Manusia sendiri banyak menggunakan ucapan untuk berinteraksi antara manusia. Cara interaksi menggunakan ucapan dianggap lebih mudah, cepat, dan lebih nyaman dilakukan. Untuk mewujudkan keinginan tersebut, ada tiga teknologi kunci yang diperlukan, yaitu sebagai berikut.

1. Teknologi *speech recognition* (pengenalan ucapan), diperlukan untuk mengenali setiap ucapan menjadi teks.
2. Teknologi *natural language processing* (NLP), diperlukan, untuk mengartikan ucapan-ucapan yang sudah dikenali oleh bagian pengenal ucapan dan mengkonversikannya ke dalam format yang dapat diolah oleh komputer. Hal ini diperlukan karena bahasa lisan yang diucapkan secara langsung sering tidak mengikuti kaidah-kaidah bahasa formal.
3. Teknologi *speech synthesizer* (pensintesa ucapan), diperlukan, untuk mengucapkan informasi-informasi teks yang dihasilkan oleh komputer. Sistem seperti ini dikenal pula dengan istilah sistem *Text To Speech* (TTS).

Penelitian yang dilakukan dalam rangka disertasi ini, adalah pengembangan sistem pensintesa ucapan Bahasa Indonesia, oleh karena itu pembahasan berikutnya dibatasi hanya pada bagian pensintesa ucapan saja. Penelitian di bidang pensintesa ucapan mengalami perjalanan, yang sangat panjang, dan telah dimulai sejak lama. Kegiatan-kegiatan penelitian tersebut dapat dibagi menjadi tiga era.

Berdasarkan hasil studi literatur dari berbagai sumber bacaan, perkembangan teknologi pensintesa ucapan dapat dibagi menjadi tiga kurun waktu. Kurun waktu pertama adalah sebelum 1930. Pada masa ini penelitian-

penelitian lebih banyak ditujukan untuk memahami karakteristik sinyal ucapan serta pengembangan pensintesa ucapan berbasis mekanik atau elektromekanik. Kurun waktu kedua dimulai sejak tahun 1930-an sampai dengan ditemukannya komputer digital. Masa ini ditandai dengan pengembangan berbagai alat pensintesa ucapan menggunakan teknologi elektronik analog. Kurun waktu ketiga dimulai sejak ditemukannya komputer digital hingga sekarang. Pada masa ini, sintesa ucapan dilakukan menggunakan pendekatan-pendekatan pemrosesan sinyal digital.

Teknologi komputer telah memperlihatkan kemajuan yang luar biasa pada pertengahan abad ini. Pada tahun 1945 tidak ada komputer dengan program tersimpan. Saat ini, hanya dengan uang beberapa ribu dolar saja dapat digunakan untuk membeli komputer pribadi yang mempunyai kelebihan unjuk kerja, dengan memori utama yang lebih besar serta penyimpan yang lebih besar dibandingkan dengan komputer yang dibeli pada tahun 1965 dengan harga satu juta dolar. Kemajuan yang sangat pesat ini diperoleh dari kemajuan teknologi yang digunakan untuk membuat komputer serta dari inovasi dalam perancangan komputer.

Sejak tahun 1985, industri komputer memperlihatkan pola arsitektur baru yang memberikan beberapa kelebihan dan memulai periode dimana unjuk kerjanya semakin meningkat dengan kecepatan yang lebih tinggi. Melalui kemajuan dalam teknologi rangkaian terpadu, perbaikan dalam teknologi *compiler*, dan ide arsitektur baru, perancang mampu menciptakan serangkaian komputer dengan perbaikan unjuk kerja dengan faktor hampir dua kali lipat setiap

tahun. Dalam tahun-tahun terakhir bahkan unjuk kerjanya meningkat dua setengah kali lipat dalam setahun.

Perkembangan teknologi komputer yang sangat pesat ini, memicu perkembangan di berbagai bidang lain yang salah satu diantaranya adalah Teknologi Bahasa Manusia (TBM).

2.2 Teknologi Bahasa Manusia (TBM)

2.2.1 Sistem Bahasa Ucapan (SBU).

Dalam kelompok pertama yaitu SBU, kegiatan yang dilakukan adalah pengembangan sistem pengenalan ucapan *(speech recognitio*n) dengan model hidden Markov dan jaringan syaraf terpadu *(neural networ*k), pengembangan sistem pembangkit ucapan dari suatu teks *(text-to-speech synthesis syste*m), sistem pemahaman bahasa alami dan pemodelan dialog.

