

## PEMBANGKIT UCAPAN MODEL ARTIKULATORI

Arif B.Putra N<sup>1</sup>, Arry Akhmad Arman<sup>2</sup>, Kuspriyanto<sup>2</sup>

<sup>1</sup>Program Studi Teknik Informatika, Fakultas Teknik, Universitas Tanjungpura  
Jl. A. Yani Pontianak, 78124

<sup>2</sup>Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung  
Jl. Ganesha 10, 40132

E-mail: ariputra@yahoo.com, arman@kupalima.com

### ABSTRAKS

Usaha untuk menghasilkan ucapan yang natural dan inteligibilitas yang baik oleh mesin pembangkit ucapan terus dilakukan, diantaranya dengan memodelkan alat ucap manusia atau *Articulatory Synthesis*. Ketertarikan peneliti pada bidang ini dikarenakan dapat mengurangi ruang memori dan kebutuhan bandwidth untuk penyimpanan dan transmisi sinyal ucapan yang dibangkitkan serta dapat meningkatkan proses pengenalan ucapan dengan cara transisi ke domain artikulatoris, di mana sinyal dapat dicirikan oleh parameter yang lebih sedikit.

Penelitian *speech synthesizer* menggunakan metoda *Articulatory synthesis* untuk menghasilkan ucapan bahasa Indonesia masih belum banyak dilakukan. Penelitian pendahuluan ini melakukan kajian pustaka mengenai pembangkit ucapan pada model artikulatori Mermelstein's, dimana untuk mengembangkan model artikulator diperlukan pengetahuan yang mendalam tentang akustik, mekanika, fisiologi, linguistik, fonetik, dan pemrosesan sinyal pada umumnya diperlukan.

Diharapkan penelitian dibidang ini dapat memberikan kontribusi bagi ilmu pengetahuan khususnya pada bidang *articulatory synthesis* bahasa Indonesia.

*Kata Kunci: articulatory, speech, synthesis, artikulatori, ucapan, sintesa*

## 1. PENDAHULUAN

### 1.1 Latar Belakang

Alat-alat ucap adalah bagian utama tubuh untuk menghasilkan dan membedakan bunyi-bunyi bahasa. Alat-alat ucap kita seperti bibir, gigi, lidah, rahang, dan pangkal tenggorokan dapat dilihat dengan foto sinar X atau MRI. Dengan alat ini setiap proses dan posisi pergerakan alat ucap manusia untuk menghasilkan bunyi ucapan dapat terlihat, sehingga menimbulkan ketertarikan untuk melakukan penelitian dibidang ini terutama pada perubahan dimensi alat ucap dan pengaruhnya terhadap bunyi yang dihasilkan.

Teknik pengembangan yang dilakukan oleh para peneliti untuk menghasilkan ucapan dilakukan dengan memodelkan bentuk fisik alat-alat ucap manusia yang disebut dengan sintesa artikulatori (*articulatory synthesis*). Sintesa artikulatori merupakan teknik komputasi untuk mensintesis ucapan dari model sistem ucapan manusia dan proses artikulasi yang terlibat dalam aktivitas itu. Bentuk sistem ucapan dapat dikontrol dalam beberapa cara, biasanya dengan memodifikasi posisi artikulator-artikulator ucapan, seperti lidah, rahang, dan bibir. Himpunan variabel artikulator-artikulator ucapan ini dinamakan parameter artikulatori. Dengan mensimulasikan aliran udara secara digital berdasarkan representasi sistem ucapan manusia maka dihasilkan ucapan.

Untuk menghasilkan sintesa ucapan, vokal dan konsonan agar dapat dikenali dilakukan dengan membangkitkan dan menggabungkan gelombang-

gelombang bunyi pada sejumlah kecil frekuensi yang berbeda-beda yang dianggap penting untuk setiap bunyi. Ada banyak komponen lain yang terlibat dalam gelombang bunyi ucapan sesungguhnya. Beberapa diantaranya belum dimengerti sepenuhnya, tetapi hakikat gelombang bunyi yang pokok dan penting untuk menyampaikan ucapan sekarang sudah diketahui dengan jelas

Pensintesa ucapan (*speech synthesis*) yaitu produksi ucapan buatan yang menyerupai ucapan manusia. Pembangkit ucapan (*speech synthesizer*) adalah sistem komputer dengan tujuan memproduksi ucapan buatan ini, dan dapat diimplementasikan ke perangkat lunak ataupun perangkat keras. Kualitas pembangkit ucapan dinilai dari kenaturalan dan inteligibilitasnya

Keuntungan utama pendekatan sintesa ucapan artikulatoris adalah bahwa :

- Model terkait langsung dengan produksi ucapan manusia, sehingga parameter model bervariasi perlahan, dan mudah diinterpolasi;
- Interaksi sumber-saluran ucap dimodelkan dengan alami.

