. Implementasi proses *synthesis* ucapan

Pada tahapan ini, semua komponen yang diperlukan untuk pembangkitan ucapan dari teks sudah dibuat. Komponen-komponen tersebut adalah pemrosesan teks serta model HMM kontekstual. Model akustik HMM tersebut secara khusus terdiri dari dua jenis *file*, yaitu *decision tree* dari *cluster* model HMM dan distribusi probabilitas untuk setiap *cluster* model HMM. Komponen-komponen tersebut digunakan pada proses *synthesis* ini untuk membangkitkan ucapan dari teks masukan. Proses ini menggunakan kakas *hts_engine* dan SoX

Berikut adalah tahapan dari proses *synthesis* yang menggunakan pada kakas *hts_engine* dan SoX



Gambar IV-16 Tahapan proses sintesis ucapan dari teks

Kakas *hts_engine* berperan dalam penelusuran *decision tree* dari model sesuai label kontekstual yang diberikan sehingga mendapatkan *cluster* dari model HMM yang sesuai. Kemudian dari distribusi probabilitas yang dimiliki oleh *cluster* tersebut, dibangkitkan parameter-parameter ucapannya yang sesuai, yaitu frekuensi fundamental, spektrum frekuensi beserta durasinya, untuk disusun menjadi sinyal ucapan dengan format '.raw'. Agar data ucapan tersebut dapat didengar dan dimengerti, maka data '.raw' tersebut diubah ke dalam '.wav' dengan kakas SoX.

BAB V.

EVALUASI SISTEM SINTESIS UCAPAN

Untuk menilai kualitas ucapan sintetis dari sistem yang dibuat pada tugas akhir ini, diperlukan adanya evaluasi. Evaluasi ini dilakukan untuk menilai tidak hanya kualitas, tapi juga akurasi dari pengucapan yang dikeluarkan apakah sesuai dengan transkripsi aslinya. Dalam rangka menilai kedua hal tersebut, digunakan dua metode yang berbeda, yaitu *Mean Opinion Score* (MOS) dan *Semantically Unpredictable Score* (SUS).

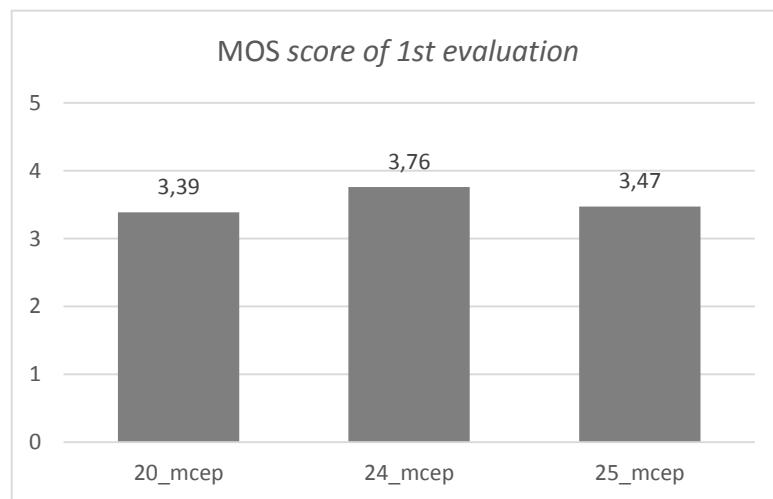
Mean Opinion Score (MOS) *test* adalah evaluasi yang dilakukan dengan memperdengarkan ucapan sintetis kepada beberapa pendengar yang ditentukan, dan meminta mereka menilai suara tersebut berdasarkan kejelasan (*clearness*) dan kealamian (*naturalness*). Nilai yang dapat diberikan memiliki *range* dari 1-5, dengan 1 berarti kualitas terburuk, dan 5 adalah kualitas terbaik dengan suara sangat jelas dan alami. Pada tes ini, digunakan 25 kalimat yang berbeda, yang dipaparkan kepada 6 orang pendengar.

Semantically Unpredictable Score (SUS) *test* adalah evaluasi yang dilakukan untuk menilai akurasi dari ucapan sintesis terhadap pengucapannya yang disesuaikan dengan teks masukan. Pada tes ini, digunakan 25 kalimat yang berbeda, yang juga dipaparkan kepada 6 orang pendengar. Setiap pendengar boleh mengulangi pengucapan suara tersebut sampai dua kali, sebelum menuliskan setiap kata yang diucapkan pada ucapan sintesis tersebut. Akurasi kata akan dihitung berdasar kata yang dapat dituliskan secara tepat. Akurasi kalimat dihitung dengan melihat apakah setiap kata pada kalimat dituliskan dengan benar semua. Hasil dari kedua metode evaluasi ini, MOS dan SUS *test* dipaparkan pada dua subbab berikut.

Untuk masing-masing dari kedua metode tersebut dilakukan dua kali evaluasi. Pada tes MOS, evaluasi pertama dilakukan untuk mencari koefisien *mel-cepstral* terbaik dari data, divariasikan antara 20, 24 dan 25. Pada evaluasi pertama

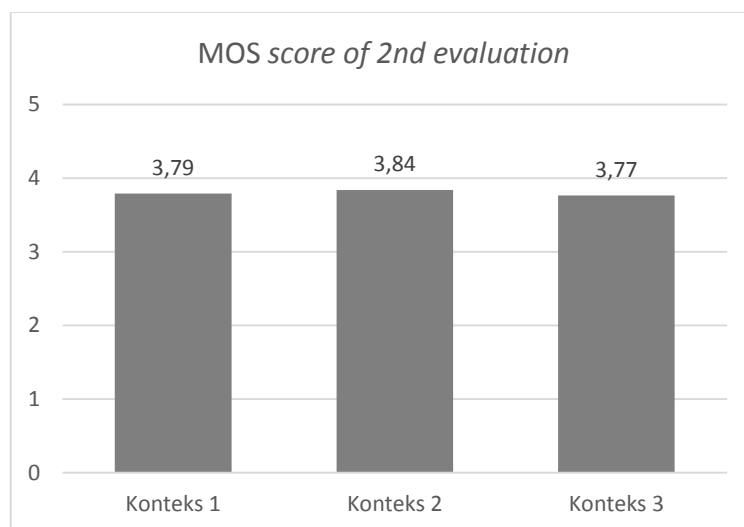
ini, digunakan set konteks 1 yang mengacu kepada (Sakti, 2008) dan *question set* yang dilampirkan pada lampiran A. Evaluasi kedua pada tes MOS dilakukan dengan membandingkan hasil ucapan sintetis dari tiga set konteks, yaitu set konteks 1 (Sakti, 2008), Vania (2011) dan kombinasinya. Pada evaluasi kedua, set konteks 1 menggunakan *question set* yang dilampirkan pada lampiran B, set konteks 2 menggunakan *question set* pada lampiran C dan set konteks 3 menggunakan *question set* pada lampiran D. Evaluasi pada tes SUS juga dilakukan dua kali. Evaluasi pertama dilakukan terhadap sistem yang menggunakan set konteks 1 dan *question set* pada lampiran A, serta koefisien *mel-cepstral* terbaik dari hasil tes MOS yang pertama. Evaluasi SUS yang kedua dilakukan terhadap sistem yang memiliki performansi terbaik pada tes MOS yang kedua.

V.1. Hasil tes *Mean Opinion Score* (MOS)



Gambar V-1 Hasil tes MOS dar evaluasi pertama dengan set konteks 1 terhadap tiga variasi koefisien *mel-cepstral*

Gambar di atas adalah hasil dari tes MOS pada evaluasi pertama. Dari hasil tersebut dapat dilihat bahwa sistem dengan 25 koefisien *mel-cepstral* menghasilkan performansi terbaik dengan nilai 3,76. Sedangkan sistem dengan koefisien 20 dan 3,47 memberikan nilai masing-masing 3,39 dan 3,47. Berdasar kepada hasil tersebut, maka sistem TTS pada Tugas Akhir ini dipastikan menggunakan koefisien 24 untuk analisis *mel-cepstral* pada proses ekstraksi parameter ucapan. Konfigurasi ini juga digunakan pada evaluasi-evaluasi yang berikutnya.



Gambar V-2 Hasil tes MOS dari evaluasi kedua untuk ketiga konteks dengan *update question* untuk decision tree

Pada tes MOS yang kedua, evaluasi diujikan terhadap tiga konfigurasi konteks dan *question set* yang berbeda. Konfigurasi pertama menggunakan set konteks 1 (sama seperti evaluasi pertama), namun menggunakan *question set* berbeda yang dilampirkan pada lampiran B. Konfigurasi kedua menggunakan set konteks 2 dan *question set* pada lampiran C. Konfigurasi ketiga menggunakan set konteks 3 dan *question set* pada lampiran D. Dari gambar tersebut, dapat dilihat bahwa konfigurasi kedua, yang menggunakan set konteks 2 (Vania, 2011) memberikan nilai tertinggi, yaitu 3,84. Yang menarik di sini adalah konfigurasi pertama memberikan nilai yang lebih tinggi dari hasil evaluasi pertama, yaitu 3,79 berbanding dengan 3,76. Hal ini menunjukkan bahwa pengembangan *question set* dapat meningkatkan kualitas pengucapan. Sedangkan kombinasi konteks yang diajukan pada Tugas Akhir ini memberikan nilai terendah, yaitu 3,77. Kalimat-kalimat yang digunakan pada tes MOS ini terlampir pada lampiran E.

Hal yang menarik dari evaluasi MOS ini adalah terdapatnya sedikit peningkatan kualitas dari sistem yang menggunakan *question set* yang berbeda. Jika dilihat pada lampiran, lampiran A menampilkan *question set* yang sebagian besar masih mengikuti dari *default question* pada bahasa Inggris, seperti jumlah fonem

pada kata maksimal kurang dari 7, dsb. Pada lampiran B, *question set* tersebut diperbarui dengan memperhatikan kemungkinan yang terjadi pada bahasa Indonesia, di mana satu kata dapat memiliki sampai 20 kata atau lebih. Ternyata, pertimbangan ini membuahkan hasil walau tidak begitu besar. Namun, ke depannya perlu diperhatikan lagi apa saja faktor krusial yang dapat menentukan dalam *question set* pada bahasa Indonesia.

V.2. Hasil tes *Semantically Unpredictable Sentence* (SUS)

Untuk tes *Semantically Unpredictable Sentence* (SUS), dipergunakan 25 kalimat yang berbeda. Kalimat-kalimat tersebut dibuat merunut kepada 5 struktur utama kalimat pada bahasa Indonesia, yaitu kalimat transitif, intransitif, pertanyaan, perintah, dan kalimat majemuk relatif. Masing-masing dari kelima struktur tersebut, dibuat 5 kalimat, dengan total menjadi 25 kalimat. Setiap kalimat mengandung kata-kata yang dibangkitkan secara acak, sesuai posisi jenis katanya. Dengan total pendengar adalah 6 orang, maka keseluruhan data uji adalah 150 kalimat. Dua puluh lima kalimat berbeda yang digunakan pada tes ini dilampirkan pada lampiran F.

Tabel V-1 Hasil tes SUS dari evaluasi pertama dengan set konteks 1

SUS score of 1st evaluation	
Total Sentences	Total Correct Sentences (%)
150	131 (87,33%)
Total Words	Total Correct Words (%)
630	606 (96,19%)

Tabel di atas menunjukkan hasil tes SUS pada evaluasi yang pertama. Pada evaluasi SUS yang pertama ini, sistem menggunakan set konteks 1 dengan *question set* dari lampiran A, dan konfigurasi koefisien *mel-cepstral* 24 sesuai dengan hasil MOS pada evaluasi pertama. Dari total 150 kalimat yang diperdengarkan kepada 6

pendengar, 131 di antaranya dituliskan dengan tepat seluruhnya, sehingga akurasi untuk kalimat mencapai 87,33%. Sedangkan untuk akurasi kata, pada evaluasi SUS yang pertama ini dicapai akurasi 96,19% dengan statistik total 606 kata dituliskan dengan benar dari 630 kata yang ada.

Dari 150 kalimat yang diperdengarkan pada evaluasi SUS yang pertama ini, 19 di antaranya ditulis/ditebak dengan tidak tepat. Kalimat-kalimat yang ditebak dengan tidak tepat itu dijabarkan dalam tabel di bawah ini.

Tabel V-2 Daftar kalimat yang salah tebak pada evaluasi SUS pertama

Kalimat	Ikan kaya melalaikan murid
	Ikan kaya melalai ikan murin
	Ikan kaya meleleh ikan murin
	Ikan kaya meleleh ikat murid
	Ikan kaya melele ikan murin
	Ikan kaya melele ikan murim
	Ikan kaya melele ikan murin
Kalimat	Lubang memonopoli niat yang mengompol
	Lubang memonopoli niat yang mengompong
Kalimat	Kepung benang dan ranjang
	Tepung benang dan ranjang
	Tepung benang dan ranjang
Kalimat	Bagaimana wibawa meyakinkan xenon

Bagaimana wibawa meyakinkan senon	
Bagaimana wibawa meyakinkan senon	
Kalimat	Paguyuban meringkus saham yang bertualang
Paguyuban meringkus sarang yang bertualang	
Kalimat	Baja pergi dari toilet
Bajak pergi dari toilet	
Bajak pergi dari toilet	
Kalimat	Ekologi memfokuskan fakultas yang bergegas
Ekologi memfokuskan fasilitas yang bergegas	
Ekologi memfokuskan fasilitas yang bergegas	
Kalimat	Visa akuatik mencolek bintang
Sisa akuatik mencolek bintang	
Kalimat	Jambak naga dengan protein
Jambak naga dengan protoin	
Kalimat	Kapan timah menguji visi
Kapan timah menguji fisi	

Dari tabel di atas dapat dilihat bahwa kesalahan sebagian besar terjadi pada kalimat “ikan kaya melalaikan murid”. Kesalahan terjadi pada kata “melalaikan” dan “murid”. Hal ini menunjukkan bahwa pengucapan fonem /ay/ dari gugus huruf “ai” masih kurang baik. Begitu pula pengucapan huruf ‘d’ yang berada di akhir kata. Selain itu, pengucapan huruf ‘k’ di awal kalimat tampak ambigu dengan ‘t’,

seperti yang terjadi dua kali pada kalimat “kepong benang dan ranjang”. Bunyi /a/ yang berada pada akhir kata, seperti pada kalimat “baja pergi dari toilet” juga dua kali salah tebak menjadi bunyi /a k/, sehingga kata “baja” menjadi “bajak”. Oleh karena itu, penanganan model ‘a’ di akhir kata juga perlu diperhatikan ke depannya. Pengucapan bunyi /ey/ juga perlu diperhatikan ke depannya. Bukan hanya karena terdapat satu kesalahan pada kalimat “jambak naga dengan protein”, tetapi juga perlu penanganan apakah gugus huruf “ei” dibaca sebagai diftong /ey/ atau sebagai fonem /e i/ yang terpisah. Selain itu, kesalahan tampak lebih kepada kalimat yang terlalu panjang seperti “ekologi memfokuskan fakultas yang bergegas” atau kata yang sulit/ambigu, seperti “xenon” atau “visi”.

Tabel V-3 Hasil tes SUS dari evaluasi kedua untuk ketiga konteks dengan *update question* untuk decision tree

SUS score of 2nd evaluation	
Total Sentences	Total Correct Sentences (%)
150	139 (92,67%)
Total Words	Total Correct Words (%)
630	619 (98,25%)

Setelah melihat hasil dari evaluasi SUS yang pertama, hasil evaluasi SUS yang kedua dapat dilihat dari tabel di atas. Pada evaluasi SUS yang kedua ini, sistem yang digunakan menerapkan set konteks 2 dengan *question set* pada lampiran C karena konfigurasi ini memiliki nilai tertinggi pada evaluasi MOS yang kedua. Dari total 150 kalimat, akurasi 92,67% diperoleh dengan 139 di antaranya ditebak secara benar. Untuk kata, akurasi berada pada angka 98,25%, dengan total 619 kata yang benar dari 630 kata. Akurasi pada tingkat kalimat memiliki kenaikan 5,34%. Sedangkan akurasi pada tingkat kata memiliki kenaikan 2,6% dari hasil SUS yang pertama. Ini menunjukkan bahwa set konteks 2 memang memberikan hasil yang lebih baik dalam akurasi pengucapan dari set konteks 1.

Berikut adalah daftar kalimat yang salah ditebak pada tes SUS yang kedua.