Banyak riset yang telah dilakukan untuk mengembangkan TBM, baik dalam bentuk ucapan maupun dalam bentuk tulisan. Dalam hal SBU terutama dalam bidang pengenalan ucapan, kemajuannya sangat berkaitan dengan proyek yang disponsori oleh DARPA *(Defense Department's Advanced Research Projects Agenc*y) di Amerika Serikat dan usaha beberapa industri besar seperti IBM dan AT&T. Antara tahun 1971 dan 1976, program ucapan DARPA yang pertama dihasilkan untuk pengenalan ucapan pada pengucap yang tertentu untuk 1000 kata yang tingkat perbedaannya rendah. Sejak tahun 1984, DARPA telah mendanai pengembangan kosakata yang besar untuk sistem pengenalan ucapan kontinyu. Sistem ini berusaha untuk menuliskan kata-kata dari berita yang

disiarkan dan percakapan telepon, dengan tanpa kendala ukuran kosakata. Kemajuan yang cukup pesat telah dibuat dalam pengenalan kata dari ucapan kontinyu yang alami, walaupun unjuk kerjanya masih kurang bagus bila dibandingkan dengan unjuk kerja manusia.

2.2.2 Sistem Bahasa Tertulis (SBT).

Kegiatan yang dilakukan dalam SBT diantaranya adalah pengembangan OCR *(Optical Character Reade*r), Pemeriksa Ejaan *(Spell Checke*r), Pemeriksa Tatabahasa *(Grammar Checke*r), Pemenggalan Kata *(hyphenatio*n), Sistem Temubalik Informasi *(Information Retrieva*l) dan juga sistem Mesin Penerjemah.

Kemajuan pengembangan SBU jauh tertinggal bila dibandingkan dengan SBT. Beberapa kendala yang membuat SBU tertinggal jauh diantaranya adalah pendefinisian sifat-sifat SBU yang merupakan interaksi antara manusia dengan mesin. Hambatan kedua adalah kurangnya alat bantu yang mudah diakses serta murah untuk mendukung riset dan transfer teknologi. Seperti kita ketahui bahwa SBU merupakan aktivitas yang kompleks, memerlukan sumber-sumber komputer tertentu, pengintegrasian pemrosesan sinyal yang canggih, pelatihan dan algoritma pengenalan, serta sumber-sumber bahasa seperti korpus ucapan dan kamus pengucapan.

Sistem pembangkitan ucapan dari suatu teks pada dasarnya lebih sederhana dibandingkan dengan sistem pengenalan ucapan. Masalah pembangkitan ucapan dari suatu teks untuk suatu bahasa tertentu dapat dipecah ke dalam dua sub masalah:

1. Hal pertama berupa konversi spesifikasi parameter bahasa (contoh: urutan fonem, parameter aksen dll.) ke dalam paramater-parameter yang dapat menggerakkan pembangkit ucapan yang nyata.
2. Masalah kedua melibatkan komputasi spesifikasi parameter bahasa dari teks masukan. Walaupun sistem pembangkitan ucapan tidak serumit sistem pengenalan ucapan namun hasil yang diperoleh dari sistem ini sejauh ini masih mempunyai aksen yang datar dan kurang alamiah bila dibandingkan dengan ucapan manusia.

Pengembangan SBT sangat berhubungan erat dengan siklus dokumen yaitu dibuat, didistribusikan, dicetak, dibaca, di-*sca*n, dan digunakan kembali. Dalam pembuatan dokumen sampai pencetakan biasanya dilakukan dengan sistem pengolah kata yang saat ini biasanya sudah terintegrasi dengan pemeriksa ejaan, pemeriksa tatabahasa, pemenggalan kata dll. Untuk penggunaan kembali suatu dokumen yang telah tercetak dapat dilakukan dengan men-*scan* dokumen yang telah tercetak dan mengubahnya ke dalam suatu teks dokumen melalui sistem OCR.

Sistem SBT yang lain berkaitan dengan kebutuhan untuk menemukan informasi tertentu dari suatu dokumen yang sangat besar. Hal ini memicu pengembangan sistem temu balik informasi sehingga pencarian suatu data dapat dilakukan dengan cepat. Sedangkan dalam lingkungan multibahasa, sistem mesin penerjemah memegang peranan yang cukup penting untuk mengubah suatu dokumen dari suatu bahasa tertentu ke dalam dokumen bahasa yang lain.

2.3 Teknologi Bahasa Manusia (TBM) – Dalam Bahasa Indonesia

Bahasa Indonesia yang memiliki populasi pengguna lebih dari 200 juta manusia, mempunyai perkembangan yang sangat cepat dan dinamis, terutama dalam era globalisasi ini sehingga membutuhkan penerapan TBM untuk dapat membantu mengembangkan beberapa aspek kebahasaan yang akhirnya dapat membantu pengembangan bahasa Indonesia itu sendiri. Walaupun penelitian di bidang TBM dengan komputer untuk bahasa Indonesia belum begitu lama dimulai yaitu sekitar tahun 1980-an, namun. saat ini Indonesia tidak begitu tertinggal dalam penguasaan teknologi terutama dalam sistem SBT dibandingkan dengan negara-negara maju seperti Amerika, Jepang, dan beberapa negara Eropa seperti Perancis, Jerman, Spanyol, Inggris dll. Hal ini terlihat dalam beberapa proyek seperti proyek mesin penerjemah multibahasa antara Jepang, Indonesia, Malaysia, Thailand, dan Cina yang disponsori oleh pihak Jepang dimana Indonesia dapat turut aktif dalam mengembangkan modul-modul dari keseluruhan sistem yang dikembangkan.