Bunyi setiap bahasa berbeda sehingga untuk representasikan bunyi tiap bahasa pun berbeda pula. Ucapan bunyi bahasa dituliskan dalam bentuk fonem. Simbol yang digunakan untuk menuturkan bunyi tiap bahasa berbeda maka fonem yang digunakan untuk membangun bunyi tiap bahasa menjadi berbeda. Sehingga jika orang Inggris mengucapkan kata atau kalimat dalam bahasa Indonesia akan terdengar berbeda, hal ini karena

fonem yang digunakan untuk membangun kata atau kalimat yang berbeda pada bahasa Inggris dan bahasa Indonesia.

Penelitian pembangkit ucapan menggunakan metoda sintesa artikulatori untuk menghasilkan ucapan bahasa Indonesia masih belum banyak dilakukan. Penelitian pendahuluan ini melakukan kajian pustaka mengenai pembangkit ucapan bahasa Indonesia pada model artikulator. Dimana untuk mengembangkan model artikulator ini diperlukan pengetahuan yang mendalam tentang akustik, mekanika, fisiologi, linguistik, fonetik, dan pemrosesan sinyal.

Diharapkan penelitian ini dapat memberikan kontribusi bagi ilmu pengetahuan khususnya pada bidang sintesa artikulatori bahasa Indonesia.

## 1.2 Tujuan Penelitian

Penelitian ini melakukan analisis model artikulatori, dan parameter yang diperlukan untuk membangkitkan ucapan fonem-fonem.

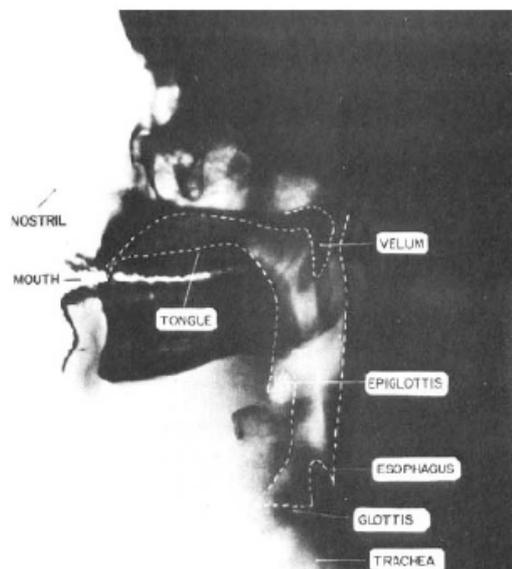
## 2. ALAT UCAP MANUSIA

Proses produksi ucapan manusia terdiri dari proses pembentuk aliran udara dari paru-paru, pengubahan aliran udara dari paru-paru menjadi suara *voice* dan *unvoice*, dan proses artikulasi atau proses modulasi pengaturan suara yang terdiri dari bunyi yang spesifik.

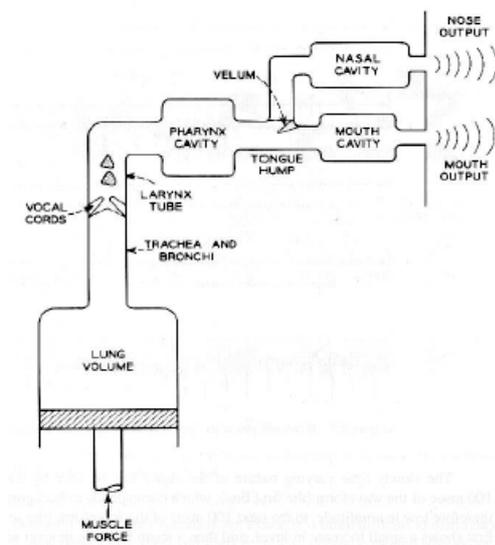
Foto sinar X alat ucap manusia diperlihatkan gambar 1. *Vocal tract* ditandai garis putus-putus, dimulai dari *vocal cords* atau *glottis*, dan berakhir pada mulut. *Vocal tract* terdiri dari *pharynx* (koneksi antara esophagus dengan mulut) dan mulut atau *oral cavity*. Pada rata-rata pria, ukuran total *vocal tract* sekitar 17 cm. Daerah pertemuan *vocal tract* ditentukan oleh posisi lidah, bibir, rahang, dan bagian belakang langit-langit; luasnya berkisar antara 0 (ketika seluruhnya tertutup) hingga sekitar 20 cm<sup>2</sup>. *Nasal tract* dimulai dari bagian belakang langit-langit lunak dan berakhir di *nostrils*. Ketika bagian belakang langit-langit lunak (organ yang memiliki fungsi sebagai pintu penghubung antara *vocal tract* dan *nasal track*) terbuka maka secara akustik *nasal track* akan bergandengan dengan *vocal track* untuk menghasilkan suara nasal.