Tabel V-4 Daftar kalimat yang salah tebak pada evaluasi SUS kedua

Kalimat	Bagaimana wibawa meyakinkan xenon
	Bagaimana wibawa meyakinkan senon
	Bagaimana wibawa meyakinkan senon
	Bagaimana wibawa meyakinkan senot
Kalimat	Gula berjalan di gambut
	Bola berjalan di gambut
	Bula berjalan di gambut
Kalimat	Ekologi memfokuskan fakultas yang bergegas
	Ekologi memfokuskan fakultas yang berbebas
Kalimat	Cari zakat tanpa debu
	Cari zakat tanpa debo
Kalimat	Kuliah pengap menudungi hakim
	Kuliah tengah menudungi hakim
Kalimat	Kapan timah menguji visi
	Kapan timah menguji fisik
	Kapan timah menguji fisik
	Kapan timah menguji fisik

Dari hasil evaluasi SUS yang kedua ini dapat dilihat bahwa akurasi pengucapan sudah lebih baik. Kesalahan terutama diakibatkan oleh belum adanya penanganan perbedaan bunyi huruf ‘e’ pada bahasa Indonesia. Hal ini mengakibatkan terjadinya salah ucap bunyi /e/ pada suatu kata. Seperti pada kalimat “ekologi memfokuskan fakultas yang bergegas”, bunyi huruf ‘e’ pada sub-kata “gegas” dibunyikan seperti bunyi huruf ‘e’ pada kata “bebas”. Begitu pula bunyi huruf ‘e’ yang tidak tepat pada kalimat “cari zakat tanpa debu”. Pengucapan bunyi dari huruf ‘b’ dan ‘g’ juga masih kurang baik, seperti terlihat bahwa terdapat tiga kesalahan tebak kata “gula” pada kalimat “gula berjalan di gambut”. Pada kalimat “kapan timah menguji visi” memang terdapat kesalahan tebak akhiran kata “visi” menjadi bunyi /i k/. Penanganan bunyi huruf ‘i’ di akhir ini juga perlu diperhatikan ke depannya, walau dapat diasumsikan juga kata “visi” tidak terlalu intuitif.

BAB VI.

PENUTUP

Dengan menyelesaikan rangkaian pengerjaan Tugas Akhir ini, dapat ditarik beberapa kesimpulan dan saran yang sesuai dengan tujuan serta harapan dari Tugas Akhir ini ke depannya, sebagai berikut.

VI.1. Kesimpulan

1. Penggunaan set konteks (Vania, 2011) memberikan hasil yang terbaik
2. Penggunaan kombinasi set konteks (Sakti, 2008) dan (Vania, 2011) tidak meningkatkan performansi sistem
3. Perubahan pertanyaan yang digunakan dalam pembuatan *decision tree* mempengaruhi kualitas dari model akustik yang dibentuk
4. Kualitas dan akurasi dari pengucapan sistem HMM yang dikembangkan pada Tugas Akhir ini cukup baik.

VI.2. Saran

1. Meningkatkan tingkatan penggunaan konteks, misalnya tingkat frase, dan penambahan fitur POS *tagging* dalam set konteksnya
2. Melakukan pengembangan terhadap perbedaan dua bunyi huruf 'e' pada bahasa Indonesia.
3. Menambah jumlah data *training* untuk meningkatkan kualitas ucapan sintetis
4. Mengembangkan pensintesis ucapan yang bervariasi berdasarkan emosi

DAFTAR REFERENSI

- Sakti, S., dkk. (2008). Development of HMM-based Indonesian Speech Synthesis. Prosiding *O-COCOSDA* 215-200, Osaka, Japan
- Vania, C. & Adriani, M. (2011) The Effect of Syllable and Word Stress on the Quality of Indonesian HMM-based Speech Synthesis System. Prosiding *ICACSIS*, 413-417
- Campbell, N. & Black, A. (1995). Prosody and the selection of source units for concatenative synthesis. *Progress in Speech Synthesis*, 279-282.
- Arman, A.A. (2001). Prosody model for Indonesian text to speech system. Prosiding *Asia Pacific Conference on Communication*. Tokyo, Japan.
- Tokuda, K., Kobayashi, T., Imai, S. (1995). Speech parameter generation from HMM using dynamic features. Prosiding *ICASSP*, 660-663.
- Zen, H. & Tokuda, H. (2009). TechWare: HMM-based speech synthesis resources. *IEEE Signal Processing Magazine*.
- Yoshimura, T., dkk. (1999). Simultaneous modelling of spectrum, pitch, and duration in HMM-based speech synthesis. Prosiding *Eurospeech*, 2347-2350.
- Tokuda, K., Zen, H. & Black, A.W. (2000). An HMM-based speech synthesis system applied to English. Prosiding *IEEE Speech Synthesis Workshop*. Santa Monica, California, USA.
- Maia, R., dkk. (2006). An HMM-based Brazilian Portuguese speech synthesizer and its characteristics. *IEEE Journal of Communication and Information Systems*, vol. 21, no. 2, pp. 58-71.
- Yamagishi, J. (2006). *An Introduction to HMM-Based Speech Synthesis*. Technical Report, Tokyo Institute of Technology, Tokyo.
- Masuko, T. (2002). *HMM-Based Speech Synthesis and Its Applications*. Doctoral thesis, Tokyo Institute of Technology, Tokyo.
- Zen, H., dkk. (2007). The HMM-based Speech Synthesis System (HTS) Version 2.0. Prosiding *6th ISCA Workshop on Speech Synthesis*, Bonn, Germany
- Yamagishi, J., dkk. (2008). *Hidden Markov Model-based Speech Synthesis*. Lecture slide, University of Edinburgh, UK: Centre for Speech Technology Research.
- Zen, H. (2010). Fundamentals & Recent Advances in HMM-based Speech Synthesis. Prosiding *FALA2010*, Vigo, Spain.
- Zen, H., Tokuda, K., Black, A. W. (2007). Statistical Parametric Speech Synthesis. Prosiding *ICASSP IV*, 1229-1232.

- Alwi, H. (2003). *Tata Bahasa Baku Bahasa Indonesia Edisi Ketiga*. Jakarta: Balai Pustaka.
- Basuki, T.A. (2000). Pengenalan Suku Kata Bahasa Indonesia Menggunakan Finite-State Automata. *Integral*, volume 5 no. 2.
- Black, A.W. & Lenzo, K. A. (2003). *Building Synthetic Voices*.
- Meyer, R. (2006). *Source Excitation Generation for a HMM-Based Synthesizer*. Diploma Thesis, Swiss Federal Institute of Technology, Zurich.
- Yamagishi, J., dkk. (2008). Performance Evaluation of the Speaker-Independent HMM-based Speech Synthesis System “HTS-2007” for the Blizzard Challenge 2007. Prosiding ICASSP, 3957-3960, Las Vegas, Nevada, USA.
- Sakti, S., dkk. (2010). Quality and Inteligibility Assessment of Indonesian HMM-based Speech Synthesis System. Prosiding *MALINDO Workshop*, Jakarta, Indonesia
- Kawahara, H., Masuda-Katsuse, I. and Cheveigne, A. (1999). Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based f0 extraction: possible role of a repetitive structure in sounds. *Speech Communication*, vol. 27, pp. 187–207.
- Tokuda, K., dkk. (2011). *HTS Slides*. Nagoya Institute of Technology, Department of Computer Science.
- Young, S., dkk. (2009). *The Hidden Markov Model Toolkit (HTK) Book (for HTK version 3.4)*. Cambridge University, Engineering Department.
- Yoshimura, T., dkk. (1999). Simultaneous modelling of spectrum, pitch, and duration in HMM-based speech synthesis. Prosiding *EUROSPEECH*, vol. 5, pp. 1315-1318.

LAMPIRAN A. Question Set 1 untuk Konteks (Sakti, 2008)

A.I. Bagian set pertanyaan terkait konteks durasi

QS "LL-Vowel"	{a [*] ,aw [*] ,ay [*] ,e [*] ,ey [*] ,i [*] ,o [*] ,oy [*] ,u [*] }	QS "LL-Not_Alveolar"	{b [*] ,c [*] ,f [*] ,g [*] ,h [*] ,j [*] ,k [*] ,kh [*] ,m [*] ,ng [*] ,ny [*] ,p [*] ,s [*] ,sy [*] ,w [*] ,y [*] }
QS "LL-Consonant"	{b [*] ,c [*] ,d [*] ,f [*] ,g [*] ,h [*] ,j [*] ,k [*] ,kh [*] ,l [*] ,m [*] ,n [*] ,ng [*] ,ny [*] ,p [*] ,r [*] ,s [*] ,sy [*] ,t [*] ,w [*] ,y [*] ,z [*] }	QS "LL-Palatal_Consonant"	{c [*] ,j [*] ,ny [*] ,sy [*] ,y [*] }
QS "LL-Short_Vowel"	{a [*] ,e [*] ,i [*] ,o [*] ,u [*] }	QS "LL-Not_Palatal"	{b [*] ,d [*] ,f [*] ,g [*] ,h [*] ,k [*] ,kh [*] ,l [*] ,m [*] ,n [*] ,ng [*] ,p [*] ,r [*] ,s [*] ,t [*] ,w [*] ,z [*] }
QS "LL-Diphthong_Vowel"	{aw [*] ,ay [*] ,ey [*] ,oy [*] }	QS "LL-Velar_Consonant"	{g [*] ,k [*] ,kh [*] }
QS "LL-High_Vowel"	{i [*] ,u [*] }	QS "LL-Not_Velar"	{b [*] ,c [*] ,d [*] ,f [*] ,h [*] ,j [*] ,l [*] ,m [*] ,n [*] ,ny [*] ,p [*] ,r [*] ,s [*] ,sy [*] ,t [*] ,w [*] ,y [*] ,z [*] }
QS "LL-Mid_Vowel"	{e [*] ,ey [*] ,o [*] }	QS "LL-Glottal_Consonant"	{h [*] }
QS "LL-Low_Vowel"	{a [*] ,aw [*] ,ay [*] ,oy [*] }	QS "LL-Voiced_Consonant"	{b [*] ,d [*] ,g [*] ,j [*] ,w [*] ,z [*] }
QS "LL-Front_Vowel"	{a [*] ,e [*] ,ey [*] ,i [*] }	QS "LL-Unvoiced_Consonant"	{c [*] ,f [*] ,h [*] ,k [*] ,kh [*] ,l [*] ,m [*] ,n [*] ,ng [*] ,ny [*] ,p [*] ,r [*] ,s [*] ,sy [*] ,t [*] ,y [*] }
QS "LL-Central_Vowel"	{aw [*] ,ay [*] }	QS "LL-Plosive_Alveolar"	{d [*] ,t [*] }
QS "LL-Back_Vowel"	{o [*] ,oy [*] ,u [*] }	QS "LL-Plosive_Bilabial"	{b [*] ,p [*] }
QS "LL-Rounded_Vowel"	{o [*] ,oy [*] ,u [*] }	QS "LL-Plosive_Velar"	{g [*] ,k [*] }
QS "LL-Unrounded_Vowel"	{a [*] ,aw [*] ,ay [*] ,e [*] ,ey [*] ,i [*] }	QS "LL-Affricate_Palatal"	{c [*] ,j [*] }
QS "LL-AVowel"	{a [*] ,aw [*] ,ay [*] }	QS "LL-Fricative_Alveolar"	{s [*] ,z [*] }
QS "LL-IVowel"	{i [*] }	QS "LL-Plosive_Voiced"	{b [*] ,d [*] ,g [*] }
QS "LL-UVowel"	{u [*] }	QS "LL-Plosive_Unvoiced"	{k [*] ,p [*] ,t [*] }
QS "LL-EVowel"	{e [*] ,ey [*] }	QS "LL-Bilabial_Voiced"	{b [*] ,w [*] }
QS "LL-OVowel"	{o [*] ,oy [*] }	QS "LL-Bilabial_Unvoiced"	{m [*] ,p [*] }
QS "LL-Plosive_Consonant"	{b [*] ,d [*] ,g [*] ,k [*] ,p [*] ,t [*] }	QS "LL-Alveolar_Voiced"	{d [*] ,z [*] }
QS "LL-Not_Plosive"	{c [*] ,f [*] ,h [*] ,j [*] ,kh [*] ,l [*] ,m [*] ,n [*] ,ng [*] ,ny [*] ,r [*] ,s [*] ,sy [*] ,w [*] ,y [*] ,z [*] }	QS "LL-Alveolar_Unvoiced"	{l [*] ,n [*] ,r [*] ,s [*] ,t [*] }
QS "LL-Affricate_Consonant"	{c [*] ,j [*] }	QS "LL-silences"	{pau [*] }
QS "LL-Fricative_Consonant"	{f [*] ,h [*] ,kh [*] ,s [*] ,sy [*] ,z [*] }	QS "LL-a"	{a [*] }
QS "LL-Not_Fricative"	{b [*] ,c [*] ,d [*] ,g [*] ,j [*] ,k [*] ,l [*] ,m [*] ,n [*] ,ng [*] ,ny [*] ,p [*] ,r [*] ,s [*] ,t [*] ,w [*] ,y [*] }	QS "LL-aw"	{aw [*] }
QS "LL-Nasal_Consonant"	{m [*] ,n [*] ,ng [*] ,ny [*] }	QS "LL-ay"	{ay [*] }
QS "LL-Not_Nasal"	{b [*] ,c [*] ,d [*] ,f [*] ,g [*] ,h [*] ,j [*] ,k [*] ,kh [*] ,l [*] ,p [*] ,r [*] ,s [*] ,sy [*] ,t [*] ,w [*] ,y [*] ,z [*] }	QS "LL-b"	{b [*] }
QS "LL-Trill_Consonant"	{r [*] }	QS "LL-c"	{c [*] }
QS "LL-Lateral_Consonant"	{l [*] }	QS "LL-d"	{d [*] }
QS "LL-SemiVowel_Consonant"	{w [*] ,y [*] }	QS "LL-e"	{e [*] }
QS "LL-Bilabial_Consonant"	{b [*] ,m [*] ,p [*] ,w [*] }	QS "LL-ey"	{ey [*] }
QS "LL-Not_Bilabial"	{c [*] ,d [*] ,f [*] ,g [*] ,h [*] ,j [*] ,k [*] ,kh [*] ,l [*] ,n [*] ,ng [*] ,ny [*] ,r [*] ,s [*] ,sy [*] ,t [*] ,y [*] ,z [*] }	QS "LL-f"	{f [*] }
QS "LL-Labiodental_Consonant"	{f [*] }	QS "LL-g"	{g [*] }
QS "LL-Alveolar_Consonant"	{d [*] ,l [*] ,n [*] ,r [*] ,s [*] ,t [*] ,z [*] }	QS "LL-h"	{h [*] }
		QS "LL-i"	{i [*] }
		QS "LL-j"	{j [*] }
		QS "LL-k"	{k [*] }
		QS "LL-kh"	{kh [*] }
		QS "LL-l"	{l [*] }
		QS "LL-m"	{m [*] }
		QS "LL-n"	{n [*] }
		QS "LL-ng"	{ng [*] }
		QS "LL-ny"	{ny [*] }
		QS "LL-o"	{o [*] }
		QS "LL-oy"	{oy [*] }
		QS "LL-p"	{p [*] }