Selain itu dalam proyek UNL *(Universal Networking Languag*e) yang disponsori oleh UNU/IAS (**United Nations University/ Institute of Advanced Stud**y) dan saat ini sedang berjalan, dimana melibatkan 13 bahasa yang mempunyai pengucap lebih dari 100 juta, Indonesia dapat berperan cukup aktif didalamnya.

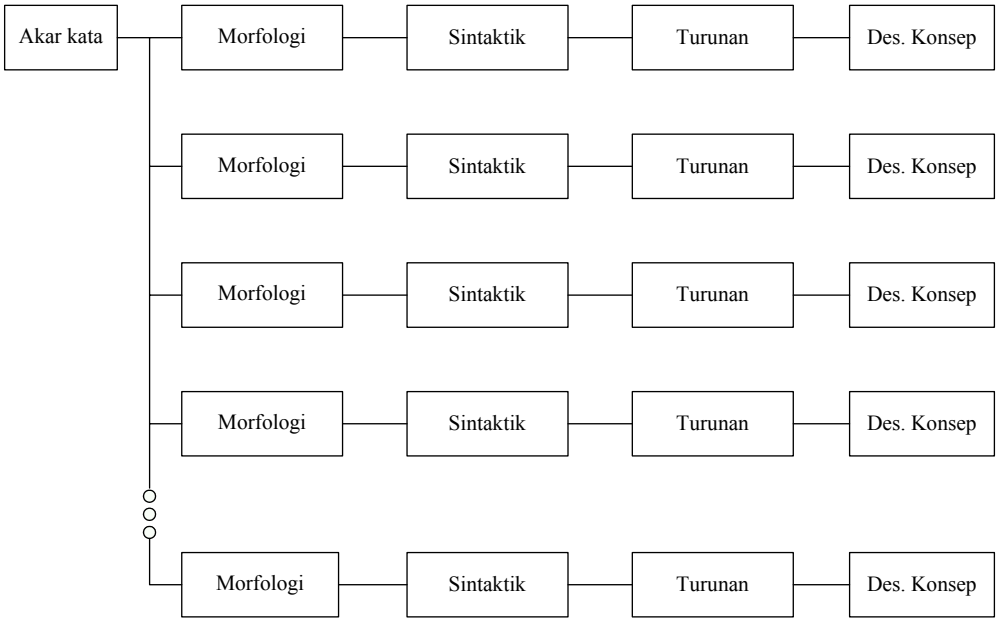
Kemajuan penelitian TBM bahasa Indonesia juga dapat dilihat dari ketersediaan sumber-sumber kebahasaan secara elektronik yang ada saat ini seperti kamus elektronik bahasa Indonesia, korpus bahasa Indonesia dll., serta

beberapa perangkat bantu untuk melakukan penelitian bahasa seperti sistem KWIC, Statistik kebahasaan, penampil grafis dari suatu representasi internal kalimat seperti UNL dll.

Dalam sub bagian berikut akan dibahas mengenai kamus elektronik bahasa Indonesia serta beberapa perangkat bantu yang dikembangkan di BPPT dan juga beberapa sistem TBM yang telah dikembangkan untuk bahasa Indonesia. Informasi semantik memuat ciri-ciri semantik suatu kata berdasarkan klasifikasi semantik bahasa Indonesia. Informasi ini sangat berguna dalam sistem pengolahan bahasa yang memerlukan analisis kalimat secara mendalam. Secara struktur, semua informasi di dalam KEBI yang mengacu pada suatu kata ditata dalam struktur hirarki untuk memperlihatkan hubungannya. Akar kata yang merupakan dasar dalam pembentukan suatu kata turunan menjadi dasar entri kamus.

Di dalam KEBI ini akar kata dapat berupa morfem , kata, atau idiom yang menjadi dasar pembentukan suatu kata. Akar kata bahasa Indonesia ini dapat membentuk beberapa kata turunan melalui penambahan suatu imbuhan tertentu yang terkadang mempunyai beberapa jenis kata dan juga mempunyai beberapa arti.

Dalam hal ini, ciri-ciri gramatik dapat digunakan untuk mengaitkan antara kata turunan dengan artinya. Struktur dari KEBI sendiri dapat dilihat dalam *gambar 2.1*.