Aliran udara yang dihasilkan oleh dorongan otot paru-paru bersifat konstan. Ketika pita suara berkontraksi maka aliran udara yang lewat akan bergetar. Aliran udara tersebut terpotong-potong oleh gerakan pita suara menjadi sinyal pulsa *quasi-periodic*. Sinyal ini mengalami modulasi frekuensi ketika melewati *pharynx*, rongga mulut maupun rongga hidung. Sinyal yang dihasilkan oleh prose ini disebut sinyal *voiced*. Tetapi jika pita suara pada keadaan relaksasi, maka aliran udara berusaha melewati celah sempit pada permukaan *vocal track* sehingga alirannya

menjadi turbulen. Proses ini menghasilkan sinyal *unvoiced*. Representasi sederhana dari mekanisme fisiologis utuh untuk menghasilkan ucapan di perlihatkan gambar 2.



Gambar 1. Foto sinar x penampang alat-alat ucap manusia [Rabiner,93]



Gambar 2. Model sistem produksi ucapan manusia [Rabiner,93]

Saat sinyal suara melalui *vocal tract*, maka kandungan frekuensi mengalami modulasi sehingga terjadi resonansi di *vocal track* yang disebut *formants*. Jika sinyal yang dihasilkan adalah *voiced* maka di selang waktu yang singkat bentuk *vocal track* cenderung konstan sehingga bentuk *vocal track* dapat diperkirakan dari bentuk *spectral* sinyal *voiced*.

### 3. SINTESA UCAPAN ARTIKULATORI

Pada dasarnya, ada tiga pendekatan yang digunakan dalam sintesa ucapan artikulatori. [Childers,2000].

Pendekatan pertama, Gelombang Tapis Digital. Pendekatan ini didasarkan pada maju dan mundur perjalanan gelombang dalam tabung akustik yang dapat menghasilkan sintesa *real-time*.

Pendekatan kedua, menggunakan Gabungan Domain Time-Frequency. Pendekatan ini memodelkan karakteristik *glotal* yang sangat nonlinear dalam domain waktu dan karakteristik getaran dinding vocal dalam domain frekuensi.

Pendekatan ketiga adalah memodelkan sistem vokal manusia sebagai himpunan perbedaan persamaan linier dan nonlinier yang besar yang harus dipecahkan di setiap interval sampling untuk memberikan sampel tekanan dan kecepatan volume pada setiap titik jalur transmisi vokal.

#### 3.1 Kriteria Sintesa Artikulatori

Tujuan utama sistem sintesa artikulatori adalah menghasilkan sinyal akustik yang menyerupai suara manusia dengan kualitas tingkat akurasi yang tinggi. Untuk mencapai tujuan tersebut, berikut ada empat kriteria yang harus dipenuhi sistem sintesa artikulator [khalis et.al, 2003].

##### a. Akurasi Konfigurasi Parameter Artikulator Statis

Sistem sintesa artikulatori harus menghasilkan konfigurasi statis parameter-parameter artikulator yang sesuai dengan konfigurasi tiap penutur berbeda. Ini tidak berarti bahwa sistem ini harus dapat mereproduksi anatomi saluran vokal untuk setiap saluran vokal penutur, melainkan untuk memparameterkan beberapa konfigurasi artikulator dengan sesedikit mungkin, tetapi masih dapat menggambarkan variabel konfigurasi parameter yang spesifik.