QS "LL-r"	{r^*}	QS "L-SemiVowel_Consonant"	{*^w-*,*^y-*
QS "LL-s"	{s^*}	QS "L-Bilabial_Consonant"	{*^b-*,*^m-*,*^p-*,*^w-*
QS "LL-sy"	{sy^*}	QS "L-Not_Bilabial"	{*^c-*,*^d-*,*^f-*,*^g-*,*^h-*,*^j-*,*^k-*,*^kh-*,*^l-*,*^n-*,*^ng-*,*^ny-*,*^r-*,*^s-*,*^sy-*,*^t-*,*^y-*,*^z-*
QS "LL-t"	{t^*}	QS "L-Labiodental_Consonant"	{*^f-*
QS "LL-u"	{u^*}	QS "L-Alveolar_Consonant"	{*^d-*,*^l-*,*^n-*,*^r-*,*^s-*,*^t-*,*^z-*
QS "LL-w"	{w^*}	QS "L-Not_Alveolar"	{*^b-*,*^c-*,*^f-*,*^g-*,*^h-*,*^j-*,*^k-*,*^kh-*,*^m-*,*^ng-*,*^ny-*,*^p-*,*^sy-*,*^w-*,*^y-*
QS "LL-y"	{y^*}	QS "L-Palatal_Consonant"	{*^c-*,*^j-*,*^ny-*,*^sy-*,*^y-*
QS "LL-z"	{z^*}	QS "L-Not_Palatal"	{*^b-*,*^d-*,*^f-*,*^g-*,*^h-*,*^k-*,*^kh-*,*^l-*,*^m-*,*^n-*,*^ng-*,*^p-*,*^r-*,*^s-*,*^t-*,*^w-*,*^z-*
QS "LL-pau"	{pau^*}	QS "L-Velar_Consonant"	{*^g-*,*^k-*,*^kh-*,*^ng-*
QS "L-Vowel"	{*^a-*,*^aw-*,*^ay-*,*^e-*,*^ey-*,*^i-*,*^o-*,*^oy-*,*^u-*		
QS "L-Consonant"	{*^b-*,*^c-*,*^d-*,*^f-*,*^g-*,*^h-*,*^j-*,*^k-*,*^kh-*,*^l-*,*^m-*,*^n-*,*^ng-*,*^ny-*,*^p-*,*^r-*,*^s-*,*^sy-*,*^t-*,*^w-*,*^y-*,*^z-*		
QS "L-Short_Vowel"	{*^a-*,*^e-*,*^i-*,*^o-*,*^u-*		
QS "L-Diphthong_Vowel"	{*^aw-*,*^ay-*,*^ey-*,*^oy-*		
QS "L-High_Vowel"	{*^i-*,*^u-*		
QS "L-Mid_Vowel"	{*^e-*,*^ey-*,*^o-*		
QS "L-Low_Vowel"	{*^a-*,*^aw-*,*^ay-*,*^oy-*		
QS "L-Front_Vowel"	{*^a-*,*^e-*,*^ey-*,*^i-*		
QS "L-Central_Vowel"	{*^aw-*,*^ay-*		
QS "L-Back_Vowel"	{*^o-*,*^oy-*,*^u-*		
QS "L-Rounded_Vowel"	{*^o-*,*^oy-*,*^u-*		
QS "L-Unrounded_Vowel"	{*^a-*,*^aw-*,*^ay-*,*^e-*,*^ey-*,*^i-*		
QS "L-AVowel"	{*^a-*,*^aw-*,*^ay-*		
QS "L-IVowel"	{*^i-*		
QS "L-UVowel"	{*^u-*		
QS "L-EVowel"	{*^e-*,*^ey-*		
QS "L-OVowel"	{*^o-*,*^oy-*		
QS "L-Plosive_Consonant"	{*^b-*,*^d-*,*^g-*,*^k-*,*^p-*,*^t-*		
QS "L-Not_Plosive"	{*^c-*,*^f-*,*^h-*,*^j-*,*^kh-*,*^l-*,*^m-*,*^n-*,*^ng-*,*^ny-*,*^r-*,*^s-*,*^sy-*,*^w-*,*^y-*,*^z-*		
QS "L-Affricate_Consonant"	{*^c-*,*^j-*		
QS "L-Fricative_Consonant"	{*^f-*,*^h-*,*^kh-*,*^s-*,*^sy-*,*^z-*		
QS "L-Not_Fricative"	{*^b-*,*^c-*,*^d-*,*^g-*,*^j-*,*^k-*,*^l-*,*^m-*,*^n-*,*^ng-*,*^ny-*,*^p-*,*^r-*,*^t-*,*^w-*,*^y-*		
QS "L-Nasal_Consonant"	{*^m-*,*^n-*,*^ng-*,*^ny-*		
QS "L-Not_Nasal"	{*^b-*,*^c-*,*^d-*,*^f-*,*^g-*,*^h-*,*^j-*,*^k-*,*^kh-*,*^l-*,*^p-*,*^r-*,*^s-*,*^sy-*,*^t-*,*^w-*,*^y-*,*^z-*		
QS "L-Trill_Consonant"	{*^r-*		
QS "L-Lateral_Consonant"	{*^l-*		
QS "L-Silences"	{*^pau-*		
QS "L-a"	{*^a-*		
QS "L-aw"	{*^aw-*		
QS "L-ay"	{*^ay-*		
QS "L-b"	{*^b-*		
QS "L-c"	{*^c-*		
QS "L-d"	{*^d-*		
QS "L-e"	{*^e-*		
QS "L-ey"	{*^ey-*		
QS "L-f"	{*^f-*		
QS "L-g"	{*^g-*		
QS "L-h"	{*^h-*		
QS "L-i"	{*^i-*		
QS "L-j"	{*^j-*		
QS "L-k"	{*^k-*		
QS "L-kh"	{*^kh-*		
QS "L-l"	{*^l-*		

QS "L-m" {^m-*}
 QS "L-n" {^n-*}
 QS "L-ng" {^ng-*}
 QS "L-ny" {^ny-*}
 QS "L-o" {^o-*}
 QS "L-oy" {^oy-*}
 QS "L-p" {^p-*}
 QS "L-r" {^r-*}
 QS "L-s" {^s-*}
 QS "L-sy" {^sy-*}
 QS "L-t" {^t-*}
 QS "L-u" {^u-*}
 QS "L-w" {^w-*}
 QS "L-y" {^y-*}
 QS "L-z" {^z-*}
 QS "L-pau" {^pau-*}

QS "C-Vowel" {*-a+*, *-aw+*, *-ay+*, *-e+*, *-ey+*, *-i+*, *-o+*, *-oy+*, *-u+*}
 QS "C-Consonant" {*-b+*, *-c+*, *-d+*, *-f+*, *-g+*, *-h+*, *-j+*, *-k+*, *-kh+*, *-l+*, *-m+*, *-n+*, *-ng+*, *-ny+*, *-p+*, *-r+*, *-s+*, *-sy+*, *-t+*, *-w+*, *-y+*, *-z+*}
 QS "C-Short_Vowel" {*-a+*, *-e+*, *-i+*, *-o+*, *-u+*}
 QS "C-Diphthong_Vowel" {*-aw+*, *-ay+*, *-ey+*, *-oy+*}
 QS "C-High_Vowel" {*-i+*, *-u+*}
 QS "C-Mid_Vowel" {*-e+*, *-ey+*, *-o+*}
 QS "C-Low_Vowel" {*-a+*, *-aw+*, *-ay+*, *-oy+*}
 QS "C-Front_Vowel" {*-a+*, *-e+*, *-ey+*, *-i+*}
 QS "C-Central_Vowel" {*-aw+*, *-ay+*}
 QS "C-Back_Vowel" {*-o+*, *-oy+*, *-u+*}
 QS "C-Rounded_Vowel" {*-o+*, *-oy+*, *-u+*}
 QS "C-Unrounded_Vowel" {*-a+*, *-aw+*, *-ay+*, *-e+*, *-ey+*, *-i+*}
 QS "C-AVowel" {*-a+*, *-aw+*, *-ay+*}
 QS "C-IVowel" {*-i+*}
 QS "C-UVowel" {*-u+*}
 QS "C-EVowel" {*-e+*, *-ey+*}
 QS "C-OVowel" {*-o+*, *-oy+*}
 QS "C-Plosive_Consonant" {*-b+*, *-d+*, *-g+*, *-k+*, *-p+*, *-t+*}
 QS "C-Not_Plosive" {*-c+*, *-f+*, *-h+*, *-j+*, *-kh+*, *-l+*, *-m+*, *-n+*, *-ng+*, *-ny+*, *-r+*, *-s+*, *-sy+*, *-w+*, *-y+*, *-z+*}
 QS "C-Affricate_Consonant" {*-c+*, *-j+*}
 QS "C-Fricative_Consonant" {*-f+*, *-h+*, *-kh+*, *-s+*, *-sy+*, *-z+*}
 QS "C-Not_Fricative" {*-b+*, *-c+*, *-d+*, *-g+*, *-j+*, *-k+*, *-l+*, *-m+*, *-n+*, *-ng+*, *-ny+*, *-p+*, *-r+*, *-t+*, *-w+*, *-y+*}

QS "C-Nasal_Consonant" {*-m+*, *-n+*, *-ng+*, *-ny+*}
 QS "C-Not_Nasal" {*-b+*, *-c+*, *-d+*, *-f+*, *-g+*, *-h+*, *-j+*, *-k+*, *-kh+*, *-l+*, *-p+*, *-r+*, *-s+*, *-sy+*, *-t+*, *-w+*, *-y+*, *-z+*}
 QS "C-Trill_Consonant" {*-r+*}
 QS "C-Lateral_Consonant" {*-l+*}
 QS "C-SemiVowel_Consonant" {*-w+*, *-y+*}
 QS "C-Bilabial_Consonant" {*-b+*, *-m+*, *-p+*, *-w+*}
 QS "C-Not_Bilabial" {*-c+*, *-d+*, *-f+*, *-g+*, *-h+*, *-j+*, *-k+*, *-kh+*, *-l+*, *-n+*, *-ng+*, *-ny+*, *-r+*, *-s+*, *-sy+*, *-t+*, *-y+*, *-z+*}
 QS "C-Labiodental_Consonant" {*-f+*}
 QS "C-Alveolar_Consonant" {*-d+*, *-l+*, *-n+*, *-r+*, *-s+*, *-t+*, *-z+*}
 QS "C-Not_Alveolar" {*-b+*, *-c+*, *-f+*, *-g+*, *-h+*, *-j+*, *-k+*, *-kh+*, *-m+*, *-ng+*, *-ny+*, *-p+*, *-sy+*, *-w+*, *-y+*}
 QS "C-Palatal_Consonant" {*-c+*, *-j+*, *-ny+*, *-sy+*, *-y+*}
 QS "C-Not_Palatal" {*-b+*, *-d+*, *-f+*, *-g+*, *-h+*, *-k+*, *-kh+*, *-l+*, *-m+*, *-n+*, *-ng+*, *-p+*, *-r+*, *-s+*, *-t+*, *-w+*, *-z+*}
 QS "C-Velar_Consonant" {*-g+*, *-k+*, *-kh+*, *-ng+*}
 QS "C-Not_Velar" {*-b+*, *-c+*, *-d+*, *-f+*, *-h+*, *-j+*, *-l+*, *-m+*, *-n+*, *-ny+*, *-p+*, *-r+*, *-s+*, *-sy+*, *-t+*, *-w+*, *-y+*, *-z+*}
 QS "C-Glottal_Consonant" {*-h+*}
 QS "C-Voiced_Consonant" {*-b+*, *-d+*, *-g+*, *-j+*, *-w+*, *-z+*}
 QS "C-Unvoiced_Consonant" {*-c+*, *-f+*, *-h+*, *-k+*, *-kh+*, *-l+*, *-m+*, *-n+*, *-ng+*, *-ny+*, *-p+*, *-r+*, *-s+*, *-sy+*, *-t+*, *-y+*}
 QS "C-Plosive_Alveolar" {*-d+*, *-t+*}
 QS "C-Plosive_Bilabial" {*-b+*, *-p+*}
 QS "C-Plosive_Velar" {*-g+*, *-k+*}
 QS "C-Affricate_Palatal" {*-c+*, *-j+*}
 QS "C-Fricative_Alveolar" {*-s+*, *-z+*}
 QS "C-Plosive_Voiced" {*-b+*, *-d+*, *-g+*}
 QS "C-Plosive_Unvoiced" {*-k+*, *-p+*, *-t+*}
 QS "C-Bilabial_Voiced" {*-b+*, *-w+*}
 QS "C-Bilabial_Unvoiced" {*-m+*, *-p+*}
 QS "C-Alveolar_Voiced" {*-d+*, *-z+*}
 QS "C-Alveolar_Unvoiced" {*-l+*, *-n+*, *-r+*, *-s+*, *-t+*}
 QS "C-silences" {*-pau+*}
 QS "C-a" {*-a+*}
 QS "C-aw" {*-aw+*}
 QS "C-ay" {*-ay+*}
 QS "C-b" {*-b+*}
 QS "C-c" {*-c+*}
 QS "C-d" {*-d+*}

QS "C-e" {*-e+*}
 QS "C-ey" {*ey+*}
 QS "C-f" {*-f+*}
 QS "C-g" {*-g+*}
 QS "C-h" {*-h+*}
 QS "C-i" {*-i+*}
 QS "C-j" {*-j+*}
 QS "C-k" {*-k+*}
 QS "C-kh" {*-kh+*}
 QS "C-l" {*-l+*}
 QS "C-m" {*-m+*}
 QS "C-n" {*-n+*}
 QS "C-ng" {*-ng+*}
 QS "C-ny" {*-ny+*}
 QS "C-o" {*-o+*}
 QS "C-oy" {*-oy+*}
 QS "C-p" {*-p+*}
 QS "C-r" {*-r+*}
 QS "C-s" {*-s+*}
 QS "C-sy" {*-sy+*}
 QS "C-t" {*-t+*}
 QS "C-u" {*-u+*}
 QS "C-w" {*-w+*}
 QS "C-y" {*-y+*}
 QS "C-z" {*-z+*}
 QS "C-pau" {*-pau+*}

QS "R-Vowel"
 {*+a=*,*+aw=*,*+ay=*,*+e=*,*+ey=*,*+i=*,*+o=*,*+oy=*,*+u=*}

QS "R-Consonant"
 {*+b=*,*+c=*,*+d=*,*+f=*,*+g=*,*+h=*,*+j=*,*+k=*,*+kh=*,*+l=*,*+m=*,*+n=*,*+ng=*,*+ny=*,*+p=*,*+r=*,*+s=*,*+sy=*,*+t=*,*+w=*,*+y=*,*+z=*}

QS "R-Short_Vowel"
 {*+a=*,*+e=*,*+i=*,*+o=*,*+u=*}

QS "R-Diphthong_Vowel"
 {*+aw=*,*+ay=*,*+ey=*,*+oy=*}

QS "R-High_Vowel" {*+i=*,*+u=*}

QS "R-Mid_Vowel"
 {*+e=*,*+ey=*,*+o=*}

QS "R-Low_Vowel"
 {*+a=*,*+aw=*,*+ay=*,*+oy=*}

QS "R-Front_Vowel"
 {*+a=*,*+e=*,*+ey=*,*+i=*}

QS "R-Central_Vowel"
 {*+aw=*,*+ay=*}

QS "R-Back_Vowel"
 {*+o=*,*+oy=*,*+u=*}

QS "R-Rounded_Vowel"
 {*+o=*,*+oy=*,*+u=*}

QS "R-Unrounded_Vowel"
 {*+a=*,*+aw=*,*+ay=*,*+e=*,*+ey=*,*+i=*}

QS "R-AVowel"
 {*+a=*,*+aw=*,*+ay=*}

QS "R-IVowel" {*+i=*}
 QS "R-UVowel" {*+u=*}
 QS "R-EVowel" {*+e=*,*+ey=*}
 QS "R-OVowel" {*+o=*,*+oy=*}

QS "R-Plosive_Consonant"
 {*+b=*,*+d=*,*+g=*,*+k=*,*+p=*,*+t=*}

QS "R-Not_Plosive"
 {*+c=*,*+f=*,*+h=*,*+j=*,*+kh=*,*+l=*,*+m=*,*+n=*,*+ng=*,*+ny=*,*+r=*,*+s=*,*+sy=*,*+w=*,*+y=*,*+z=*}

QS "R-Affricate_Consonant" {*+c=*,*+j=*}

QS "R-Fricative_Consonant"
 {*+f=*,*+h=*,*+kh=*,*+s=*,*+sy=*,*+z=*}

QS "R-Not_Fricative"
 {*+b=*,*+c=*,*+d=*,*+g=*,*+j=*,*+k=*,*+l=*,*+m=*,*+n=*,*+ng=*,*+ny=*,*+p=*,*+r=*,*+s=*,*+t=*,*+w=*,*+y=*}

QS "R-Nasal_Consonant"
 {*+m=*,*+n=*,*+ng=*,*+ny=*}

QS "R-Not_Nasal"
 {*+b=*,*+c=*,*+d=*,*+f=*,*+g=*,*+h=*,*+j=*,*+k=*,*+kh=*,*+l=*,*+m=*,*+p=*,*+r=*,*+s=*,*+sy=*,*+t=*,*+w=*,*+y=*,*+z=*}

QS "R-Trill_Consonant" {*+r=*}

QS "R-Lateral_Consonant" {*+l=*}

QS "R-SemiVowel_Consonant"
 {*+w=*,*+y=*}

QS "R-Bilabial_Consonant"
 {*+b=*,*+m=*,*+p=*,*+w=*}

QS "R-Not_Bilabial"
 {*+c=*,*+d=*,*+f=*,*+g=*,*+h=*,*+j=*,*+k=*,*+kh=*,*+l=*,*+m=*,*+ng=*,*+ny=*,*+r=*,*+s=*,*+sy=*,*+t=*,*+y=*,*+z=*}