Gambar 2.1 Struktur KEBI

Berdasarkan struktur di atas, maka informasi yang termuat di dalam KEBI dapat disusun dalam tabel seperti terlihat pada *tabel 2.1*. Saat ini KEBI memuat 22.500 akar kata yang membentuk sekitar 44.000 kata turunan dan pengguna Internet dapat mengakses beberapa informasi di dalamnya secara langsung pada alamat http://www.aia.bppt.go.id/kebi .

KEBI telah digunakan dalam pengembangan sistem mesin penerjemah multibahasa yang meliputi bahasa Jepang, Indonesia, Malaysia, Thailand, dan Cina. Hasil prototipe dari proyek sistem mesin penerjemahan multibahasa ini dapat dilihat secara langsung melalui Internet pada alamat http://www.aia.bppt.go.id/mmts.

Tabel 2.1 Informasi KEBI

Informasi Akar Kata	Informasi Morfologi	Informasi Sintaktik	Informasi Semantik	Informasi Kata Turunan
Akar Kata Nomor Akar Kata	Kode Konjugasi Aktif Kode Kunjugasi Pasif Kode Konjugasi Lain Kode Reduplikasi	Jenis Kata	Ciri Semantik	Deskripsi Konsep

2.4 Sistem KWIC (Key Word In Context)

KWIC merupakan suatu alat bantu dalam penelitian pola kalimat suatu bahasa berdasarkan suatu kata didalam suatu korpus. Sistem ini pada dasarnya melakukan pengindeksan terhadap suatu korpus, sehingga pengguna dapat menampilkan semua kalimat yang terdapat dalam korpus tersebut yang mengandung unsur suatu kata tertentu, sesuai dengan permintaan pengguna.

KWIC terdiri atas dua bagian utama yaitu:

1. Sistem Pengindeks

   Sistem pengindeks berfungsi untuk membuat indeks kata dari suatu korpus, sehingga letak kata tertentu dalam korpus dapat dicari dengan cepat.

2. Sistem Q*uer*y.

   Sistem query akan menerima masukan kata yang ingin dilihat pola kalimatnya oleh pengguna dan berdasarkan indeks yang telah dibuat pada sistem pengindeks, sistem ini akan mengakses korpus dan mengambil semua kalimat yang mempunyai unsur kata tersebut yang merupakan keluaran dari sistem ini.

Peneliti bahasa dapat menggunakan KWIC sebagai alat bantu penelitian pola kalimat. Sebagai contoh, untuk mendapatkan semua kalimat yang

mengandung unsur **makan** maka kita tinggal memasukkan kata **makan** dalam sistem query sehingga diperoleh semua kalimat yang mengandung kata makan di dalam korpus yang diolah. Selanjutnya peneliti tersebut dapat menganalisa pola kalimat yang menggunakan kata **makan**. Tentu saja hasil yang diperoleh sangat tergantung pada kelengkapan korpus yang tersedia.

2.5 Perangkat Statistik Bahasa

Penggunaan statistik sangat penting dalam berbagai penelitian yang juga dirasakan dalam penelitian bahasa. Penggunaan statistik dalam bidang Bahasa diantaranya adalah pembuatan data frekuensi suatu kata dalam suatu korpus tertentu. Salah satu pemanfaatan data frekuensi kata ini adalah dalam sistem penerjemah yaitu untuk pemilihan kata yang dibangkitkan jika terdapat berbagai pilihan kata yang akan dibangkitkan.

Perangkat ststistik ini belum ada yang dikembangkan secara khusus untuk Bahasa Indonesia, namun saat ini telah tersedia beberapa perangkat statistik yang dapat kita dapatkan secara gratis dan salah satu diantaranya adalah **CMU Statistical Language Modeling Toolkit Version.2** yang dikembangkan oleh Roni Rosenfeld di universitas Carnegie Mellon.

Pemanfaatan program statistik telah dirasakan sangat membantu dalam mengembangkan berbagai aplikasi dan saat ini telah banyak dihasilkan produk pengolahan bahasa yang berdasarkan pada metoda statistik.

2.6 Pemeriksa Ejaan, Pemeriksa Tatabahasa, dan Pemenggalan Kata

Saat ini banyak terdapat program pengolah kata yang beredar di Indonesia. Sistem-sistem ini biasanya disertai dengan program untuk pemenggalan kata secara otomatis *(Hyphenatio*n) , pemeriksa ejaan *(Spell Checke*r) dan Pemeriksa Tatabahasa (Grammar Checker) dan terpadukan di dalam sistem tersebut.

Fasilitas ini sangat bergantung pada bahasa dimana sistem itu dirancang. Sebagai contoh, untuk pemenggalan kata, Bahasa Indonesia mempunyai aturan tersendiri yang berbeda dengan bahasa lain. Karena sebagian besar sistem yang beredar dibangun untuk Bahasa Inggris, maka jika kita menggunakan fasilitas pemenggalan kata pada waktu mengolah Bahasa Indonesia sering terjadi kesalahan.