##### b. Akurasi Gerakan Dinamis

Sistem ini juga harus mampu mereproduksi secara akurat konfigurasi perpindahan dari artikulator-artikulator antara dua target fonetis. Hal ini sangat penting, karena pemeriksaan spektrogram bahasa alami menunjukkan bahwa sebagian besar dari ujaran ucapan terdiri dari variasi dinamis, bukan stabil.

##### c. Kemampuan dikonfigurasi

Parameter kontrol model artikulatoris harus dapat dikonfigurasi untuk menyesuaikan variasi biologis alami dari saluran alat ucapan manusia. Variasi pembicara tentunya dihasilkan pada sinyal akustik, tetapi variasi pada tingkat tersebut memberikan konsekuensi pada variasi ukuran saluran vokal dan kekhasannya. Variasi pada satu parameter artikulatori secara bersamaan dapat mempengaruhi beberapa parameter akustik di segmen yang berbeda, misalnya bentuk bibir akan mempengaruhi formant vokal, spektrum suara frikatif, dan spektrum letupan.

##### d. Parameter untuk Kendali Ragam Bahasa

Model pengendalian parameter dibutuhkan untuk sedekat mungkin dengan karakterisasi suatu bahasa, misalnya kita harus mampu menentukan tempat artikulasi dan parameter artikulator secara langsung. Hal ini menimbulkan *tradeoff* antara karakterisasi geometri dan linguistik dari parameter sintesa artikulatori. Akurasi pemodelan geometrik bentuk organ ucapan memerlukan parameter seperti jarak dan lokasi pusat lingkaran untuk mencirikan bentuk lidah, namun spesifikasi linguistik vokal biasanya dilakukan pada ciri artikulatori.

#### 3.2 Proses Sintesa Ucapan Artikulator

Ada empat tahapan proses untuk menghasilkan ucapan dari pembangkit ucapan artikulatori seperti diperlihatkan pada gambar 3. yaitu.

##### a. Analisis

Tahap analisis yaitu mengekstrak formant target untuk file ucapan dengan menentukan saluran formant dari sinyal ucapan target. Selanjutnya menandai lintasan formant target pada interval yang diinginkan (frame). Lintasan formant file ini ditandai dan disimpan sebagai target file formant yang merepresentasikan fonem tertentu.

##### b. *Speech Invers Filtering*

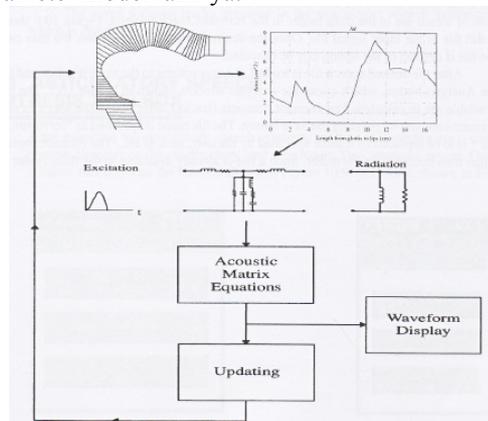
*Speech Invers Filtering* dilakukan untuk menentukan parameter model artikulatori. Untuk mendapatkan parameter model artikulatori menggunakan algoritma *simulated annealing* untuk memperkecil jarak (kesalahan) antara formant target dan formant model.

##### c. Penentuan Eksitasi

Proses selanjutnya yaitu menentukan jenis eksitasi yang digunakan untuk sintesa. Proses ini berkaitan dengan pembagian frame pada proses analisis sebelumnya.

##### d. Sintesa

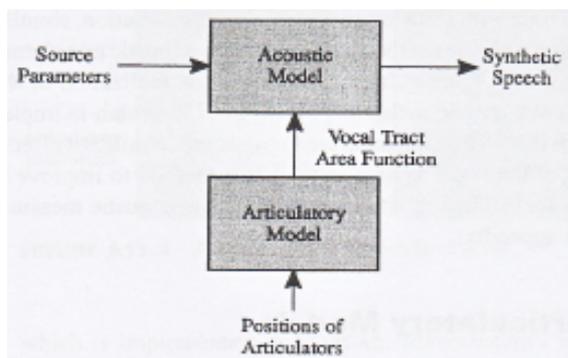
Hasil parameter model artikulatori pada proses *speech invers filtering* sebagai masukan untuk proses sintesa ini. Keluarannya menghasilkan tampilan animasi dari konfigurasi saluran suara yang digunakan untuk sintesa serta menampilkan parameter model lainnya.