QS "R-Labiodental_Consonant" {*+f=*}

QS "R-Alveolar_Consonant"
 {*+d=*,*+l=*,*+n=*,*+r=*,*+s=*,*+t=*,*+z=*}

QS "R-Not_Alveolar"
 {*+b=*,*+c=*,*+f=*,*+g=*,*+h=*,*+j=*,*+k=*,*+kh=*,*+m=*,*+ng=*,*+ny=*,*+p=*,*+sy=*,*+w=*,*+y=*}

QS "R-Palatal_Consonant"
 {*+c=*,*+j=*,*+ny=*,*+sy=*,*+y=*}

QS "R-Not_Palatal"
 {*+b=*,*+d=*,*+f=*,*+g=*,*+h=*,*+k=*,*+kh=*,*+l=*,*+m=*,*+n=*,*+ng=*,*+p=*,*+r=*,*+s=*,*+t=*,*+w=*,*+z=*}

QS "R-Velar_Consonant"
 {*+g=*,*+k=*,*+kh=*,*+ng=*}

QS "R-Not_Velar"
 {*+b=*,*+c=*,*+d=*,*+f=*,*+h=*,*+j=*,*+l=*,*+m=*,*+n=*,*+ny=*,*+p=*,*+r=*,*+s=*,*+sy=*,*+t=*,*+w=*,*+y=*,*+z=*}

QS "R-Glottal_Consonant" {*+h=*}

QS "R-Voiced_Consonant"
 {*+b=*,*+d=*,*+g=*,*+j=*,*+w=*,*+z=*}

QS "R-Unvoiced_Consonant"
 {*+c=*,*+f=*,*+h=*,*+k=*,*+kh=*,*+l=*,*+m=*,*+n=*,*+ng=*,*+ny=*,*+p=*,*+r=*,*+s=*,*+sy=*,*+t=*,*+y=*}

QS "R-Plosive_Alveolar" {*+d=*,*+t=*}

QS "R-Plosive_Bilabial" {*+b=*,*+p=*}

QS "R-Plosive_Velar" {*+g=*,*+k=*}

QS "R-Affricate_Palatal" {*+c=*,*+j=*}

QS "R-Fricative_Alveolar" {*+s=*,*+z=*}

QS "R-Plosive_Voiced" { *+b=* , *+d=* , *+g=* }
 QS "R-Plosive_Unvoiced" { *+k=* , *+p=* , *+t=* }
 QS "R-Bilabial_Voiced" { *+b=* , *+w=* }
 QS "R-Bilabial_Unvoiced" { *+m=* , *+p=* }
 QS "R-Alveolar_Voiced" { *+d=* , *+z=* }
 QS "R-Alveolar_Unvoiced" { *+l=* , *+n=* , *+r=* , *+s=* , *+t=* }
 QS "R-silences" { *+pau=* }
 QS "R-a" { *+a=* }
 QS "R-aw" { *+aw=* }
 QS "R-ay" { *+ay=* }
 QS "R-b" { *+b=* }
 QS "R-c" { *+c=* }
 QS "R-d" { *+d=* }
 QS "R-e" { *+e=* }
 QS "R-ey" { *+ey=* }
 QS "R-f" { *+f=* }
 QS "R-g" { *+g=* }
 QS "R-h" { *+h=* }
 QS "R-i" { *+i=* }
 QS "R-j" { *+j=* }
 QS "R-k" { *+k=* }
 QS "R-kh" { *+kh=* }
 QS "R-l" { *+l=* }
 QS "R-m" { *+m=* }
 QS "R-n" { *+n=* }
 QS "R-ng" { *+ng=* }
 QS "R-ny" { *+ny=* }
 QS "R-o" { *+o=* }
 QS "R-oy" { *+oy=* }
 QS "R-p" { *+p=* }
 QS "R-r" { *+r=* }
 QS "R-s" { *+s=* }
 QS "R-sy" { *+sy=* }
 QS "R-t" { *+t=* }
 QS "R-u" { *+u=* }
 QS "R-w" { *+w=* }
 QS "R-y" { *+y=* }
 QS "R-z" { *+z=* }
 QS "R-pau" { *+pau=* }

QS "RR-Vowel" { *+a@* , *+aw@* , *+ay@* , *+e@* , *+ey@* , *+i@* , *+o@* , *+oy@* , *+u@* }
 QS "RR-Consonant" { *+b@* , *+c@* , *+d@* , *+f@* , *+g@* , *+h@* , *+j@* , *+k@* , *+kh@* , *+l@* , *+m@* , *+n@* , *+ng@* , *+ny@* , *+p@* , *+r@* , *+s@* , *+sy@* , *+t@* , *+w@* , *+y@* , *+z@* }
 QS "RR-Short_Vowel" { *+a@* , *+e@* , *+i@* , *+o@* , *+u@* }
 QS "RR-Diphthong_Vowel" { *+aw@* , *+ay@* , *+ey@* , *+oy@* }
 QS "RR-High_Vowel" { *+i@* , *+u@* }
 QS "RR-Mid_Vowel" { *+e@* , *+ey@* , *+o@* }

QS "RR-Low_Vowel" { *+a@* , *+aw@* , *+ay@* , *+oy@* }
 QS "RR-Front_Vowel" { *+a@* , *+e@* , *+ey@* , *+i@* }
 QS "RR-Central_Vowel" { *+aw@* , *+ay@* }
 QS "RR-Back_Vowel" { *+o@* , *+oy@* , *+u@* }
 QS "RR-Rounded_Vowel" { *+o@* , *+oy@* , *+u@* }
 QS "RR-Unrounded_Vowel" { *+a@* , *+aw@* , *+ay@* , *+e@* , *+ey@* , *+i@* }
 QS "RR-AVowel" { *+a@* , *+aw@* , *+ay@* }
 QS "RR-IVowel" { *+i@* }
 QS "RR-UVowel" { *+u@* }
 QS "RR-EVowel" { *+e@* , *+ey@* }
 QS "RR-OVowel" { *+o@* , *+oy@* }
 QS "RR-Plosive_Consonant" { *+b@* , *+d@* , *+g@* , *+k@* , *+p@* , *+t@* }
 QS "RR-Not_Plosive" { *+c@* , *+f@* , *+h@* , *+j@* , *+kh@* , *+l@* , *+m@* , *+n@* , *+ng@* , *+ny@* , *+r@* , *+s@* , *+sy@* , *+w@* , *+y@* , *+z@* }
 QS "RR-Affricate_Consonant" { *+c@* , *+j@* }
 QS "RR-Fricative_Consonant" { *+f@* , *+h@* , *+kh@* , *+s@* , *+sy@* , *+z@* }
 QS "RR-Not_Fricative" { *+b@* , *+c@* , *+d@* , *+g@* , *+j@* , *+k@* , *+l@* , *+m@* , *+n@* , *+ng@* , *+ny@* , *+p@* , *+r@* , *+t@* , *+w@* , *+y@* }
 QS "RR-Nasal_Consonant" { *+m@* , *+n@* , *+ng@* , *+ny@* }
 QS "RR-Not_Nasal" { *+b@* , *+c@* , *+d@* , *+f@* , *+g@* , *+h@* , *+j@* , *+k@* , *+kh@* , *+l@* , *+p@* , *+r@* , *+s@* , *+sy@* , *+t@* , *+w@* , *+y@* , *+z@* }
 QS "RR-Trill_Consonant" { *+r@* }
 QS "RR-Lateral_Consonant" { *+l@* }
 QS "RR-SemiVowel_Consonant" { *+w@* , *+y@* }
 QS "RR-Bilabial_Consonant" { *+b@* , *+m@* , *+p@* , *+w@* }
 QS "RR-Not_Bilabial" { *+c@* , *+d@* , *+f@* , *+g@* , *+h@* , *+j@* , *+k@* , *+kh@* , *+l@* , *+n@* , *+ng@* , *+ny@* , *+p@* , *+r@* , *+s@* , *+sy@* , *+t@* , *+y@* , *+z@* }
 QS "RR-Labiodental_Consonant" { *+f@* }
 QS "RR-Alveolar_Consonant" { *+d@* , *+l@* , *+n@* , *+r@* , *+s@* , *+t@* , *+z@* }
 QS "RR-Not_Alveolar" { *+b@* , *+c@* , *+f@* , *+g@* , *+h@* , *+j@* , *+k@* , *+kh@* , *+m@* , *+ng@* , *+ny@* , *+p@* , *+sy@* , *+w@* , *+y@* }
 QS "RR-Palatal_Consonant" { *+c@* , *+j@* , *+ny@* , *+sy@* , *+y@* }

QS "RR-Not_Palatal"
 {*=b@*,*=d@*,*=f@*,*=g@*,*=h@*,*=k@*,
 =kh@,*=l@*,*=m@*,*=n@*,*=ng@*,*=p@
 ,=r@*,*=s@*,*=t@*,*=w@*,*=z@*}
 QS "RR-Velar_Consonant"
 {*=g@*,*=k@*,*=kh@*,*=ng@*}
 QS "RR-Not_Velar"
 {*=b@*,*=c@*,*=d@*,*=f@*,*=h@*,*=j@*,
 =l@,*=m@*,*=n@*,*=ny@*,*=p@*,*=r@*,
 =s@,*=sy@*,*=t@*,*=w@*,*=y@*,*=z@*}
 QS "RR-Glottal_Consonant" {*=h@*}
 QS "RR-Voiced_Consonant"
 {*=b@*,*=d@*,*=g@*,*=j@*,*=w@*,*=z@*
 }
 QS "RR-Unvoiced_Consonant"
 {*=c@*,*=f@*,*=h@*,*=k@*,*=kh@*,*=l@*,
 =m@,*=n@*,*=ng@*,*=ny@*,*=p@*,*=r@*,
 =s@,*=sy@*,*=t@*,*=y@*}
 QS "RR-Plosive_Alveolar"
 {*=d@*,*=t@*}
 QS "RR-Plosive_Bilabial"
 {*=b@*,*=p@*}
 QS "RR-Plosive_Velar"
 {*=g@*,*=k@*}
 QS "RR-Affricate_Palatal" {*=c@*,*=j@*}
 QS "RR-Fricative_Alveolar"
 {*=s@*,*=z@*}
 QS "RR-Plosive_Voiced"
 {*=b@*,*=d@*,*=g@*}
 QS "RR-Plosive_Unvoiced"
 {*=k@*,*=p@*,*=t@*}
 QS "RR-Bilabial_Voiced"
 {*=b@*,*=w@*}
 QS "RR-Bilabial_Unvoiced"
 {*=m@*,*=p@*}
 QS "RR-Alveolar_Voiced"
 {*=d@*,*=z@*}

QS "RR-Alveolar_Unvoiced"
 {*=l@*,*=n@*,*=r@*,*=s@*,*=t@*}
 QS "RR-silences" {*=pau@*}
 QS "RR-a" {*=a@*}
 QS "RR-aw" {*=aw@*}
 QS "RR-ay" {*=ay@*}
 QS "RR-b" {*=b@*}
 QS "RR-c" {*=c@*}
 QS "RR-d" {*=d@*}
 QS "RR-e" {*=e@*}
 QS "RR-ey" {*=ey@*}
 QS "RR-f" {*=f@*}
 QS "RR-g" {*=g@*}
 QS "RR-h" {*=h@*}
 QS "RR-i" {*=i@*}
 QS "RR-j" {*=j@*}
 QS "RR-k" {*=k@*}
 QS "RR-kh" {*=kh@*}
 QS "RR-l" {*=l@*}
 QS "RR-m" {*=m@*}
 QS "RR-n" {*=n@*}
 QS "RR-ng" {*=ng@*}
 QS "RR-ny" {*=ny@*}
 QS "RR-o" {*=o@*}
 QS "RR-oy" {*=oy@*}
 QS "RR-p" {*=p@*}
 QS "RR-r" {*=r@*}
 QS "RR-s" {*=s@*}
 QS "RR-sy" {*=sy@*}
 QS "RR-t" {*=t@*}
 QS "RR-u" {*=u@*}
 QS "RR-w" {*=w@*}
 QS "RR-y" {*=y@*}
 QS "RR-z" {*=z@*}
 QS "RR-pau" {*=pau@*}

A.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.

QS "Word_Fw==x" {*@x_*}	QS "L-Word_Num-Segs==3" {*/A:3/B:*}
QS "Word_Fw==1" {@1_*}	QS "L-Word_Num-Segs==4" {*/A:4/B:*}
QS "Word_Fw==2" {@2_*}	QS "L-Word_Num-Segs==5" {*/A:5/B:*}
QS "Word_Fw==3" {@3_*}	QS "L-Word_Num-Segs==6" {*/A:6/B:*}
QS "Word_Fw==4" {@4_*}	QS "L-Word_Num-Segs==7" {*/A:7/B:*}
QS "Word_Fw==5" {@5_*}	
QS "Word_Fw==6" {@6_*}	QS "L-Word_Num-Segs<=1" {*/A:0/B:*,*/A:1/B:*}
QS "Word_Fw==7" {@7_*}	QS "L-Word_Num-Segs<=2" {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*}
	QS "L-Word_Num-Segs<=3" {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*}
QS "Word_Fw<=1" {@x_*,*@1_*}	QS "L-Word_Num-Segs<=4" {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*}
QS "Word_Fw<=2" {@x_*,*@1_*,*@2_*}	QS "L-Word_Num-Segs<=5" {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*}
{*@x_*,*@1_*,*@2_*}	QS "L-Word_Num-Segs<=6" {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*}
QS "Word_Fw<=3" {@x_*,*@1_*,*@2_*,*@3_*}	QS "L-Word_Num-Segs<=7" {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*}
{*@x_*,*@1_*,*@2_*,*@3_*}	
QS "Word_Fw<=4" {@x_*,*@1_*,*@2_*,*@3_*,*@4_*}	
{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*}	
QS "Word_Fw<=5" {@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*}	
{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*}	
QS "Word_Fw<=6" {@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*}	
{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*}	
QS "Word_Bw==x" {*_x/A:*}	QS "C-Word_Num-Segs==x" {*/B:x@*}
QS "Word_Bw==1" {*_1/A:*}	QS "C-Word_Num-Segs==1" {*/B:1@*}
QS "Word_Bw==2" {*_2/A:*}	QS "C-Word_Num-Segs==2" {*/B:2@*}
QS "Word_Bw==3" {*_3/A:*}	QS "C-Word_Num-Segs==3" {*/B:3@*}
QS "Word_Bw==4" {*_4/A:*}	QS "C-Word_Num-Segs==4" {*/B:4@*}
QS "Word_Bw==5" {*_5/A:*}	QS "C-Word_Num-Segs==5" {*/B:5@*}
QS "Word_Bw==6" {*_6/A:*}	QS "C-Word_Num-Segs==6" {*/B:6@*}
QS "Word_Bw==7" {*_7/A:*}	QS "C-Word_Num-Segs==7" {*/B:7@*}
QS "Word_Bw<=1" {*_x/A:*,*_1/A:*}	QS "C-Word_Num-Segs<=1" {*/B:x@*,*/B:1@*}
QS "Word_Bw<=2" {*_x/A:*,*_1/A:*,*_2/A:*}	QS "C-Word_Num-Segs<=2" {*/B:x@*,*/B:1@*,*/B:2@*}
QS "Word_Bw<=3" {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*}	QS "C-Word_Num-Segs<=3" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*}
QS "Word_Bw<=4" {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*,*_4/A:*}	QS "C-Word_Num-Segs<=4" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*}
QS "Word_Bw<=5" {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*,*_4/A:*,*_5/A:*}	QS "C-Word_Num-Segs<=5" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*,*/B:5@*}
QS "Word_Bw<=6" {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*,*_4/A:*,*_5/A:*,*_6/A:*}	QS "C-Word_Num-Segs<=6" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*,*/B:5@*,*/B:6@*}
QS "Word_Bw<=7" {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*,*_4/A:*,*_5/A:*,*_6/A:*,*_7/A:*}	QS "C-Word_Num-Segs<=7" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*,*/B:5@*,*/B:6@*,*/B:7@*}
QS "L-Word_Num-Segs==0" {*/A:0/B:*}	
QS "L-Word_Num-Segs==1" {*/A:1/B:*}	QS "Pos_C-Word_in_C-Utterance(Fw)==x" {*@x_*}
QS "L-Word_Num-Segs==2" {*/A:2/B:*}	

