

✓✓

## **Konstruksi Bunyi Bahasa Indonesia: Strategi Pemenggalan Kata<sup>1</sup>**

*Multamia RMT Lauder*

### **1. Pembuka**

Konstruksi Bunyi Bahasa Indonesia yang akan saya paparkan pada kesempatan ini sesungguhnya merupakan laporan utak-atik saya mengenai upaya mengenali konstruksi bunyi untuk kepentingan pengomputerisasian pemenggalan kata. Hal-hal yang akan saya sampaikan di sini sesungguhnya lebih merupakan lanjutan dari apa yang pernah saya sajikan pada Sewindu Pertemuan Linguistik Lembaga Bahasa Atmajaya 1994. Pada makalah tersebut saya telah menelusuri kendala apa saja yang menyebabkan proses komputerisasi pemenggalan kata tersendat-sendat atau bahkan terhenti. Penelusuran itu dilaksanakan berdasarkan pedoman pemenggalan kata yang mutakhir dari Pusat Pembinaan dan Pengembangan Bahasa. Pedoman itu sebetulnya merupakan hasil rapat kerja ke-31 Panitia Kerja Sama Kebahasaan antara Brunei Darussalam, Indonesia, dan Malaysia pada tanggal 16--20 Desember 1991.

Dalam makalah itu, saya hanya mencoba menerapkan Pedoman Pemenggalan Kata 1991 pada semua lema yang berjumlah 32.489 dari A hingga Z pada Kamus Besar Bahasa Indonesia (1993). Untuk sementara, berdasarkan ujicoba yang telah dilaksanakan, saya menyimpulkan bahwa komputerisasi pemenggalan kata dalam bahasa Indonesia sulit untuk direalisasikan mengingat sifat dari kaidahnya itu sendiri (lihat ringkasan makalah terdahulu pada lampiran 1). Dugaan saya, kaidah-kaidah pemenggalan kata bahasa Indonesia yang dibuat oleh para ahli bahasa kita, kemungkinan besar tidak dapat atau sulit sekali dipindahkan dalam bentuk perintah-perintah terstruktur untuk mengenali jajaran huruf yang harus dipenggal oleh komputer. Dalam hal ini harus diakui bahwa penyusunan kaidah pemenggalan kata bahasa Indonesia bukanlah hal

---

<sup>1</sup> Makalah ini disajikan untuk Seminar Penerapan Teknologi Komputer pada Pengembangan dan Penelitian Bahasa, Depok, 24--25 Januari 1995.

yang mudah, mengingat sifat bahasa Indonesia yang terus berkembang dan terbuka pada unsur-unsur daerah maupun asing. Akan tetapi usaha penyusunan itu bukanlah hal yang tak mungkin.

## 2. Konstruksi Bunyi

Konstruksi bunyi bahasa dapat dilihat pada berbagai tataran. Tataran yang terkecil biasanya terlihat pada satuan-satuan bunyi yang bergabung dalam satu suku kata. Suku kata dalam bahasa Indonesia sekurang-kurangnya harus terdiri atas satu bunyi vokal. Dapat pula, sebuah suku kata itu misalnya terdiri dari V (vokal) dan K (konsonan) atau KV, KVK, KKVK, KKKV, KKKVK. Keberadaan vokal dalam konstruksi itu sebagai puncak kenyaringan bunyi yang berfungsi menandai satu kesatuan suku kata.

Para ahli bahasa terdahulu telah melakukan observasi secara cermat mengenai profil bahasa Indonesia. Berdasarkan pengamatan mereka, diperkirakan konstruksi kata dasar bahasa Indonesia yang dianggap asli adalah kata yang terdiri dari dua suku kata. Kemudian konstruksi suku kata bahasa Indonesia yang frekuensi kemunculannya tinggi adalah V, VK, KV, dan KVK.

Halim (1974) dalam penelitiannya mengenai intonasi juga menjabarkan profil konstruksi bunyi bahasa Indonesia. Dalam hal ini profil bahasa Indonesia di tahun 1970-an. Pada dasarnya, hasil penelitian Halim itu memperkuat observasi para ahli terdahulu lalu ditambah dengan beberapa rincian lainnya.