Gambar 3. Langkah-langkah sintesa ucapan artikulatori [Childers,2000]

#### 4. PEMODELAN ARTIKULATORI

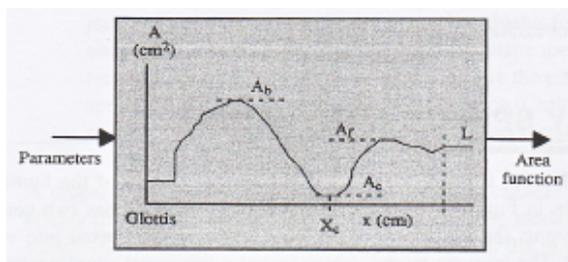
Pembangkit ucapan artikulatori didasarkan pada model fisiologi dari proses produksi ujaran manusia. Seperti ditunjukkan dalam Gambar 4, *synthesizer* artikulatori memiliki dua komponen. Model artikulatori menggambarkan posisi artikulator, yang dikonversi ke sistem vokal fungsi daerah lintasan. Model akustik, yang meliputi *couplin-subglottal*, interkasi sumber-saluran, saluran vokal, saluran hidung dengan rongga *sinus*, dan radiasi akustik, mensimulasikan produksi ucapan dan propagasi secara fisik transformasi fisiologis-ke-akustik.



Gambar 4. Model sintesa ucapan artikulatori [Childers,2000]

##### a. Model Area Parametric

Model *Area parametric* bukan menunjukkan posisi artikulatori secara langsung, melainkan menunjukkan pemodelan fungsi kawasan sebagai fungsi dari jarak di sepanjang saluran utama dengan batasan tertentu. Sebuah ciri umum dari model ini adalah spesifikasi dari penyempitan daerah minimum dan lokasi aksial nya. Wilayah alat suara biasanya diwakili oleh fungsi kontinu seperti hiperbola, parabola, atau sinusoida. Artikulasi konsonan umumnya belum diimplementasikan. Gambar 5 menunjukkan salah satu contoh dari model daerah parametrik.

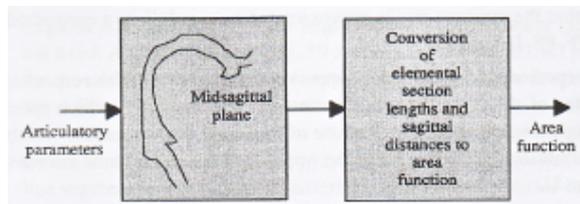


Gambar 5. Model daerah parametrik [Childers,2000]

##### b. Model jarak *midsagittal*

Model Jarak *midsagittal* biasanya didasarkan pada representasi bidang *midsagittal* seperti terlihat dari gambar sinar x. Penggambaran gerakan artikulator ucapan di bidang *midsagittal* membutuhkan spesifikasi posisi artikulator-artikulator atau aturan untuk mengontrol gerakan

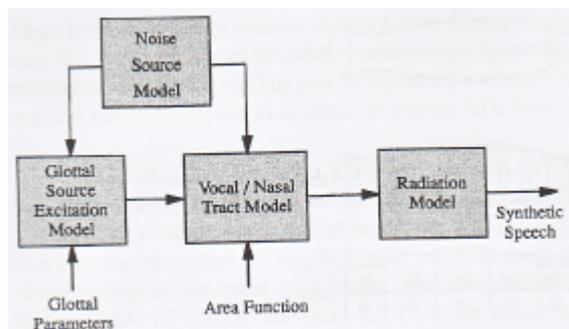
artikulator-artikulator. Keluaran model ini adalah perkiraan daerah batasan pergerakan alat ucap di saluran vokal. Visualisasi dan interpretasi artikulatori adalah keuntungan utama dari model ini. Gambar 6 memperlihatkan model jarak *midsagittal*



Gambar 6. Model jarak *midsagittal*, [Childers,2000]

#### 4.1 Model Akustik

Pada dasarnya model akustik sistem vokal manusia diwujudkan dalam beberapa submodel yang diperlihatkan pada gambar 7. Model *vocal track* dan *nasal track* mensimulasikan propagasi/perambatan suara pada saluran model ini. Model *excitation source* merepresentasikan dan membangkitkan bentuk gelombang eksitasi suara pada saluran vokal. Letupan aliran *turbulent* udara bergolak menghasilkan bunyi desah yang dihasilkan dari model *noise source*. Model radiasi mensimulasikan radiasi energi akustik dari bibir dan lubang hidung.