QS "Pos_C-Word_in_C-Utterance(Fw)==1"
 {*@1_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==2"
 {*@2_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==3"
 {*@3_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==4"
 {*@4_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==5"
 {*@5_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==6"
 {*@6_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==7"
 {*@7_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==8"
 {*@8_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==9"
 {*@9_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==10"
 {*@10_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==11"
 {*@11_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==12"
 {*@12_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==13"
 {*@13_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==14"
 {*@14_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==15"
 {*@15_*}

QS "Pos_C-Word_in_C-Utterance(Fw)<=1"
 {*@x_*,*@1_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=2"
 {*@x_*,*@1_*,*@2_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=3"
 {*@x_*,*@1_*,*@2_*,*@3_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=4"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=5"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=6"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=7"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=8"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=9"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=10"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=11"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*}

QS "Pos_C-Word_in_C-Utterance(Fw)<=12"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=13"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=14"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=15"
 {*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*}

QS "Pos_C-Word_in_C-Utterance(Bw)==x"
 {*_x/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==1"
 {*_1/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==2"
 {*_2/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==3"
 {*_3/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==4"
 {*_4/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==5"
 {*_5/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==6"
 {*_6/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==7"
 {*_7/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==8"
 {*_8/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==9"
 {*_9/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==10"
 {*_10/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==11"
 {*_11/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==12"
 {*_12/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==13"
 {*_13/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==14"
 {*_14/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)==15"
 {*_15/C:}*

QS "Pos_C-Word_in_C-Utterance(Bw)<=1"
 {*_x/C:*,*_1/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)<=2"
 {*_x/C:*,*_1/C:*,*_2/C:}*
 QS "Pos_C-Word_in_C-Utterance(Bw)<=3"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:}*

QS "Pos_C-Word_in_C-Utterance(Bw)<=4"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=5"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=6"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=7"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=8"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=9"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*,*_9/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=10"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*,*_9/C:*,*_10/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=11"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*,*_9/C:*,*_10/C:*,*_11/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=12"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*,*_9/C:*,*_10/C:*,*_11/C:*,*_12/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=13"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*,*_9/C:*,*_10/C:*,*_11/C:*,*_12/C:*,*_13/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=14"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*,*_9/C:*,*_10/C:*,*_11/C:*,*_12/C:*,*_13/C:*,*_14/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=15"
 {*_x/C:*,*_1/C:*,*_2/C:*,*_3/C:*,*_4/C:*,*_5/C:*,*_6/C:*,*_7/C:*,*_8/C:*,*_9/C:*,*_10/C:*,*_11/C:*,*_12/C:*,*_13/C:*,*_14/C:*,*_15/C:*}

QS "R-Word_Num-Segs==0" {*/C:0/D:*}
 QS "R-Word_Num-Segs==1" {*/C:1/D:*}
 QS "R-Word_Num-Segs==2" {*/C:2/D:*}
 QS "R-Word_Num-Segs==3" {*/C:3/D:*}
 QS "R-Word_Num-Segs==4" {*/C:4/D:*}
 QS "R-Word_Num-Segs==5" {*/C:5/D:*}
 QS "R-Word_Num-Segs==6" {*/C:6/D:*}
 QS "R-Word_Num-Segs==7" {*/C:7/D:*}

QS "R-Word_Num-Segs<=1"
 {*/C:0/D:* */C:1/D:*}
 QS "R-Word_Num-Segs<=2"
 {*/C:0/D:* */C:1/D:* */C:2/D:*}
 QS "R-Word_Num-Segs<=3"
 {*/C:0/D:* */C:1/D:* */C:2/D:* */C:3/D:*}

QS "R-Word_Num-Segs<=4"
 {*/C:0/D:* */C:1/D:* */C:2/D:* */C:3/D:* */C:4/D:*}
 QS "R-Word_Num-Segs<=5"
 {*/C:0/D:* */C:1/D:* */C:2/D:* */C:3/D:* */C:4/D:* */C:5/D:*}
 QS "R-Word_Num-Segs<=6"
 {*/C:0/D:* */C:1/D:* */C:2/D:* */C:3/D:* */C:4/D:* */C:5/D:* */C:6/D:*}
 QS "R-Word_Num-Segs<=7"
 {*/C:0/D:* */C:1/D:* */C:2/D:* */C:3/D:* */C:4/D:* */C:5/D:* */C:6/D:* */C:7/D:*}

QS "Num-Words_in_Utterance==1" {*/D:1}
 QS "Num-Words_in_Utterance==2" {*/D:2}
 QS "Num-Words_in_Utterance==3" {*/D:3}
 QS "Num-Words_in_Utterance==4" {*/D:4}
 QS "Num-Words_in_Utterance==5" {*/D:5}
 QS "Num-Words_in_Utterance==6" {*/D:6}
 QS "Num-Words_in_Utterance==7" {*/D:7}
 QS "Num-Words_in_Utterance==8" {*/D:8}
 QS "Num-Words_in_Utterance==9" {*/D:9}
 QS "Num-Words_in_Utterance==10" {*/D:10}
 QS "Num-Words_in_Utterance==11" {*/D:11}
 QS "Num-Words_in_Utterance==12" {*/D:12}
 QS "Num-Words_in_Utterance==13" {*/D:13}

QS "Num-Words_in_Utterance<=2"
 {*/D:x,*/D:1,*/D:2}
 QS "Num-Words_in_Utterance<=3"
 {*/D:x,*/D:1,*/D:2,*/D:3}
 QS "Num-Words_in_Utterance<=4"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4}
 QS "Num-Words_in_Utterance<=5"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5}
 QS "Num-Words_in_Utterance<=6"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6}
 QS "Num-Words_in_Utterance<=7"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:7}
 QS "Num-Words_in_Utterance<=8"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:7,*/D:8}
 QS "Num-Words_in_Utterance<=9"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:7,*/D:8,*/D:9}
 QS "Num-Words_in_Utterance<=10"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:7,*/D:8,*/D:9,*/D:10}
 QS "Num-Words_in_Utterance<=11"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:7,*/D:8,*/D:9,*/D:10,*/D:11}
 QS "Num-Words_in_Utterance<=12"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12}
 QS "Num-Words_in_Utterance<=13"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13}

LAMPIRAN B. Question Set 2 untuk Konteks (Sakti, 2008)

B.I. Bagian set pertanyaan terkait konteks fonetis

Sama dengan lampiran A bagian I.

B.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.

QS "Word_Fw==x"	{*@x_*}	QS "Word_Fw<=10"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*}
QS "Word_Fw==1"	{*@1_*}	QS "Word_Fw<=12"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*}
QS "Word_Fw==2"	{*@2_*}	QS "Word_Fw<=13"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*}
QS "Word_Fw==3"	{*@3_*}	QS "Word_Fw<=14"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*}
QS "Word_Fw==4"	{*@4_*}	QS "Word_Fw<=15"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*}
QS "Word_Fw==5"	{*@5_*}	QS "Word_Fw<=16"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*}
QS "Word_Fw==6"	{*@6_*}	QS "Word_Fw<=17"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*,*@16_*}
QS "Word_Fw==7"	{*@7_*}	QS "Word_Fw<=18"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@17_*}
QS "Word_Fw==8"	{*@8_*}	QS "Word_Fw<=19"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@17_*,*@18_*}
QS "Word_Fw==9"	{*@9_*}	QS "Word_Fw<=20"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@17_*,*@18_*,*@20_*}
QS "Word_Fw==10"	{*@10_*}	QS "Word_Fw<=21"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@17_*,*@18_*,*@20_*,*@21_*}
QS "Word_Fw==12"	{*@11_*}	QS "Word_Fw<=22"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@17_*,*@18_*,*@20_*,*@21_*}
QS "Word_Fw==13"	{*@12_*}		
QS "Word_Fw==14"	{*@13_*}		
QS "Word_Fw==15"	{*@14_*}		
QS "Word_Fw==16"	{*@15_*}		
QS "Word_Fw==17"	{*@16_*}		
QS "Word_Fw==18"	{*@17_*}		
QS "Word_Fw==19"	{*@18_*}		
QS "Word_Fw==20"	{*@20_*}		
QS "Word_Fw==21"	{*@21_*}		
QS "Word_Fw==22"	{*@22_*}		
QS "Word_Fw==23"	{*@23_*}		
QS "Word_Fw==24"	{*@24_*}		
QS "Word_Fw==25"	{*@25_*}		
QS "Word_Fw<=1"	{*@x_*,*@1_*}		
QS "Word_Fw<=2"	{*@x_*,*@1_*,*@2_*}		
QS "Word_Fw<=3"	{*@x_*,*@1_*,*@2_*,*@3_*}		
QS "Word_Fw<=4"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*}		
QS "Word_Fw<=5"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*}		
QS "Word_Fw<=6"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*}		
QS "Word_Fw<=7"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*}		
QS "Word_Fw<=8"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*}		
QS "Word_Fw<=9"	{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*@6_*,*@7_*,*@8_*,*@9_*}		

```

@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@1
7_*,*@18_*,*@20_*,*@21_*,*@22_*}
QS "Word_Fw<=23"
{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*
@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*
@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@1
7_*,*@18_*,*@20_*,*@21_*,*@22_*,*@23_
*}
QS "Word_Fw<=24"
{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*
@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*
@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@1
7_*,*@18_*,*@20_*,*@21_*,*@22_*,*@23_
*,*@24_*}
QS "Word_Fw<=25"
{*@x_*,*@1_*,*@2_*,*@3_*,*@4_*,*@5_*,*
@6_*,*@7_*,*@8_*,*@9_*,*@10_*,*@11_*,*
@12_*,*@13_*,*@14_*,*@15_*,*@16_*,*@1
7_*,*@18_*,*@20_*,*@21_*,*@22_*,*@23_
*,*@24_*,*@25_*}

QS "Word_Bw==x"      {*_x/A:*}
QS "Word_Bw==1"     {*_1/A:*}
QS "Word_Bw==2"     {*_2/A:*}
QS "Word_Bw==3"     {*_3/A:*}
QS "Word_Bw==4"     {*_4/A:*}
QS "Word_Bw==5"     {*_5/A:*}
QS "Word_Bw==6"     {*_6/A:*}
QS "Word_Bw==7"     {*_7/A:*}
QS "Word_Bw==8"     {*_8/A:*}
QS "Word_Bw==9"     {*_9/A:*}
QS "Word_Bw==10"    {*_10/A:*}
QS "Word_Bw==12"    {*_11/A:*}
QS "Word_Bw==13"    {*_12/A:*}
QS "Word_Bw==14"    {*_13/A:*}
QS "Word_Bw==15"    {*_14/A:*}
QS "Word_Bw==16"    {*_15/A:*}
QS "Word_Bw==17"    {*_16/A:*}
QS "Word_Bw==18"    {*_17/A:*}
QS "Word_Bw==19"    {*_18/A:*}
QS "Word_Bw==20"    {*_20/A:*}
QS "Word_Bw==21"    {*_21/A:*}
QS "Word_Bw==22"    {*_22/A:*}
QS "Word_Bw==23"    {*_23/A:*}
QS "Word_Bw==24"    {*_24/A:*}
QS "Word_Bw==25"    {*_25/A:*}

QS "Word_Bw<=1"
{*_x/A:,*_1/A:*}
QS "Word_Bw<=2"
{*_x/A:,*_1/A:,*_2/A:*}
QS "Word_Bw<=3"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:*}
QS "Word_Bw<=4"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:*}
QS "Word_Bw<=5"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:*}
QS "Word_Bw<=6"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:*}

```

```

QS "Word_Bw<=7"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:*}
QS "Word_Bw<=8"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:*}
QS "Word_Bw<=9"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:*}
QS "Word_Bw<=10"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:*}
}
QS "Word_Bw<=12"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:*}
QS "Word_Bw<=13"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:*}
QS "Word_Bw<=14"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:*}
QS "Word_Bw<=15"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:,*_14/A:*}
QS "Word_Bw<=16"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:,*_14/A:,*_15/A:*}
}
QS "Word_Bw<=17"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:,*_14/A:,*_15/A:,*
_16/A:*}
QS "Word_Bw<=18"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:,*_14/A:,*_15/A:,*
_16/A:,*_17/A:*}
QS "Word_Bw<=19"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:,*_14/A:,*_15/A:,*
_16/A:,*_17/A:,*_18/A:*}
QS "Word_Bw<=20"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:,*_14/A:,*_15/A:,*
_16/A:,*_17/A:,*_18/A:,*_20/A:*}
}
QS "Word_Bw<=21"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/
A:,*_6/A:,*_7/A:,*_8/A:,*_9/A:,*_10/A:,*
_11/A:,*_12/A:,*_13/A:,*_14/A:,*_15/A:,*
_16/A:,*_17/A:,*_18/A:,*_20/A:,*_21/A:
*}
QS "Word_Bw<=22"
{*_x/A:,*_1/A:,*_2/A:,*_3/A:,*_4/A:,*_5/

```

A:*,*_6/A:*,*_7/A:*,*_8/A:*,*_9/A:*,*_10/A:*,*_11/A:*,*_12/A:*,*_13/A:*,*_14/A:*,*_15/A:*,*_16/A:*,*_17/A:*,*_18/A:*,*_20/A:*,*_21/A:*,*_22/A:*

QS "Word_Bw<=23"
 {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*,*_4/A:*,*_5/A:*,*_6/A:*,*_7/A:*,*_8/A:*,*_9/A:*,*_10/A:*,*_11/A:*,*_12/A:*,*_13/A:*,*_14/A:*,*_15/A:*,*_16/A:*,*_17/A:*,*_18/A:*,*_20/A:*,*_21/A:*,*_22/A:*,*_23/A:*

QS "Word_Bw<=24"
 {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*,*_4/A:*,*_5/A:*,*_6/A:*,*_7/A:*,*_8/A:*,*_9/A:*,*_10/A:*,*_11/A:*,*_12/A:*,*_13/A:*,*_14/A:*,*_15/A:*,*_16/A:*,*_17/A:*,*_18/A:*,*_20/A:*,*_21/A:*,*_22/A:*,*_23/A:*,*_24/A:*

QS "Word_Bw<=25"
 {*_x/A:*,*_1/A:*,*_2/A:*,*_3/A:*,*_4/A:*,*_5/A:*,*_6/A:*,*_7/A:*,*_8/A:*,*_9/A:*,*_10/A:*,*_11/A:*,*_12/A:*,*_13/A:*,*_14/A:*,*_15/A:*,*_16/A:*,*_17/A:*,*_18/A:*,*_20/A:*,*_21/A:*,*_22/A:*,*_23/A:*,*_24/A:*,*_25/A:*

QS "L-Word_Num-Segs==0" {*/A:0/B:*

QS "L-Word_Num-Segs==1" {*/A:1/B:*

QS "L-Word_Num-Segs==2" {*/A:2/B:*

QS "L-Word_Num-Segs==3" {*/A:3/B:*

QS "L-Word_Num-Segs==4" {*/A:4/B:*

QS "L-Word_Num-Segs==5" {*/A:5/B:*

QS "L-Word_Num-Segs==6" {*/A:6/B:*

QS "L-Word_Num-Segs==7" {*/A:7/B:*

QS "L-Word_Num-Segs==8" {*/A:8/B:*

QS "L-Word_Num-Segs==9" {*/A:9/B:*

QS "L-Word_Num-Segs==10" {*/A:10/B:*

QS "L-Word_Num-Segs==11" {*/A:11/B:*

QS "L-Word_Num-Segs==12" {*/A:12/B:*

QS "L-Word_Num-Segs==13" {*/A:13/B:*

QS "L-Word_Num-Segs==14" {*/A:14/B:*

QS "L-Word_Num-Segs==15" {*/A:15/B:*

QS "L-Word_Num-Segs==16" {*/A:16/B:*

QS "L-Word_Num-Segs==17" {*/A:17/B:*

QS "L-Word_Num-Segs==18" {*/A:18/B:*

QS "L-Word_Num-Segs==19" {*/A:19/B:*

QS "L-Word_Num-Segs==20" {*/A:20/B:*

QS "L-Word_Num-Segs==21" {*/A:21/B:*

QS "L-Word_Num-Segs==22" {*/A:22/B:*

QS "L-Word_Num-Segs==23" {*/A:23/B:*

QS "L-Word_Num-Segs==24" {*/A:24/B:*

QS "L-Word_Num-Segs==25" {*/A:25/B:*

QS "L-Word_Num-Segs<=1"
 {*/A:0/B:*,*/A:1/B:*

QS "L-Word_Num-Segs<=2"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*

QS "L-Word_Num-Segs<=3"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*

QS "L-Word_Num-Segs<=4"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*

QS "L-Word_Num-Segs<=5"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*

QS "L-Word_Num-Segs<=6"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*

QS "L-Word_Num-Segs<=7"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*

QS "L-Word_Num-Segs<=8"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*

QS "L-Word_Num-Segs<=9"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*

QS "L-Word_Num-Segs<=10"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*,*/A:10/B:*