Fasilitas pemeriksa ejaan dan pemeriksa tatabahasa saat ini menjadi bagian yang sangat penting dalam sistem pengolah kata. Walaupun demikian belum terdapat satupun sistem pengolah kata yang memberikan fasilitas lokal untuk Bahasa Indonesia terutama untuk pemeriksa ejaan dan pemeriksa tatabahasa. Hal ini mengakibatkan pengolahan kata Bahasa Indonesia menjadi kurang optimal dan pengguna tidak dapat mencegah adanya kesalahan eja di dalam dokumen.

Keterkaitan sistem pemeriksa ejaan dan pemeriksa tatabahasa dengan sistem pengolah kata menimbulkan ketergantungan dengan pengembang pengolah kata tersebut jika kita ingin menambahkan fasilitas lokal untuk bahasa Indonesia.

Saat ini *Lotus Corporation* telah mengadakan kerjasama pengembangan pemeriksa ejaan Bahasa Indonesia dengan Fakultas Ilmu Komputer Universitas Indonesia untuk digunakan pada sistem pengolah kata yang dikembangkan oleh Lotus. Kerjasama ini baru dimulai dan saat ini masih dalam proses pengembangan.

Pada dasarnya pengembangan program pemeriksa ejaan tidaklah rumit. Sistem ini terdiri atas kamus dan suatu program untuk menangani proses morfologi bahasa, memeriksa apakah suatu kata ada di kamus atau merupakan kata turunan dari kata yang ada di kamus, dan jika tidak ada maka akan memberikan kandidat kata yang benar kepada pengguna.

Untuk fasilitas pengolahan kata yang lain yang berhubungan dengan Bahasa Indonesia, team peneliti dari Universitas Gadjah Mada telah mengembangkan program *macro* pada *Microsoft Word* untuk sistem pemenggalan kata Bahasa Indonesia yang dinamakan ***Indotool***. Selain itu mereka juga ingin meningkatkan Indotool ini untuk sistem pemeriksa tatabahasa Bahasa Indonesia.

2.7 Sistem Temu Balik Informasi Bahasa Indonesia

Komputer saat ini telah menjadi alat utama dalam siklus dokumen yaitu dibuat, didistribusikan, dicetak, dibaca, di-*sca*n, dan digunakan kembali. Penggunaan komputer dalam pembuatan dokumen ini membuat ketersediaan dokumen secara elektronik dari berbagai bidang. Hal ini berlangsung cukup lama sehingga dokumen yang tersedia menjadi sangat banyak. Suatu hal yang tidak dapat dihindari dalam hal ini adalah keinginan untuk mendapatkan suatu

informasi dari dokumen yang tersedia tersebut. Karena dokumen yang tersedia sangat banyak, maka diperlukan suatu sistem tersendiri untuk dapat mencari informasi tertentu dari dokumen yang tersedia tersebut. Sistem pencarian informasi dari dokumen-dokumen yang tersedia ini disebut sistem temu-balik Informasi.

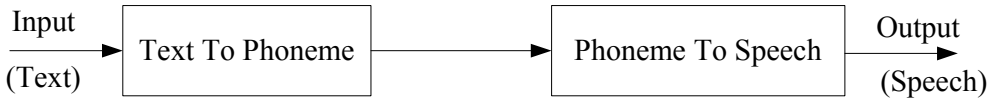
Sistem temu balik informasi tentu saja tergantung pada bahasa yang digunakan dalam dokumen yang akan dicari. Sistem temu balik informasi yang dikembangkan untuk Bahasa Inggris tentu saja kurang optimal apabila langsung diterapkan untuk penanganan dokumen Bahasa Indonesia. Bahasa Indonesia mempunyai **"kata buang"** tersendiri serta sistem pembentukan kata yang sangat lain dengan bahasa Inggris, sehingga diperlukan sistem temu balik informasi yang khusus untuk Bahasa Indonesia.

Dalam sistem temu balik informasi hal yang paling utama adalah pengindeksan dokumen dan sistem pencarian dokumen berdasarkan indeks yang telah dibuat tersebut. Dalam sistem temu balik Bahasa Indonesia hal pertama yang perlu diperhatikan adalah pengumpulan kata buang bahasa Indonesia, dan menghilangkan kata tersebut dalam pengindeksan. Dengan menghilangkan kata buang dari indeks yang dibuat maka kata yang diindeks menjadi lebih sedikit sehingga ruang yang diperlukan untuk penyimpanan indeks lebih sedikit dan pencarian informasi dapat dilakukan dengan lebih cepat.