Gambar 7. Model akustik [Childers,2000]

##### a. Model *Vocal Track* (saluran alat ucap)

Model saluran alat ucap dapat dinyatakan sebagai tabung lurus dengan luas daerah yang berbeda-beda pada tiap titik tabung (cross sectional area). Perubahan luas daerah dalam tabung ini mempengaruhi perambatan suara dalam alat ucap.

##### b. Model *Nasal Tract* dan Rongga Hidung

Model *nasal tract* merupakan percabangan pada sisi pergerakan alat ucap. Velum digunakan untuk mengontrol hubungan antara *vocal track* dan *nasal tract*.

##### c. Model Radiasi Mulut dan Hidung

Energi akustik dilepaskan dari *vocal track* melalui mulut. Dari analogi saluran transmisi, mulut menghasilkan impedansi radiasi pada *vocal track*. Impedansi radiasi terdiri dari resistansi yang

dinyatakan sebagai loss energi akustik dan reaktansi dinyatakan sebagai massa inersia udara dimulut. Model radiasi yang sama berlaku juga pada hidung.

#### d. Model *Source Excitation*

Pada dasarnya terdapat dua jenis suara yaitu voiced(bersuara) yang meliputi vibrasi *quasi-periodic* dari pita suara, dan unvoiced(tak bersuara) yang meliputi pembakitan dilakukan pada dan voiceless(tak bersuara) yang melibatkan pembangkitan turbulensi noise dikarenakan cepatnya aliran udara melewati batasan yang sempit. Untuk ucapan bersuara sumber eksitasi adalah rentetan pulsa *quasi-periodic* pada celah udara.

#### e. Model Impedansi *glottal* dan Model *subglottal*

Sistem subglotal mengikuti *glottal*, dimana ketika daerah *glottal* mengecil maka impedansi *glottal* cenderung tinggi. Sedangkan pengaruh pada fungsi transfer akustik dapat diabaikan. Ketika daerah *glottal* membesar, maka bandwidth cenderung meningkat.

#### f. Model sumber *noise*

Pada dasarnya, model sumber *noise* merupakan karakteristik sumber *noise* sebagai fungsi aliran udara yang melewati daerah yang dibatasi. Jika aliran udara melewati daerah sempit atau dihalangi, maka akan terbentuklah turbulen. Ada tiga tipe konsonan yang dihasilkan oleh keadaan ini yaitu frikatif, stop, dan afrikatif.

## 4.2 Parameter Model Artikulatori

Model artikulatori dalam analisis ini adalah model Mermelstein's (1973), karena model ini memberikan kecocokan antara sinar x rekaman dan garis besar saluran vokal *midsagittal*. [Mermelstein, 1973].

Model artikulatori digunakan untuk mentransformasi parameter artikulatori menjadi representasi vektor dari *vocal tract cross-section* yang kemudian berubah menjadi karakteristik akustik dalam alat suara.

Pada model artikulatori, sekelompok variabel digunakan untuk mengatur bentuk dari saluran vokal. Parameter-parameter yang dapat dilihat pada Gambar 7 tersebut adalah sebagai berikut.

#### a. Badan lidah

Badan lidah direpresentasikan oleh busur (DL-B) dari sebuah lingkaran dengan titik pusat yang dapat bergerak dan jari-jari tetap. Pusat dari badan lidah, yang disimbolkan dengan tongc, memiliki koordinat polar (sc, thetaj+thetab) yang berpusat pada titik F. Meskipun demikian, koordinat kartesian (tbodyx, tbodyy) digunakan dalam tampilan dan optimasi.

#### b. Ujung lidah

Ujung lidah direpresentasikan oleh koordinat kartesian (tipx, tipy) dari titik T. Busur B-T dan TPF memberi bentuk dari bagian depan lidah. Oleh karena letak titik B bervariasi tergantung pada pusat badan lidah (tongc) dan sudut rahang (JAW), pergerakan bagian depan lidah tergantung pada badan lidah dan posisi rahang.

#### c. Rahang

Titik JAW dengan koordinat polar (sj, thetaj) digunakan untuk merepresentasikan letak rahang. Jarak sj tetap untuk sebagian besar fonem. Parameter rahang digunakan untuk menyatakan sudut dari thetaj. Perhatikan bahwa lekukan rahang didekati dengan beberapa segmen garis yang berhubungan (PF-PS-JAW-L6).