QS "L-Word_Num-Segs<=11"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*,*/A:10/B:*,*/A:11/B:*

QS "L-Word_Num-Segs<=12"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*,*/A:10/B:*,*/A:11/B:*,*/A:12/B:*

QS "L-Word_Num-Segs<=13"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*,*/A:10/B:*,*/A:11/B:*,*/A:12/B:*,*/A:13/B:*

QS "L-Word_Num-Segs<=14"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*,*/A:10/B:*,*/A:11/B:*,*/A:12/B:*,*/A:13/B:*,*/A:14/B:*

QS "L-Word_Num-Segs<=15"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*,*/A:10/B:*,*/A:11/B:*,*/A:12/B:*,*/A:13/B:*,*/A:14/B:*,*/A:15/B:*

QS "L-Word_Num-Segs<=16"
 {*/A:0/B:*,*/A:1/B:*,*/A:2/B:*,*/A:3/B:*,*/A:4/B:*,*/A:5/B:*,*/A:6/B:*,*/A:7/B:*,*/A:8/B:*,*/A:9/B:*,*/A:10/B:*,*/A:11/B:*,*/A:12/B:*,*/A:13/B:*,*/A:14/B:*,*/A:15/B:*,*/A:16/B:*

A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/

QS "L-Word_Num-Segs<=17"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/}

QS "L-Word_Num-Segs<=18"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/}

QS "L-Word_Num-Segs<=19"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/A:19/B:*/}

QS "L-Word_Num-Segs<=20"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/A:19/B:*/A:20/B:*/}

QS "L-Word_Num-Segs<=21"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/A:19/B:*/A:20/B:*/A:21/B:*/}

QS "L-Word_Num-Segs<=22"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/A:19/B:*/A:20/B:*/A:21/B:*/A:22/B:*/}

QS "L-Word_Num-Segs<=23"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/A:19/B:*/A:20/B:*/A:21/B:*/A:22/B:*/A:23/B:*/}

QS "L-Word_Num-Segs<=24"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/A:19/B:*/A:20/B:*/A:21/B:*/A:22/B:*/A:23/B:*/A:24/B:*/}

QS "L-Word_Num-Segs<=25"
 {*/A:0/B:*/A:1/B:*/A:2/B:*/A:3/B:*/A:4/B:*/A:5/B:*/A:6/B:*/A:7/B:*/A:8/B:*/A:9/B:*/A:10/B:*/A:11/B:*/A:12/B:*/A:13/B:*/A:14/B:*/A:15/B:*/A:16/B:*/A:17/B:*/A:18/B:*/A:19/B:*/A:20/B:*/A:21/B:*/A:22/B:*/A:23/B:*/A:24/B:*/A:25/B:*/}

QS "C-Word_Num-Segs==x" {*/B:x@*}

QS "C-Word_Num-Segs==1" {*/B:1@*}

QS "C-Word_Num-Segs==2" {*/B:2@*}

QS "C-Word_Num-Segs==3" {*/B:3@*}

QS "C-Word_Num-Segs==4" {*/B:4@*}

QS "C-Word_Num-Segs==5" {*/B:5@*}

QS "C-Word_Num-Segs==6" {*/B:6@*}

QS "C-Word_Num-Segs==7" {*/B:7@*}

QS "C-Word_Num-Segs==8" {*/B:8@*}

QS "C-Word_Num-Segs==9" {*/B:9@*}

QS "C-Word_Num-Segs==10" {*/B:10@*}

QS "C-Word_Num-Segs==11" {*/B:11@*}

QS "C-Word_Num-Segs==12" {*/B:12@*}

QS "C-Word_Num-Segs==13" {*/B:13@*}

QS "C-Word_Num-Segs==14" {*/B:14@*}

QS "C-Word_Num-Segs==15" {*/B:15@*}

QS "C-Word_Num-Segs==16" {*/B:16@*}

QS "C-Word_Num-Segs==17" {*/B:17@*}

QS "C-Word_Num-Segs==18" {*/B:18@*}

QS "C-Word_Num-Segs==19" {*/B:19@*}

QS "C-Word_Num-Segs==20" {*/B:20@*}

QS "C-Word_Num-Segs==21" {*/B:21@*}

QS "C-Word_Num-Segs==22" {*/B:22@*}

QS "C-Word_Num-Segs==23" {*/B:23@*}

QS "C-Word_Num-Segs==24" {*/B:24@*}

QS "C-Word_Num-Segs==25" {*/B:25@*}

QS "C-Word_Num-Segs<=1" {*/B:x@*,*/B:1@*}

QS "C-Word_Num-Segs<=2" {*/B:x@*,*/B:1@*,*/B:2@*}

QS "C-Word_Num-Segs<=3" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*}

QS "C-Word_Num-Segs<=4" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*}

QS "C-Word_Num-Segs<=5" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*,*/B:5@*}

QS "C-Word_Num-Segs<=6" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*,*/B:5@*,*/B:6@*}

QS "C-Word_Num-Segs<=7" {*/B:x@*,*/B:1@*,*/B:2@*,*/B:3@*,*/B:4@*,*/B:5@*,*/B:6@*,*/B:7@*}

QS "Pos_C-Word_in_C-Utterance(Fw)==15"
 {*@15-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==16"
 {*@16-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==17"
 {*@17-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==18"
 {*@18-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==19"
 {*@19-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)==20"
 {*@20-*}

QS "Pos_C-Word_in_C-Utterance(Fw)<=1"
 {*@x-*,*@1-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=2"
 {*@x-*,*@1-*,*@2-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=3"
 {*@x-*,*@1-*,*@2-*,*@3-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=4"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=5"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=6"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=7"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=8"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=9"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=10"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=11"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=12"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=13"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=14"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*,*@14-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=15"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*,*@14-*,*@15-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=16"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*,*@14-*,*@15-*,*@16-*}

QS "Pos_C-Word_in_C-Utterance(Fw)<=17"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*,*@14-*,*@15-*,*@16-*,*@17-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=18"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*,*@14-*,*@15-*,*@16-*,*@17-*,*@18-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=19"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*,*@14-*,*@15-*,*@16-*,*@17-*,*@18-*,*@19-*}
 QS "Pos_C-Word_in_C-Utterance(Fw)<=20"
 {*@x-*,*@1-*,*@2-*,*@3-*,*@4-*,*@5-*,*@6-*,*@7-*,*@8-*,*@9-*,*@10-*,*@11-*,*@12-*,*@13-*,*@14-*,*@15-*,*@16-*,*@17-*,*@18-*,*@19-*,*@20-*}

QS "Pos_C-Word_in_C-Utterance(Bw)==x"
 {*_x/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==1"
 {*_1/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==2"
 {*_2/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==3"
 {*_3/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==4"
 {*_4/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==5"
 {*_5/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==6"
 {*_6/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==7"
 {*_7/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==8"
 {*_8/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==9"
 {*_9/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==10"
 {*_10/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==11"
 {*_11/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==12"
 {*_12/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==13"
 {*_13/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==14"
 {*_14/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==15"
 {*_15/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==16"
 {*_16/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==17"
 {*_17/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==18"
 {*_18/C:*}

QS "Pos_C-Word_in_C-Utterance(Bw)==19"
 {*_19/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)==20"
 {*_20/C:*}

QS "Pos_C-Word_in_C-Utterance(Bw)<=1"
 {*_x/C:**_1/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=2"
 {*_x/C:**_1/C:**_2/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=3"
 {*_x/C:**_1/C:**_2/C:**_3/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=4"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=5"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=6"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=7"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=8"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=9"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=10"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=11"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:**_11/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=12"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:**_11/C:**_12/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=13"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:**_11/C:**_12/C:**_13/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=14"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:**_11/C:**_12/C:**_13/C:**_14/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=15"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:**_11/C:**_12/C:**_13/C:**_14/C:**_
 _15/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=16"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:**_11/C:**_12/C:**_13/C:**_14/C:**_
 _15/C:**_16/C:*}
 QS "Pos_C-Word_in_C-Utterance(Bw)<=17"
 {*_x/C:**_1/C:**_2/C:**_3/C:**_4
 /C:**_5/C:**_6/C:**_7/C:**_8/C:**_9/C:**_
 _10/C:**_11/C:**_12/C:**_13/C:**_14/C:**_
 _15/C:**_16/C:**_17/C:*}

QS "R-Word_Num-Segs==0" {*/C:0/D:*}
 QS "R-Word_Num-Segs==1" {*/C:1/D:*}
 QS "R-Word_Num-Segs==2" {*/C:2/D:*}
 QS "R-Word_Num-Segs==3" {*/C:3/D:*}
 QS "R-Word_Num-Segs==4" {*/C:4/D:*}
 QS "R-Word_Num-Segs==5" {*/C:5/D:*}
 QS "R-Word_Num-Segs==6" {*/C:6/D:*}
 QS "R-Word_Num-Segs==7" {*/C:7/D:*}
 QS "R-Word_Num-Segs==8" {*/C:8/D:*}
 QS "R-Word_Num-Segs==9" {*/C:9/D:*}
 QS "R-Word_Num-Segs==10" {*/C:10/D:*}
 QS "R-Word_Num-Segs==11" {*/C:11/D:*}
 QS "R-Word_Num-Segs==12" {*/C:12/D:*}
 QS "R-Word_Num-Segs==13" {*/C:13/D:*}
 QS "R-Word_Num-Segs==14" {*/C:14/D:*}
 QS "R-Word_Num-Segs==15" {*/C:15/D:*}
 QS "R-Word_Num-Segs==16" {*/C:16/D:*}
 QS "R-Word_Num-Segs==17" {*/C:17/D:*}
 QS "R-Word_Num-Segs==18" {*/C:18/D:*}
 QS "R-Word_Num-Segs==19" {*/C:19/D:*}
 QS "R-Word_Num-Segs==20" {*/C:20/D:*}
 QS "R-Word_Num-Segs==21" {*/C:21/D:*}
 QS "R-Word_Num-Segs==22" {*/C:22/D:*}
 QS "R-Word_Num-Segs==23" {*/C:23/D:*}
 QS "R-Word_Num-Segs==24" {*/C:24/D:*}
 QS "R-Word_Num-Segs==25" {*/C:25/D:*}

QS "Num-Words_in_Utterance==5" {*/D:5}
 QS "Num-Words_in_Utterance==6" {*/D:6}
 QS "Num-Words_in_Utterance==7" {*/D:7}
 QS "Num-Words_in_Utterance==8" {*/D:8}
 QS "Num-Words_in_Utterance==9" {*/D:9}
 QS "Num-Words_in_Utterance==10" {*/D:10}
 QS "Num-Words_in_Utterance==11" {*/D:11}
 QS "Num-Words_in_Utterance==12" {*/D:12}
 QS "Num-Words_in_Utterance==13" {*/D:13}
 QS "Num-Words_in_Utterance==14" {*/D:14}
 QS "Num-Words_in_Utterance==15" {*/D:15}
 QS "Num-Words_in_Utterance==16" {*/D:16}
 QS "Num-Words_in_Utterance==17" {*/D:17}
 QS "Num-Words_in_Utterance==18" {*/D:18}
 QS "Num-Words_in_Utterance==19" {*/D:19}
 QS "Num-Words_in_Utterance==20" {*/D:20}

QS "Num-Words_in_Utterance<=2"
 {*/D:x,*/D:1,*/D:2}
 QS "Num-Words_in_Utterance<=3"
 {*/D:x,*/D:1,*/D:2,*/D:3}
 QS "Num-Words_in_Utterance<=4"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4}
 QS "Num-Words_in_Utterance<=5"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5}
 QS "Num-Words_in_Utterance<=6"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6}
 QS "Num-Words_in_Utterance<=7"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7}
 QS "Num-Words_in_Utterance<=8"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8}
 QS "Num-Words_in_Utterance<=9"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9}
 QS "Num-Words_in_Utterance<=10"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10}

QS "Num-Words_in_Utterance<=11"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11}
 QS "Num-Words_in_Utterance<=12"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12}
 QS "Num-Words_in_Utterance<=13"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13}
 QS "Num-Words_in_Utterance<=14"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13,*/D:
 14}
 QS "Num-Words_in_Utterance<=15"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13,*/D:
 14,*/D:15}
 QS "Num-Words_in_Utterance<=16"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13,*/D:
 14,*/D:15,*/D:16}
 QS "Num-Words_in_Utterance<=17"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13,*/D:
 14,*/D:15,*/D:16,*/D:17}
 QS "Num-Words_in_Utterance<=18"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13,*/D:
 14,*/D:15,*/D:16,*/D:17,*/D:18}
 QS "Num-Words_in_Utterance<=19"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13,*/D:
 14,*/D:15,*/D:16,*/D:17,*/D:18,*/D:19}
 QS "Num-Words_in_Utterance<=20"
 {*/D:x,*/D:1,*/D:2,*/D:3,*/D:4,*/D:5,*/D:6,*/D:
 7,*/D:8,*/D:9,*/D:10,*/D:11,*/D:12,*/D:13,*/D:
 14,*/D:15,*/D:16,*/D:17,*/D:18,*/D:19,*/D:20}

LAMPIRAN C. Question Set 3 untuk Konteks (Vania, 2008)

C.I. Bagian set pertanyaan terkait konteks fonetis

Sama dengan lampiran A bagian I.

C.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.

QS "Pos_In_Syl==x"	{*@x/A:*}	QS "L-Syl_Num-Segs<=3"	{*/A:0_*,*/A:1_*,*/A:2_*,*/A:3_*}
QS "Pos_In_Syl==1"	{*@1/A:*}	QS "L-Syl_Num-Segs<=4"	{*/A:0_*,*/A:1_*,*/A:2_*,*/A:3_*,*/A:4_*}
QS "Pos_In_Syl==2"	{*@2/A:*}	QS "L-Syl_Num-Segs<=5"	{*/A:0_*,*/A:1_*,*/A:2_*,*/A:3_*,*/A:4_*,*/A:5_*}
QS "Pos_In_Syl==3"	{*@3/A:*}	QS "L-Syl_Num-Segs<=6"	{*/A:0_*,*/A:1_*,*/A:2_*,*/A:3_*,*/A:4_*,*/A:5_*,*/A:6_*}
QS "Pos_In_Syl==4"	{*@4/A:*}	QS "L-Syl_Num-Segs<=7"	{*/A:0_*,*/A:1_*,*/A:2_*,*/A:3_*,*/A:4_*,*/A:5_*,*/A:6_*,*/A:7_*}
QS "Pos_In_Syl==5"	{*@5/A:*}	QS "L-Syl_Num-Segs<=8"	{*/A:0_*,*/A:1_*,*/A:2_*,*/A:3_*,*/A:4_*,*/A:5_*,*/A:6_*,*/A:7_*,*/A:8_*}
QS "Pos_In_Syl==6"	{*@6/A:*}	QS "L-Syl_Stress==0"	{*_0/B:*}
QS "Pos_In_Syl==7"	{*@7/A:*}	QS "L-Syl_Stress==1"	{*_1/B:*}
QS "Pos_In_Syl==8"	{*@8/A:*}		
QS "Pos_In_Syl<=1"	{*@x/A:*,*@1/A:*}	QS "C-Syl_Num-Segs==x"	{*/B:x_*}
QS "Pos_In_Syl<=2"	{*@x/A:*,*@1/A:*,*@2/A:*}	QS "C-Syl_Num-Segs==1"	{*/B:1_*}
QS "Pos_In_Syl<=3"	{*@x/A:*,*@1/A:*,*@2/A:*,*@3/A:*}	QS "C-Syl_Num-Segs==2"	{*/B:2_*}
QS "Pos_In_Syl<=4"	{*@x/A:*,*@1/A:*,*@2/A:*,*@3/A:*,*@4/A:*}	QS "C-Syl_Num-Segs==3"	{*/B:3_*}
QS "Pos_In_Syl<=5"	{*@x/A:*,*@1/A:*,*@2/A:*,*@3/A:*,*@4/A:*,*@5/A:*}	QS "C-Syl_Num-Segs==4"	{*/B:4_*}
QS "Pos_In_Syl<=6"	{*@x/A:*,*@1/A:*,*@2/A:*,*@3/A:*,*@4/A:*,*@5/A:*,*@6/A:*}	QS "C-Syl_Num-Segs==5"	{*/B:5_*}
QS "Pos_In_Syl<=7"	{*@x/A:*,*@1/A:*,*@2/A:*,*@3/A:*,*@4/A:*,*@5/A:*,*@6/A:*,*@7/A:*}	QS "C-Syl_Num-Segs==6"	{*/B:6_*}
QS "Pos_In_Syl<=8"	{*@x/A:*,*@1/A:*,*@2/A:*,*@3/A:*,*@4/A:*,*@5/A:*,*@6/A:*,*@7/A:*,*@8/A:*}	QS "C-Syl_Num-Segs==7"	{*/B:7_*}
		QS "C-Syl_Num-Segs==8"	{*/B:8_*}
QS "L-Syl_Num-Segs==0"	{*/A:0_*}	QS "C-Syl_Num-Segs<=1"	{*/B:x_*,*/B:1_*}
QS "L-Syl_Num-Segs==1"	{*/A:1_*}	QS "C-Syl_Num-Segs<=2"	{*/B:x_*,*/B:1_*,*/B:2_*}
QS "L-Syl_Num-Segs==2"	{*/A:2_*}	QS "C-Syl_Num-Segs<=3"	{*/B:x_*,*/B:1_*,*/B:2_*,*/B:3_*}
QS "L-Syl_Num-Segs==3"	{*/A:3_*}	QS "C-Syl_Num-Segs<=4"	{*/B:x_*,*/B:1_*,*/B:2_*,*/B:3_*,*/B:4_*}
QS "L-Syl_Num-Segs==4"	{*/A:4_*}	QS "C-Syl_Num-Segs<=5"	{*/B:x_*,*/B:1_*,*/B:2_*,*/B:3_*,*/B:4_*,*/B:5_*}
QS "L-Syl_Num-Segs==5"	{*/A:5_*}	QS "C-Syl_Num-Segs<=6"	{*/B:x_*,*/B:1_*,*/B:2_*,*/B:3_*,*/B:4_*,*/B:5_*,*/B:6_*}
QS "L-Syl_Num-Segs==6"	{*/A:6_*}	QS "C-Syl_Num-Segs<=7"	{*/B:x_*,*/B:1_*,*/B:2_*,*/B:3_*,*/B:4_*,*/B:5_*,*/B:6_*,*/B:7_*}
QS "L-Syl_Num-Segs==7"	{*/A:7_*}		
QS "L-Syl_Num-Segs==8"	{*/A:8_*}		
QS "L-Syl_Num-Segs<=1"	{*/A:0_*,*/A:1_*}		
QS "L-Syl_Num-Segs<=2"	{*/A:0_*,*/A:1_*,*/A:2_*}		