Konstruksi bunyi kata dasar bahasa Indonesia yang terdiri dari dua suku kata dapat diramalkan kemunculannya berdasarkan matriks ini:

	KV	KVK	VK	V
KV	+	+	+	+
KVK	+	+	-	-
VK	+	+	-	-
V	+	+	+	+

(Halim 1974:182)

Berdasarkan matriks tersebut di atas, Halim menjabarkan kemungkinan beberapa konstruksi kata dasar bahasa Indonesia yaitu,

Nomor	Konstruksi Kata	Contoh
1.	KV-KV	lu-pa
2.	KV-KVK	ma-kan
3.	KV-VK	ka-in
4.	KV-V	ma-u
5.	KVK-KV	tan-da
6.	KVK-KVK	lom-pat
7.	VK-KV	ang-ka
8.	VK-KVK	ar-wah
9.	V-KV	i-ni
10.	V-KVK	a-nak
11.	V-VK	a-ir
12.	V-V	i-a

(Halim 1974:182)

Secara umum kata dasar yang terdiri dari dua suku kata memperlihatkan 12 macam konstruksi. Sebagaimana kita ketahui bersama kosa kata dasar bahasa Indonesia yang dianggap asli tidak mengenal gugus konsonan dalam konstruksi suku katanya. Berdasarkan permutasi kemungkinan munculnya konstruksi kata dasar dapat diamati bahwa konstruksi nomor 5 sampai dengan 8 mengandung dua konsonan yang berdampingan. Suku kata pertama diakhiri oleh sebuah konsonan dan suku kata kedua diawali oleh sebuah konsonan.

Berdasarkan penelitian itu, dari semua data yang kata dasarnya bersuku kata dua (500 kata), terpilah dalam dua kelompok yang memperlihatkan jenis konsonan pendamping yang berbeda. Kelompok pertama berjumlah 70% menunjukkan bahwa suku kata pertama selalu mengandung konsonan nasal dan suku keduanya mengandung konsonan non-nasal yang homorganik. Sedangkan kelompok kedua berjumlah 30% menunjukkan bahwa suku kata pertama cenderung mengandung konsonan getar.

Dalam hal ini yang dimaksud dengan bunyi yang homorganik adalah bunyi yang dihasilkan proses yang mirip yaitu dengan menggunakan titik artikulasi serta artikulator aktif yang sama. Bunyi konsonan nasal dan

konsonan non-nasal yang homorganik dapat didaftarkan sebagai berikut:  
bunyi  $m \Rightarrow p$  dan  $b$ ;  $n \Rightarrow t, d, c, j$ , dan  $s$ ;  $ng \Rightarrow k$  dan  $g$

### I. Nasal dan Non-Nasal (70%)

Nomor	Konstruksi	Persentase	Contoh
1.	ng-g	24%	eng-gak
2.	n-t	18%	nan-ti
3.	n-d	8%	un-dang
4.	m-p	8%	lom-pat
5.	ng-k	4%	ang-kat
6.	m-b	3%	am-bil
7.	n-c	2%	in-ci
8.	n-j	2%	pin-jam
9.	n-s	1%	pm-sil

(Halim 1974:183)

### II. Cenderung Getar (30%)

Nomor	Konstruksi	Persentase	Contoh
1.	r-l	5%	ker-ja
2.	s-t	4%	mus-ti
3.	r-g	2%	har-ga
4.	r-l	2%	per-lu
5.	r-t	2%	ker-tas
6.	k-s	2%	sak-si
7.	k-t	2%	wak-tu
8.	r-k	1%	ber-kas
9.	r-n	1%	per-nah
10.	r-d	1%	ker-dil
11.	l-m	1%	il-mu
12.	l-p	1%	tel-pun
13.	m-r	1%	lum-rah
14.	t-w	1%	jad-wal
15.	p-t	1%	sap-ta
16.	r-w	0,5%	ar-wah
17.	h-w	0,5%	bah-wa
18.	r-s	0,5%	per-sen
19.	r-b	0,5%	per-ban
20.	r-m	0,5%	ter-mos
21.	k-l	0,5%	ik-lan

(Halim 1974:184)

Jika ditilik secara seksama dapat terlihat adanya 21 pola. Terdapat 11 pola suku pertamanya mengandung konsonan getar [r]. Ada 7 pola yang suku pertamanya cenderung mengandung konsonan tak bersuara [p, t, k, s, dan h]. Sisanya 2 pola yang suku pertamanya mengandung konsonan sampingan [l] serta 1 pola yang suku pertamanya mengandung konsonan nasal [m].