2.8 Konversi *Text To Speech*

Sistem *Text to Speech* pada prinsipnya terdiri dari dua sub sistem, yaitu :

1. Bagian Konverter Teks ke Fonem *(Text to Phonem*e), serta
2. Bagian Konverter Fonem to Ucapan *(Phoneme to Speec*h).



Gambar 2.2 Sistem Text To Speech

Bagian *Konverter Teks ke Fonem* berfungsi untuk mengubah kalimat masukan dalam suatu bahasa tertentu yang berbentuk teks menjadi rangkaian kode-kode bunyi yang biasanya direpresentasikan dengan kode fonem, durasi serta *pitc*h-nya. Bagian ini bersifat sangat *language dependan*t. Untuk suatu bahasa baru, bagian ini harus dikembangkan secara lengkap khusus untuk bahasa tersebut.

Bagian *Konverter Fonem ke Ucapan* akan menerima masukan berupa kode-kode fonem serta pitch dan durasi yang dihasilkan oleh bagian sebelumnya. Berdasarkan kode-kode tersebut, bagian Konverter Fonem ke Ucapan akan menghasilkan bunyi atau sinyal ucapan yang sesuai dengan kalimat yang ingin diucapkan. Ada beberapa alternatif teknik yang dapat digunakan untuk implementasi bagian ini. Dua teknik yang banyak digunakan adalah:

1. *Formant Synthesize*r

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

2. *Diphone Concatenatio*n.

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

Struktur sistem seperti di atas pada prinsipnya merupakan konfigurasi tipikal yang digunakan pada berbagai sistem *Text to Speech* berbagai bahasa. Namun demikian, pada setiap sub-sistem terdapat sifat-sifat serta proses-proses yang sangat spesifik dan sangat tergantung dari bahasanya.

Konversi dari teks ke fonem sangat dipengaruhi oleh aturan-aturan yang berlaku dalam suatu bahasa. Pada prinsipnya proses ini melakukan konversi dari simbol-simbol tekstual menjadi simbol-simbol fonetik yang merepresentasikan unit bunyi terkecil dalam suatu bahasa. Setiap bahasa memiliki aturan cara pembacaan dan cara pengucapan teks yang sangat spesifik. Hal ini menyebabkan

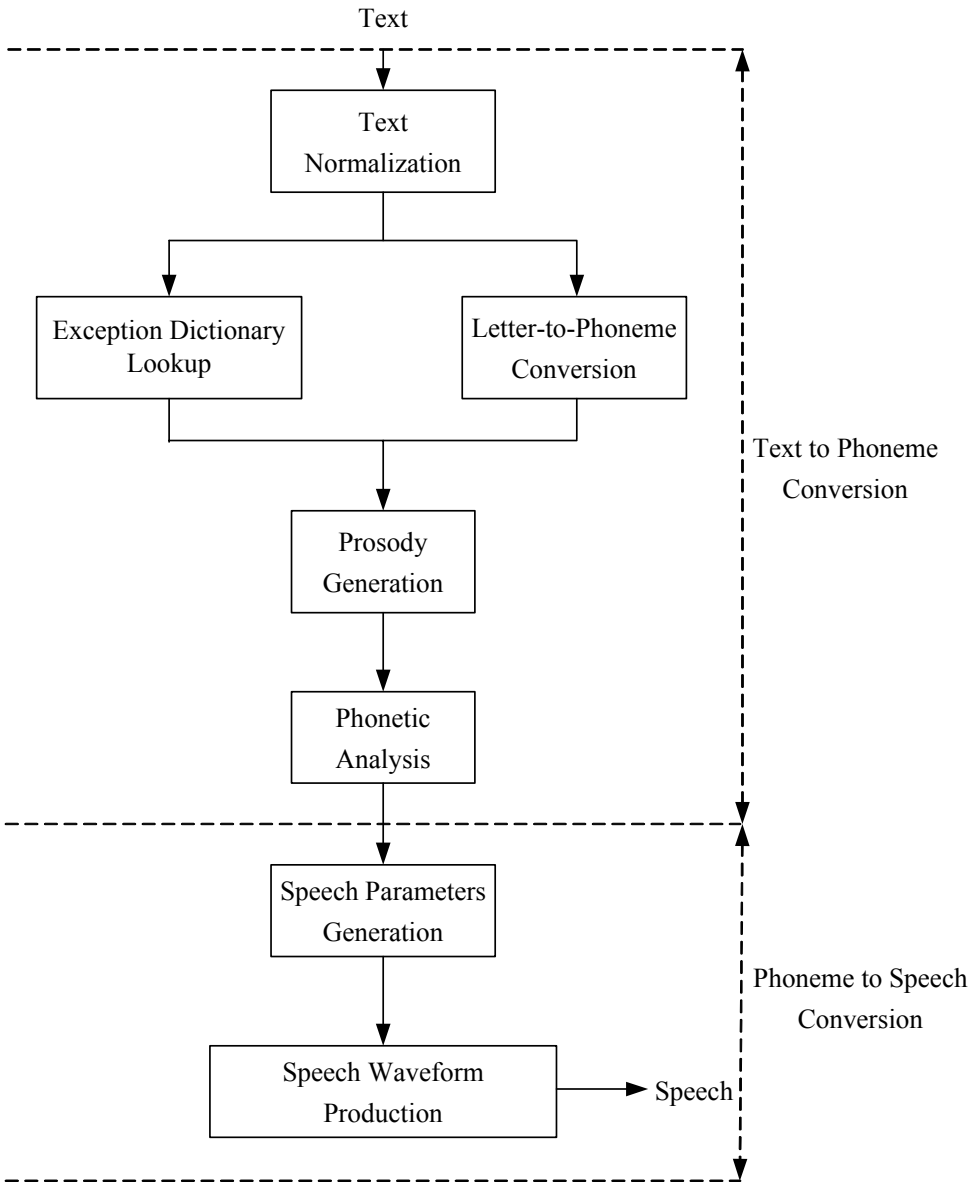
implementasi unit konverter teks ke fonem menjadi sangat spesifik terhadap suatu bahasa.