QS "C-Syl_Num-Segs<=8"
 {*/B:x_*,*/B:1_*,*/B:2_*,*/B:3_*,*/B:4_*,*/B:5_*,*/B:6_*,*/B:7_*,*/B:8_*}

QS "C-Syl_Stress==0" {*_0@*}
 QS "C-Syl_Stress==1" {*_1@*}

QS "Pos_C-Syl_in_C-Word==x" {*@x!*}
 QS "Pos_C-Syl_in_C-Word==1" {*@1!*}
 QS "Pos_C-Syl_in_C-Word==2" {*@2!*}
 QS "Pos_C-Syl_in_C-Word==3" {*@3!*}
 QS "Pos_C-Syl_in_C-Word==4" {*@4!*}
 QS "Pos_C-Syl_in_C-Word==5" {*@5!*}
 QS "Pos_C-Syl_in_C-Word==6" {*@6!*}
 QS "Pos_C-Syl_in_C-Word==7" {*@7!*}
 QS "Pos_C-Syl_in_C-Word==8" {*@8!*}
 QS "Pos_C-Syl_in_C-Word==9" {*@9!*}
 QS "Pos_C-Syl_in_C-Word==10" {*@10!*}

QS "Pos_C-Syl_in_C-Word<=1"
 {*@x!*,*@1!*}

QS "Pos_C-Syl_in_C-Word<=2"
 {*@x!*,*@1!*,*@2!*}

QS "Pos_C-Syl_in_C-Word<=3"
 {*@x!*,*@1!*,*@2!*,*@3!*}

QS "Pos_C-Syl_in_C-Word<=4"
 {*@x!*,*@1!*,*@2!*,*@3!*,*@4!*}

QS "Pos_C-Syl_in_C-Word<=5"
 {*@x!*,*@1!*,*@2!*,*@3!*,*@4!*,*@5!*}

QS "Pos_C-Syl_in_C-Word<=6"
 {*@x!*,*@1!*,*@2!*,*@3!*,*@4!*,*@5!*,*@6!*}

QS "Pos_C-Syl_in_C-Word<=7"
 {*@x!*,*@1!*,*@2!*,*@3!*,*@4!*,*@5!*,*@6!*,*@7!*}

QS "Pos_C-Syl_in_C-Word<=8"
 {*@x!*,*@1!*,*@2!*,*@3!*,*@4!*,*@5!*,*@6!*,*@7!*,*@8!*}

QS "Pos_C-Syl_in_C-Word<=9"
 {*@x!*,*@1!*,*@2!*,*@3!*,*@4!*,*@5!*,*@6!*,*@7!*,*@8!*,*@9!*}

QS "Pos_C-Syl_in_C-Word<=10"
 {*@x!*,*@1!*,*@2!*,*@3!*,*@4!*,*@5!*,*@6!*,*@7!*,*@8!*,*@9!*,*@10!*}

QS "Num-Syl_from_Prev-Stress-Syl<=0"
 {*!x-*,*!0-*}

QS "Num-Syl_from_Prev-Stress-Syl<=1"
 {*!x-*,*!0-*,*!1-*}

QS "Num-Syl_from_Prev-Stress-Syl<=2"
 {*!x-*,*!0-*,*!1-*,*!2-*}

QS "Num-Syl_from_Prev-Stress-Syl<=3"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*}

QS "Num-Syl_from_Prev-Stress-Syl<=4"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*,*!4-*}

QS "Num-Syl_from_Prev-Stress-Syl<=5"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*,*!4-*,*!5-*}

QS "Num-Syl_from_Prev-Stress-Syl<=6"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*,*!4-*,*!5-*,*!6-*}

QS "Num-Syl_from_Prev-Stress-Syl<=7"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*,*!4-*,*!5-*,*!6-*,*!7-*}

QS "Num-Syl_from_Prev-Stress-Syl<=8"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*,*!4-*,*!5-*,*!6-*,*!7-*,*!8-*}

QS "Num-Syl_from_Prev-Stress-Syl<=9"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*,*!4-*,*!5-*,*!6-*,*!7-*,*!8-*,*!9-*}

QS "Num-Syl_from_Prev-Stress-Syl<=10"
 {*!x-*,*!0-*,*!1-*,*!2-*,*!3-*,*!4-*,*!5-*,*!6-*,*!7-*,*!8-*,*!9-*,*!10-*}

QS "Num-Syl_from_Next-Stress-Syl==x"
 {*_x^*}

QS "Num-Syl_from_Next-Stress-Syl==0"
 {*_0^*}

QS "Num-Syl_from_Next-Stress-Syl==1"
 {*_1^*}

QS "Num-Syl_from_Next-Stress-Syl==2"
 {*_2^*}

QS "Num-Syl_from_Next-Stress-Syl==3"
 {*_3^*}

QS "Num-Syl_from_Next-Stress-Syl==4"
 {*_4^*}

QS "Num-Syl_from_Next-Stress-Syl==5"
 {*_5^*}

QS "Num-Syl_from_Next-Stress-Syl==6"
 {*_6^*}

QS "Num-Syl_from_Next-Stress-Syl==7"
 {*_7^*}

QS "Num-Syl_from_Next-Stress-Syl==8"
 {*_8^*}

QS "Num-Syl_from_Next-Stress-Syl==9"
 {*_9^*}

QS "Num-Syl_from_Next-Stress-Syl==10"
 {*_10^*}

QS "Num-Syl_from_Next-Stress-Syl<=0"
 {*_x^*,*_0^*}

QS "Num-Syl_from_Next-Stress-Syl<=1"
 {*_x^*,*_0^*,*_1^*}

QS "Num-Syl_from_Next-Stress-Syl<=2"
 {*_x^*,*_0^*,*_1^*,*_2^*}

QS "Num-Syl_from_Next-Stress-Syl<=3"
 {*_x^*,*_0^*,*_1^*,*_2^*,*_3^*}

QS "Num-Syl_from_Next-Stress-Syl<=4"
 {*_x^*,*_0^*,*_1^*,*_2^*,*_3^*,*_4^*}

QS "Num-Syl_from_Next-Stress-Syl<=5"
 {*_x^*,*_0^*,*_1^*,*_2^*,*_3^*,*_4^*,*_5^*}

QS "Num-Syl_from_Next-Stress-Syl<=6"
 {*_x^*,*_0^*,*_1^*,*_2^*,*_3^*,*_4^*,*_5^*,*_6^*}

QS "Num-Syl_from_Next-Stress-Syl<=7"
 {*_x^*,*_0^*,*_1^*,*_2^*,*_3^*,*_4^*,*_5^*,*_6^*,*_7^*}

QS "Num-Syl_from_Next-Stress-Syl<=8"
 {*_x^*,*_0^*,*_1^*,*_2^*,*_3^*,*_4^*,*_5^*,*_6^*,*_7^*,*_8^*}

QS "Num-Syl_from_Next-Stress-Syl<=9"
 {*_x^* ,*_0^* ,*_1^* ,*_2^* ,*_3^* ,*_4^* ,*_5^* ,*_6^* ,*_7^* ,*_8^* ,*_9^* }
 QS "Num-Syl_from_Next-Stress-Syl<=10"
 {*_x^* ,*_0^* ,*_1^* ,*_2^* ,*_3^* ,*_4^* ,*_5^* ,*_6^* ,*_7^* ,*_8^* ,*_9^* ,*_10^* }

 QS "C-Syl_Short_Vowel"
 {*_a/C: * ,*_e/C: * ,*_i/C: * ,*_o/C: * ,*_u/C: * }
 QS "C-Syl_Diphthong_Vowel"
 {*_aw/C: * ,*_ay/C: * ,*_ey/C: * ,*_oy/C: * }
 QS "C-Syl_High_Vowel"
 {*_i/C: * ,*_u/C: * }
 QS "C-Syl_Mid_Vowel"
 {*_e/C: * ,*_ey/C: * ,*_o/C: * }
 QS "C-Syl_Low_Vowel"
 {*_a/C: * ,*_aw/C: * ,*_ay/C: * ,*_oy/C: * }
 QS "C-Syl_Front_Vowel"
 {*_a/C: * ,*_e/C: * ,*_ey/C: * ,*_i/C: * }
 QS "C-Syl_Central_Vowel"
 {*_aw/C: * ,*_ay/C: * }
 QS "C-Syl_Back_Vowel"
 {*_o/C: * ,*_oy/C: * ,*_u/C: * }
 QS "C-Syl_Rounded_Vowel"
 {*_o/C: * ,*_oy/C: * ,*_u/C: * }
 QS "C-Syl_Unrounded_Vowel"
 {*_a/C: * ,*_aw/C: * ,*_ay/C: * ,*_e/C: * ,*_ey/C: * ,*_i/C: * }
 QS "C-Syl_AVowel"
 {*_a/C: * ,*_aw/C: * ,*_ay/C: * }
 QS "C-Syl_IVowel" {*_i/C: * }
 QS "C-Syl_UVowel" {*_u/C: * }
 QS "C-Syl_EVowel"
 {*_e/C: * ,*_ey/C: * }
 QS "C-Syl_OVowel"
 {*_o/C: * ,*_oy/C: * }
 QS "C-Syl_Vowel==a" {*_a/C: * }
 QS "C-Syl_Vowel==i" {*_i/C: * }
 QS "C-Syl_Vowel==u" {*_u/C: * }
 QS "C-Syl_Vowel==e" {*_e/C: * }
 QS "C-Syl_Vowel==o" {*_o/C: * }
 QS "C-Syl_Vowel==ay" {*_ay/C: * }
 QS "C-Syl_Vowel==aw" {*_aw/C: * }
 QS "C-Syl_Vowel==ey" {*_ey/C: * }
 QS "C-Syl_Vowel==oy" {*_oy/C: * }

 QS "R-Syl_Num-Segs==x" {*/C:x_* }
 QS "R-Syl_Num-Segs==1" {*/C:1_* }
 QS "R-Syl_Num-Segs==2" {*/C:2_* }
 QS "R-Syl_Num-Segs==3" {*/C:3_* }
 QS "R-Syl_Num-Segs==4" {*/C:4_* }
 QS "R-Syl_Num-Segs==5" {*/C:5_* }
 QS "R-Syl_Num-Segs==6" {*/C:6_* }
 QS "R-Syl_Num-Segs==7" {*/C:7_* }
 QS "R-Syl_Num-Segs==8" {*/C:8_* }

 QS "R-Syl_Num-Segs<=1"
 {*/C:x_* ,*/C:1_* }
 QS "R-Syl_Num-Segs<=2"
 {*/C:x_* ,*/C:1_* ,*/C:2_* }
 QS "R-Syl_Num-Segs<=3"
 {*/C:x_* ,*/C:1_* ,*/C:2_* ,*/C:3_* }

QS "R-Syl_Num-Segs<=4"
 {*/C:x_* ,*/C:1_* ,*/C:2_* ,*/C:3_* ,*/C:4_* }
 QS "R-Syl_Num-Segs<=5"
 {*/C:x_* ,*/C:1_* ,*/C:2_* ,*/C:3_* ,*/C:4_* ,*/C:5_* }
 QS "R-Syl_Num-Segs<=6"
 {*/C:x_* ,*/C:1_* ,*/C:2_* ,*/C:3_* ,*/C:4_* ,*/C:5_* ,*/C:6_* }
 QS "R-Syl_Num-Segs<=7"
 {*/C:x_* ,*/C:1_* ,*/C:2_* ,*/C:3_* ,*/C:4_* ,*/C:5_* ,*/C:6_* ,*/C:7_* }
 QS "R-Syl_Num-Segs<=8"
 {*/C:x_* ,*/C:1_* ,*/C:2_* ,*/C:3_* ,*/C:4_* ,*/C:5_* ,*/C:6_* ,*/C:7_* ,*/C:8_* }

 QS "R-Syl_Stress==0" {*_0/D: * }
 QS "R-Syl_Stress==1" {*_1/D: * }

 QS "L-Word_Num-Syl==1" {*/D:1/E: * }
 QS "L-Word_Num-Syl==2" {*/D:2/E: * }
 QS "L-Word_Num-Syl==3" {*/D:3/E: * }
 QS "L-Word_Num-Syl==4" {*/D:4/E: * }
 QS "L-Word_Num-Syl==5" {*/D:5/E: * }
 QS "L-Word_Num-Syl==6" {*/D:6/E: * }
 QS "L-Word_Num-Syl==7" {*/D:7/E: * }
 QS "L-Word_Num-Syl==8" {*/D:8/E: * }
 QS "L-Word_Num-Syl==9" {*/D:9/E: * }
 QS "L-Word_Num-Syl==10" {*/D:10/E: * }

 QS "L-Word_Num-Syl<=1" {*/D:1/E: * }
 QS "L-Word_Num-Syl<=2"
 {*/D:1/E: * ,*/D:2/E: * }
 QS "L-Word_Num-Syl<=3"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * }
 QS "L-Word_Num-Syl<=4"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * ,*/D:4/E: * }
 QS "L-Word_Num-Syl<=5"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * ,*/D:4/E: * ,*/D:5/E: * }
 QS "L-Word_Num-Syl<=6"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * ,*/D:4/E: * ,*/D:5/E: * ,*/D:6/E: * }
 QS "L-Word_Num-Syl<=7"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * ,*/D:4/E: * ,*/D:5/E: * ,*/D:6/E: * ,*/D:7/E: * }
 QS "L-Word_Num-Syl<=8"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * ,*/D:4/E: * ,*/D:5/E: * ,*/D:6/E: * ,*/D:7/E: * ,*/D:8/E: * }
 QS "L-Word_Num-Syl<=9"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * ,*/D:4/E: * ,*/D:5/E: * ,*/D:6/E: * ,*/D:7/E: * ,*/D:8/E: * ,*/D:9/E: * }
 QS "L-Word_Num-Syl<=10"
 {*/D:1/E: * ,*/D:2/E: * ,*/D:3/E: * ,*/D:4/E: * ,*/D:5/E: * ,*/D:6/E: * ,*/D:7/E: * ,*/D:8/E: * ,*/D:9/E: * ,*/D:10/E: * }