#### d. Bibir

Bibir direpresentasikan oleh titik L5 (atas) dan L7 (bawah). Dengan mengacu pada titik JAW, koordinat dari bibir bawah direpresentasikan oleh (lipp, lipo), yang memberikan keterangan mengenai protrusi bibir dan bukaan bibir. Penggunaan lipp dan lipo sebagai variabel yang terpisah memungkinkan representasi dari bibir yang terkutup, bibir yang terbuka, dan bibir yang membulat. Bibir atas L5 memiliki koordinat yang sama dengan mengacu pada titik U.

#### e. Hyoid

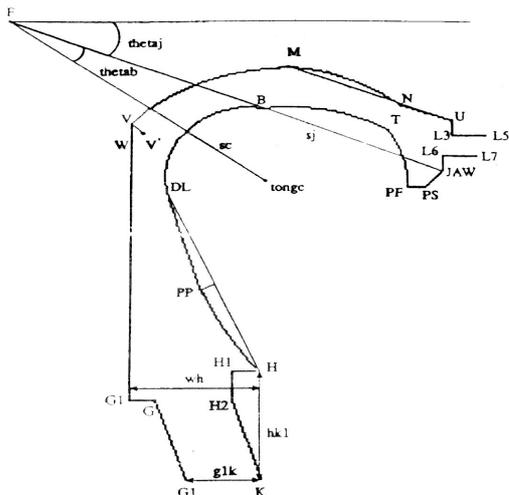
Hyoid direpresentasikan oleh parameter hyoid, yaitu jarak dari titik PP ke garis H-DL. Titik PP terdapat pada titik tengah dari segmen garis H-DL, yang merupakan garis singgung dari busur badan lidah pada titik DL. Segmen garis DL-PP dan busur PP-H, serta badan lidah, menentukan bagian depan dari faring. Titik H merepresentasikan perpotongan dari bagian depan epiglottis dan bagian atas tulang hyoid. Titik K merepresentasikan perkiraan dari batas bagian depan dari laring.

#### f. Bagian atas dari saluran vokal

Direpresentasikan oleh letak gigi atas, U, busur langit-langit rongga mulut UN-M (hard palate), titik tertinggi pada maxilla M, busur langit-langit rongga mulut M-V (soft palate), letak bagian belakang langit-langit rongga mulut (velum) V, letak dinding belakang faring W, dan titik tertinggi dari perarytenoid G. Pada busur hard palate, titik N terletak pada segmen garis M-U sedemikian rupa sehingga jarak M-N adalah dua kali jarak N-U. Busur lingkaran M-V dan M-N memiliki pusat yang terletak pada garis vertikal melalui M. Secara umum, bentuk bagian atas dan bagian belakang dianggap tetap, kecuali untuk busur soft palate yang berada dekat dengan titik V. Untuk memberikan keterangan mengenai area bukaan velopharyngeal port, bagian belakang langit-langit rongga mulut (velum) menjadi sebuah parameter artikulatori.

#### g. Bagian belakang langit-langit rongga mulut (velum)

Kondisi bagian belakang langit-langit rongga mulut (velum) direpresentasikan oleh letak V dari ujung uvula yang bergerak pada segmen garis V-V'. Area bukaan velar diasumsikan proporsional terhadap jarak antara titik V dan titik tertinggi dari velum.



Gambar 7. Model artikulator Marmelstein  
[Marmelstein,1972]

**5. BUNYI VOKAL BAHASA INDONESIA**

Vokal adalah jenis buyi bahasa yang ketika dihasilkan atau diproduksi dengan cara setelah arus udara keluar dari *glotis* tidak mendapat hambatan dari alat ucap melainkan hanya diganggu oleh posisi lidah dan bentuk mulut.

Vokal kardinal adalh konsep untuk menentukan bunyi vokal berdasarkan posisi lidah yang berguna untuk membandingkan vokal-vokal suatu bahasa diantara bahasa lain.

Konsep vokal kardinal menjelaskan adanya posisi lidah tertinggi, terendah dan terdepan dalam memproduksi buyi vokal tersebut.