Untuk mendapatkan ucapan yang lebih alami, ucapan yang dihasilkan harus memiliki intonasi *(prosod*y). Secara kuantisasi, prosodi adalah perubahan nilai *pitch* (frekuensi dasar) selama pengucapan kalimat dilakukan atau pitch sebagai fungsi waktu. Pada prakteknya, informasi pembentuk prosodi berupa data-data pitch serta durasi pengucapannya untuk setiap fonem yang dibangkitkan. Nilai-nilai yang dihasilkan diperoleh dari suatu model prosodi. Prosodi bersifat sangat spesifik untuk setiap bahasa, sehingga model yang diperlukan untuk membangkitkan data prosodi menjadi sangat spesifik juga untuk suatu bahasa. Beberapa model umum prosodi pernah dikembangkan, tetapi untuk digunakan pada suatu bahasa masih perlu banyak penyesuaian yang harus dilakukan.

Konverter fonem ke ucapan berfungsi untuk membangkitkan sinyal ucapan berdasarkan kode-kode fonem yang dihasilkan dari proses sebelumnya. Sub sistem ini harus memiliki pustaka setiap unit ucapan dari suatu bahasa. Pada sistem yang menggunakan teknik *diphone concatenatio*n, sistem harus didukung oleh suatu *diphone database* yang berisi rekaman segmen-segmen ucapan yang berupa diphone. Ucapan dalam suatu bahasa dibentuk dari satu set bunyi yang mungkin berbeda untuk setiap bahasa, oleh karena itu setiap bahasa harus dilengkapi dengan diphone database yang berbeda.

Tahapan-tahapan utama konversi dari teks menjadi ucapan dapat dinyatakan dengan diagram seperti terlihat pada *Gambar 2.3*. Tahap normalisasi

teks berfungsi untuk mengubah semua teks kalimat yang ingin diucapkan menjadi teks yang secara lengkap memperlihatkan cara pengucapannya.



Gambar 2.3. Urutan Proses Konversi dari Teks ke Ucapan (dimodifikasi dari Pelton, 1992)

Tahap berikutnya adalah melakukan konversi dari teks yang sudah secara lengkap merepresentasikan kalimat yang ingin diucapkan menjadi kode-kode fonem. Konversiteks menjadi fonem biasanya dilakukan dengan dua cara. Sebagian proses konversi dapat dilakukan dengan aturan konversi yang sederhana dan berlaku umum untuk berbagai kondisi. Sebagian proses lainnya bersifat kondisional, tergantung dari huruf-huruf atau fonem-fonem tetangganya, bahkan terdapat bentuk-bentuk translasi yang tidak dapat ditemukan keteraturannya.

Konversi yang teratur dapat diimplementasikan dengan tabel konversi yang berisi pasangan antara urutan huruf dan urutan fonem, bahkan mungkin hanya berisi satu huruf dan satu fonem. Aturan yang lebih sulit biasanya diimplementasikan dengan tabel konversi yang akan diterapkan jika kondisi rangkaian huruf tetangga kiri dan kanannya terpenuhi. Contoh bentuk aturan konversi huruf ke fonem yang memenuhi teknik tersebut adalah sebagai berikut.