 QS "Pos_C-Word_in_C-Utterance==x"
 {*@x/F: * }
 QS "Pos_C-Word_in_C-Utterance==1"
 {*@1/F: * }

QS "Pos_C-Word_in_C-Utterance==2"
 {*@2/F:*}
 QS "Pos_C-Word_in_C-Utterance==3"
 {*@3/F:*}
 QS "Pos_C-Word_in_C-Utterance==4"
 {*@4/F:*}
 QS "Pos_C-Word_in_C-Utterance==5"
 {*@5/F:*}
 QS "Pos_C-Word_in_C-Utterance==6"
 {*@6/F:*}
 QS "Pos_C-Word_in_C-Utterance==7"
 {*@7/F:*}
 QS "Pos_C-Word_in_C-Utterance==8"
 {*@8/F:*}
 QS "Pos_C-Word_in_C-Utterance==9"
 {*@9/F:*}
 QS "Pos_C-Word_in_C-Utterance==10"
 {*@10/F:*}
 QS "Pos_C-Word_in_C-Utterance==11"
 {*@11/F:*}
 QS "Pos_C-Word_in_C-Utterance==12"
 {*@12/F:*}
 QS "Pos_C-Word_in_C-Utterance==13"
 {*@13/F:*}
 QS "Pos_C-Word_in_C-Utterance==14"
 {*@14/F:*}
 QS "Pos_C-Word_in_C-Utterance==15"
 {*@15/F:*}
 QS "Pos_C-Word_in_C-Utterance==16"
 {*@16/F:*}
 QS "Pos_C-Word_in_C-Utterance==17"
 {*@17/F:*}
 QS "Pos_C-Word_in_C-Utterance==18"
 {*@18/F:*}
 QS "Pos_C-Word_in_C-Utterance==19"
 {*@19/F:*}
 QS "Pos_C-Word_in_C-Utterance==20"
 {*@20/F:*}

QS "Pos_C-Word_in_C-Utterance<=1"
 {*@x/F:*,*@1/F:*}
 QS "Pos_C-Word_in_C-Utterance<=2"
 {*@x/F:*,*@1/F:*,*@2/F:*}
 QS "Pos_C-Word_in_C-Utterance<=3"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*}
 QS "Pos_C-Word_in_C-Utterance<=4"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*}
 QS "Pos_C-Word_in_C-Utterance<=5"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*}
 QS "Pos_C-Word_in_C-Utterance<=6"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*}
 QS "Pos_C-Word_in_C-Utterance<=7"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*}
 QS "Pos_C-Word_in_C-Utterance<=8"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*}
 QS "Pos_C-Word_in_C-Utterance<=9"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*}

QS "Pos_C-Word_in_C-Utterance<=10"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*}
 QS "Pos_C-Word_in_C-Utterance<=11"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*}
 QS "Pos_C-Word_in_C-Utterance<=12"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*}
 QS "Pos_C-Word_in_C-Utterance<=13"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*}
 QS "Pos_C-Word_in_C-Utterance<=14"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*,*@14/F:*}
 QS "Pos_C-Word_in_C-Utterance<=15"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*,*@14/F:*,*@15/F:*}
 QS "Pos_C-Word_in_C-Utterance<=16"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*,*@14/F:*,*@15/F:*,*@16/F:*}
 QS "Pos_C-Word_in_C-Utterance<=17"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*,*@14/F:*,*@15/F:*,*@16/F:*,*@17/F:*}
 QS "Pos_C-Word_in_C-Utterance<=18"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*,*@14/F:*,*@15/F:*,*@16/F:*,*@17/F:*,*@18/F:*}
 QS "Pos_C-Word_in_C-Utterance<=19"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*,*@14/F:*,*@15/F:*,*@16/F:*,*@17/F:*,*@18/F:*,*@19/F:*}
 QS "Pos_C-Word_in_C-Utterance<=20"
 {*@x/F:*,*@1/F:*,*@2/F:*,*@3/F:*,*@4/F:*,*@5/F:*,*@6/F:*,*@7/F:*,*@8/F:*,*@9/F:*,*@10/F:*,*@11/F:*,*@12/F:*,*@13/F:*,*@14/F:*,*@15/F:*,*@16/F:*,*@17/F:*,*@18/F:*,*@19/F:*,*@20/F:*}

QS "R-Word_Num-Syl==1" {*/F:1/G:*}
 QS "R-Word_Num-Syl==2" {*/F:2/G:*}
 QS "R-Word_Num-Syl==3" {*/F:3/G:*}
 QS "R-Word_Num-Syl==4" {*/F:4/G:*}
 QS "R-Word_Num-Syl==5" {*/F:5/G:*}
 QS "R-Word_Num-Syl==6" {*/F:6/G:*}

QS "R-Word_Num-Syll==7" {*/F:7/G:*}
 QS "R-Word_Num-Syll==8" {*/F:8/G:*}
 QS "R-Word_Num-Syll==9" {*/F:9/G:*}
 QS "R-Word_Num-Syll==10" {*/F:10/G:*}

 QS "R-Word_Num-Syll<=1" {*/F:1/G:*}
 QS "R-Word_Num-Syll<=2" {*/F:1/G:*,*/F:2/G:*}
 QS "R-Word_Num-Syll<=3" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*}
 QS "R-Word_Num-Syll<=4" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*,*/F:4/G:*}
 QS "R-Word_Num-Syll<=5" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*,*/F:4/G:*,*/F:5/G:*}
 QS "R-Word_Num-Syll<=6" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*,*/F:4/G:*,*/F:5/G:*,*/F:6/G:*}
 QS "R-Word_Num-Syll<=7" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*,*/F:4/G:*,*/F:5/G:*,*/F:6/G:*,*/F:7/G:*}
 QS "R-Word_Num-Syll<=8" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*,*/F:4/G:*,*/F:5/G:*,*/F:6/G:*,*/F:7/G:*,*/F:8/G:*}
 QS "R-Word_Num-Syll<=9" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*,*/F:4/G:*,*/F:5/G:*,*/F:6/G:*,*/F:7/G:*,*/F:8/G:*,*/F:9/G:*}
 QS "R-Word_Num-Syll<=10" {*/F:1/G:*,*/F:2/G:*,*/F:3/G:*,*/F:4/G:*,*/F:5/G:*,*/F:6/G:*,*/F:7/G:*,*/F:8/G:*,*/F:9/G:*,*/F:10/G:*}

 QS "Num-Sylls_in_Utterance==1" {*/G:1}
 QS "Num-Sylls_in_Utterance==2" {*/G:2}
 QS "Num-Sylls_in_Utterance==3" {*/G:3}
 QS "Num-Sylls_in_Utterance==4" {*/G:4}
 QS "Num-Sylls_in_Utterance==5" {*/G:5}
 QS "Num-Sylls_in_Utterance==6" {*/G:6}
 QS "Num-Sylls_in_Utterance==7" {*/G:7}
 QS "Num-Sylls_in_Utterance==8" {*/G:8}
 QS "Num-Sylls_in_Utterance==9" {*/G:9}
 QS "Num-Sylls_in_Utterance==10" {*/G:10}
 QS "Num-Sylls_in_Utterance==11" {*/G:11}
 QS "Num-Sylls_in_Utterance==12" {*/G:12}
 QS "Num-Sylls_in_Utterance==13" {*/G:13}
 QS "Num-Sylls_in_Utterance==14" {*/G:14}
 QS "Num-Sylls_in_Utterance==15" {*/G:15}
 QS "Num-Sylls_in_Utterance==16" {*/G:16}
 QS "Num-Sylls_in_Utterance==17" {*/G:17}
 QS "Num-Sylls_in_Utterance==18" {*/G:18}
 QS "Num-Sylls_in_Utterance==19" {*/G:19}
 QS "Num-Sylls_in_Utterance==20" {*/G:20}
 QS "Num-Sylls_in_Utterance==21" {*/G:21}
 QS "Num-Sylls_in_Utterance==22" {*/G:22}
 QS "Num-Sylls_in_Utterance==23" {*/G:23}
 QS "Num-Sylls_in_Utterance==24" {*/G:24}
 QS "Num-Sylls_in_Utterance==25" {*/G:25}
 QS "Num-Sylls_in_Utterance==26" {*/G:26}
 QS "Num-Sylls_in_Utterance==27" {*/G:27}
 QS "Num-Sylls_in_Utterance==28" {*/G:28}

QS "Num-Sylls_in_Utterance==29" {*/G:29}
 QS "Num-Sylls_in_Utterance==30" {*/G:30}
 QS "Num-Sylls_in_Utterance==31" {*/G:31}
 QS "Num-Sylls_in_Utterance==32" {*/G:32}
 QS "Num-Sylls_in_Utterance==33" {*/G:33}
 QS "Num-Sylls_in_Utterance==34" {*/G:34}
 QS "Num-Sylls_in_Utterance==35" {*/G:35}
 QS "Num-Sylls_in_Utterance==36" {*/G:36}
 QS "Num-Sylls_in_Utterance==37" {*/G:37}
 QS "Num-Sylls_in_Utterance==38" {*/G:38}
 QS "Num-Sylls_in_Utterance==39" {*/G:39}
 QS "Num-Sylls_in_Utterance==40" {*/G:40}
 QS "Num-Sylls_in_Utterance==41" {*/G:41}
 QS "Num-Sylls_in_Utterance==42" {*/G:42}
 QS "Num-Sylls_in_Utterance==43" {*/G:43}
 QS "Num-Sylls_in_Utterance==44" {*/G:44}
 QS "Num-Sylls_in_Utterance==45" {*/G:45}
 QS "Num-Sylls_in_Utterance==46" {*/G:46}
 QS "Num-Sylls_in_Utterance==47" {*/G:47}
 QS "Num-Sylls_in_Utterance==48" {*/G:48}
 QS "Num-Sylls_in_Utterance==49" {*/G:49}
 QS "Num-Sylls_in_Utterance==50" {*/G:50}

 QS "Num-Sylls_in_Utterance<=1" {*/G:x,*/G:1}
 QS "Num-Sylls_in_Utterance<=2" {*/G:x,*/G:1,*/G:2}
 QS "Num-Sylls_in_Utterance<=3" {*/G:x,*/G:1,*/G:2,*/G:3}
 QS "Num-Sylls_in_Utterance<=4" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4}
 QS "Num-Sylls_in_Utterance<=5" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5}
 QS "Num-Sylls_in_Utterance<=6" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6}
 QS "Num-Sylls_in_Utterance<=7" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7}
 QS "Num-Sylls_in_Utterance<=8" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8}
 QS "Num-Sylls_in_Utterance<=9" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8,*/G:9}
 QS "Num-Sylls_in_Utterance<=10" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8,*/G:9,*/G:10}
 QS "Num-Sylls_in_Utterance<=11" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8,*/G:9,*/G:10,*/G:11}
 QS "Num-Sylls_in_Utterance<=12" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8,*/G:9,*/G:10,*/G:11,*/G:12}
 QS "Num-Sylls_in_Utterance<=13" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8,*/G:9,*/G:10,*/G:11,*/G:12,*/G:13}
 QS "Num-Sylls_in_Utterance<=14" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8,*/G:9,*/G:10,*/G:11,*/G:12,*/G:13,*/G:14}
 QS "Num-Sylls_in_Utterance<=15" {*/G:x,*/G:1,*/G:2,*/G:3,*/G:4,*/G:5,*/G:6,*/G:7,*/G:8,*/G:9,*/G:10,*/G:11,*/G:12,*/G:13,*/G:14,*/G:15}

LAMPIRAN D. Question Set 4 untuk Kombinasi Konteks (Sakti, 2008) dan (Vania, 2011)

D.I. Bagian set pertanyaan terkait konteks fonetis

Sama dengan lampiran A bagian I.

D.II. Bagian set pertanyaan terkait konteks posisi, jumlah, dsb.

Gabungan lampiran B bagian II dan lampiran C bagian II.

LAMPIRAN E. Daftar Kalimat yang Digunakan pada Evaluasi

MOS

1. Menurutnya, kegiatan tersebut bisa dinikmati oleh semua warga masyarakat.
2. Parade GNC tahun ini menghadirkan sesuatu yang berbeda dari sebelumnya.
3. Dia lantas mencontohkan korupsi yang didorong rasa keserakahan.
4. Pekan lalu, Sharapova juara di Madrid dengan mengalahkan Simona Halep di final.
5. Ferguson tampak duduk berdampingan dengan pelatih Juventus, Antonio Conte.
6. Sebelumnya festival tersebut selalu digelar di Pulau Jawa, seperti Jakarta dan Bandung.
7. Saat ini belasan senjata rakitan sudah diamankan di Polres Klaten.
8. Padahal di negara tersebut angka kejadian penyakit jantungnya relatif rendah.
9. Namun, warga belum bisa memastikan jenis hewan piaraan yang dimakan ular tersebut.
10. Kalau dilihat secara tampilan, ubahan desain memang tidak tergolong revolusioner.
11. Boro-boro mau ibadah, berkumpul saja tidak boleh.
12. Saya rasa hal tersebut membawa perubahan.
13. Oleh karena itu, KPK menerapkan skala prioritas.
14. Serangan udara sudah banyak dilakukan, terutama di media sosial.
15. Van Gaal masih sedang membentuk staf timnya di ruang ganti.
16. Karena sekarang ini masih konsep, masih mimpi-mimpi.
17. Menurut dia, program ini bersifat struktural, dari atas sampai bawah.
18. Sekarang dia malah sibuk tebar pesona ke sana kemari.
19. Wardah sebagai brand komestik saat ini sedang naik daun.
20. Nahrowi menambahkan, tersangka telah diobati, tetapi belum siap diperiksa hingga Kamis sore.
21. Wisatawan bisa melihat langsung obyek wisata di Labuan Bajo.
22. Mourinho melontarkan pernyataan tersebut pada awal paruh kedua musim ini.
23. Hingga saat ini kami masih melakukan proses perbaikan tersebut.
24. Ratusan kerabat penumpang kapal sudah berada di tepi sungai, tempat terjadinya kecelakaan tersebut.
25. Sebelum menghadiri pertunjukan wayang tersebut, Fadli Zon mengaku baru saja melakukan makan malam bersama Rachmawati.
26. Buah-buahan merupakan sumber serat, vitamin, dan juga antioksidan.
27. Kualitas vokal Nowela bahkan tak terganggu dengan insiden kecil di panggung.
28. Ponsel ini ditujukan bagi tuna netra dan pengguna lain yang memiliki keterbatasan visual.
29. Anjloknya pertumbuhan harga rumah juga terjadi di Manado, Sulawesi Utara.
30. Berangkat pukul lima pagi, lupa pake sarung tangan.

LAMPIRAN F. Daftar Kalimat yang Digunakan pada Evaluasi SUS

I. Intransitive (noun+verb(intr.)+preposition+noun)

1. bumi menjerit untuk dokumen
2. listrik bercerita kepada obat
3. baja pergi dari toilet
4. gula berjalan di gambut
5. hukum mengigau ke masyarakat
6. upacara bervariasi dalam wanita

II. Transitive (noun+adjective+verb(trans.)+noun)

1. kuliah pengap menudungi hakim
2. visa akuatik mencolek bintang
3. ikan kaya melalaikan murid
4. sungai seimbang memeriksa wasit
5. contoh makmur memegang kodok
6. yayasan absurd membalut yunani

III. Imperative (verb(no affix)+noun+conjunction+noun)

1. kepung benang dan ranjang
2. percaya retina sama susu
3. buka dampak serta kalsium
4. jambak naga dengan protein
5. tumpuk unggas atau adonan
6. cari zakat tanpa debu

IV. Interrogative (question word+noun+verb(trans.)+noun)

1. mengapa cahaya memaklumi ongkos
2. dimana patung merindukan status
3. kapan timah menguji visi

4. bagaimana wibawa meyakinkan xenon
5. apakah zodiak mendaratkan elektronik
6. akankah eksploitasi memfasilitasi gedung

V. *Relative (noun+verb(trans.)+noun+relative pronoun+verb(intr.))*

1. isyarat membengkokkan cacing yang berdalil
2. ekologi memfokuskan fakultas yang bergegas
3. harimau mengimpor jabatan yang berkabut
4. lubang memonopoli niat yang mengompol
5. paguyuban meringkus saham yang bertualang
6. koperasi melahirkan monyet yang menganga