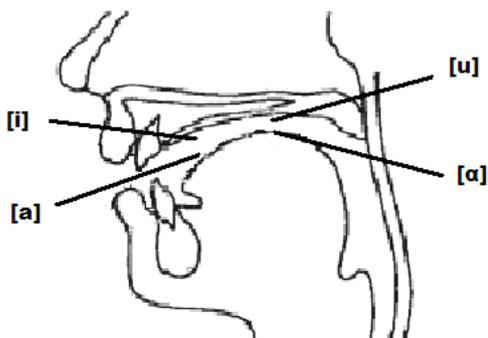
Bunyi vokal [i] diucapkan dnegan meninggikan lidah depan setinggi mungkin tanpa menyebabkan terjadinya konsonan gesekan.

Bunyi vokal [a] diucapkan dengan merendahkan lidah depan (ujung lidah) serendah mungkin.

Bunyi vokal [ɑ] diucapkan dengan merendahkan pangkal lidah dibawah mungkin.

Bunyi vokal [u] diucapkan dengan menaikkan pangkal lidah setinggi mungkin.

Posisi ucapan lidah untuk keempat vokal tersebut dapat digambarkan seperti pada gambar 8.



Gambar 7. Posisi Ucapan Lidah untuk Vokal [i],[a],[ɑ] dan [u] . [Chaer, 2009]

**6. PENUTUP**

Sebuah model artikulatori ideal memerlukan lebih banyak parameter agar dapat menghasilkkan ucapan buatan yang mirip dengan ucapan manusia. Model Marmelstein pada kajian paper ini memiliki 7 parameter artikulatori, yaitu badan lidah, ujung lidah, rahang, bibir, hyoid, velum dan parameter antara gigi atas dan langit-langit rongga mulut yang dapat digunakan untuk mensisntesa ucapan berbagai bahasa. Penelitian selanjutnya akan melakukan analisis perpindahan nilai parameter artikulator pada pensintesaan fonem-fonem vokal dan fonem ucapan kata Bahasa Indonesia.

**PUSTAKA**

Ahmad Arman, A. (2003). Perkembangan Teknologi TTS Dari Masa ke Masa. Diakses pada 9 Februari 2009 dari [http://indotts.melsa.net.id/perkembangan\\_TTS.pdf](http://indotts.melsa.net.id/perkembangan_TTS.pdf)

Ahmad Arman, A. (2003). Proses Pembentukan dan Karakteristik Sinyal Ucapan. Diakses pada 9 Februari 2009 dari [http://indotts.melsa.net.id/Karakteristik\\_Sinyal\\_Ucapan.pdf](http://indotts.melsa.net.id/Karakteristik_Sinyal_Ucapan.pdf)

Arie Nugraha, A. (2008). Penentuan Parameter Pembangkit Ucapan Model Artikulatori untuk Fonem-Fonem Bahasa Indonesia. *Skripsi ITB*, Bandung.

Berlianti, (2008). Penentuan Nilai-Nilai Parameter Articulatory Synthesizer Pada Beberapa Fonem Bahasa Indonesia. *Tesis ITB* Bandung.

. Childers, Donald G (2000). Speech Processing and Synthesis Toolboxes. *John Wiley & Sons, Inc.*, New York.

Chaer, Abdul. (2009). Fonologi Bahasa Indonesia, Reneka Cipta. Jakarta.

Christine H. Shadle; Robert I. Damper. (2001), Prospects for Articulatory synthesis: A Position paper, ISCA Workshop on Speech Synthesis 4-2001, p116. , Perthshire, Scotland.

Helmut Ploner-Bernard, Speech Synthesis by Articulatory Model, Diakses pada 23 Mei 2010 dari [http://www2.spsc.tugraz.at/www-archive/AdvancedSignalProcessing/WS03-WirelessCommunication/talks/ploner-bernhard\\_report.pdf](http://www2.spsc.tugraz.at/www-archive/AdvancedSignalProcessing/WS03-WirelessCommunication/talks/ploner-bernhard_report.pdf)

Jianwu Dang, Kiyoshi Honda, (2004), Construction and control of a physiological articulatory model, *Journal of Acoustical Society of America*, vol.115, no.2, pp.853-870.

Khalil Iskarous, Lous M. Goldstein, DH. Whalen, Mark K. Tiede and Philip E. Rubin (2003), Casey : The Haskins Configurable Articulatory Synthesizer, *Proceeding of the 15th International Congress of phonetic Science, Universitat autonima de Barcelona, Barcelona, Spain.*

Mermelstein, P. (1972), Artikulator Model For The Study of Speech Production, *The journal of the acoustical Society of America*, Volume 53, No.4: 1070-1082