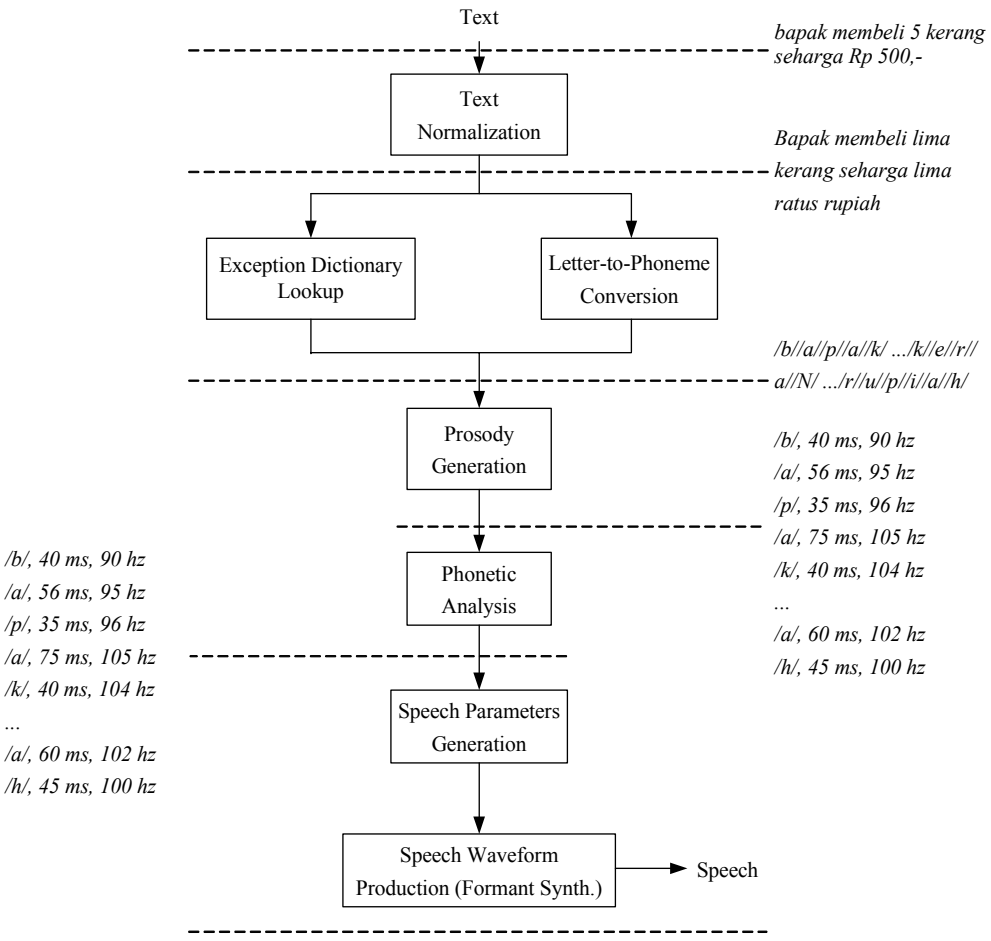
Left-context [letter-set] right-context = phoneme string

Huruf tertentu yang ditunjuk dalam posisi ***[letter-set]*** akan dikonversikan menjadi suatu fonem dalam ***"phoneme string***" jika ***left-context*** dan ***right context*** terpenuhi. Bahasa Inggris termasuk bahasa yang mempunyai keteraturan yang rendah untuk proses konversi teks ke fonem. Suatu TTS bahasa Inggris biasanya dilengkapi dengan suatu basis data yang berisi ribuan kata serta konversi padanan urutan fonemnya. Bahasa Indonesia termasuk bahasa yang jelas aturan konversinya. Sebagian besar kata dalam Bahasa Indonesia dapat dikonversikan menjadi fonem dengan aturan yang jelas dan sederhana, walaupun tetap ada kondisi-kondisi yang tidak dapat ditemukan keteraturannya. Sebagai contoh,

simbol huruf e dapat diucapkan sebagai *e pepet* atau *e talin*g, artinya harus dikonversikan menjadi fonem yang berbeda untuk kondisi yang berbeda. Dalam blok diagram di atas, kondisi yang masih dapat ditangani oleh aturan diimplementasikan dengan blok *Letter to Phoneme Conversion.* Konversi yang tidak teratur ditangani oleh bagian *Exception Dictionary Lookup*.

Hasil dari tahap tersebut adalah rangkaian fonem yang merepresentasikan bunyi kalimat yang ingin diucapkan. Bagian *prosody generator* akan melengkapi setiap unit fonem yang dihasilkan dengan data durasi pengucapannya serta pitchnya. Data durasi serta pitch diperoleh berdasarkan kombinasi antara tabel atau database serta model prosodi. Secara simbolik, hasil dari bagian ini sudah menghasilkan informasi yang cukup untuk menghasilkan ucapan yang diinginkan.

Satu tahap berikutnya yang masih sering dilakukan adalah *Phonetic Analysi*s. Tahap ini dapat dikatakan sebagai tahap penyempurnaan, yaitu melakukan perbaikan di tingkat bunyi. Sebagai contoh, dalam bahasa Indonesia, fonem /k/ dalam kata *bapak* tidak pernah diucapkan secara tegas, atau adanya sisipan fonem /y/ dalam pengucapan kata *alamiah* antara fonem /i/ dan /a/.



Gambar 2.4 Besaran-besaran Dalam Setiap Tahap Proses Konversi dari Teks ke Ucapan (dimodifikasi dari Pelton, 1992)