Tabel-tabel tersebut di atas merupakan profil konstruksi bunyi bahasa Indonesia 20 tahun yang lalu. Bagaimanakah profil konstruksi bunyi pada tahun 1990-an ini? Agar mendapatkan gambaran yang tepat, semua lema dalam KBBI 1993<sup>2</sup> ditransfer dalam bentuk satuan V atau K kemudian dihitung frekuensi kemunculannya, hasilnya adalah sebagai berikut:

Nomor	Konstruksi suku kata	Persentase	Contoh
1.	KV	49,52%	ra-pat
2.	KVK	35,48%	du-duk
3.	VK	5,05%	im-bang
4.	V	4,92%	i-tik
5.	KKV	2,14%	sas-tra
6.	KVV	1,11%	san-tai
7.	KKVK	1,07%	trak-tor
8.	KVKK	0,21%	kon-teks
9.	VV	0,17%	au-la
10.	VKK	0,08%	eks
11.	KKVKK	0,06%	kom-pleks
12.	KKKVK	0,05%	struk-tur
13.	KKKV	0,05%	in-stru-men

Berdasarkan pelacakan konstruksi bunyi bahasa Indonesia, terlihat dengan jelas bahwa 85% konstruksi suku kata bahasa Indonesia cenderung berbentuk KV (49,52%) dan KVK (35,48%). Sehingga kita dapat menyimpulkan bahwa konstruksi dominan dari suku kata bahasa

<sup>2</sup> Berbagai perhitungan serta ujicoba penerapan kaidah pemenggalan kata bahasa Indonesia pada Kamus Besar Bahasa Indonesia (A hingga Z) sebanyak 32.489 lema, tidak mungkin terwujud dalam waktu singkat tanpa bantuan Sdr. Bobby Nazief, PhD dan Sdr. Mirna Adriani, BSc dari Pusat Ilmu Komputer Universitas Indonesia.

Indonesia adalah KV dan KVK. Kedua konstruksi ini ditambah dengan beberapa bentukan lagi yaitu VK dan V masing-masing di atas 4%. Hasil perhitungan ini, lagi-lagi, memperkuat observasi para ahli bahasa terdahulu yang mendeteksi konstruksi dasar suku kata bahasa Indonesia adalah V, VK, KV, dan KVK.

Selain konstruksi dasar ditemukan pula bentukan KKV, KVV, dan KKVK yang masing-masing di atas 1 %. Sedangkan konstruksi lainnya yang di bawah 1% yaitu KVKK, VV, VKK, KKVKK, KKKVK, dan KKKV. Pengenalan tentang konstruksi bunyi merupakan salah satu langkah yang harus dilaksanakan untuk mempermudah membuat program komputerisasi pemenggalan kata.

Sebagaimana telah kita amati pada persentase kemunculan konstruksi bunyi, dari 13 jenis konstruksi ternyata di atas 90% hanya didominasi oleh empat konstruksi. Kesembilan konstruksi lainnya kemunculannya berkisar antara 1--0%. Melihat komposisi ini, jelaslah bahwa bahasa Indonesia mempunyai sistem tata bunyi yang mapan dan jelas konstruksinya.

Berdasarkan pemahaman itu, ketakutan akan merajalelanya kosa kata asing dalam himpunan kosa kata Indonesia kelihatannya tidak didukung oleh data. Kemudian, tergambar pula bahwa konstruksi bunyi lain di luar konstruksi dasar kemungkinan besar merupakan konstruksi bunyi baru. Konstruksi-konstruksi ini dapat disejajarkan dengan kosa kata pinjaman. Kosa kata pinjaman kita harapkan mengalami adaptasi fonologis sehingga mengikuti konstruksi dasar bunyi bahasa Indonesia.

### 3. Konstruksi Adaptasi Bunyi

Saya tidak begitu merisaukan jumlah kosa kata asing yang diperkirakan akan mendominasi perbendaharaan kosa kata Indonesia. Selama kata-kata asing itu tunduk pada konstruksi bunyi bahasa Indonesia, tidak akan menjadi masalah. Kata asing itu tidak akan tumbuh subur di kalangan pemakai bahasa Indonesia jika tidak mempunyai fungsi tertentu. Bahasa hanyalah media atau alat untuk mengutarakan perasaan dan pikiran si penutur, jika kosa kata Indonesia tak ada satu pun yang dapat mengekspresikan pikirannya secara tepat, maka sah saja jika si penutur meminjam kosa kata bahasa daerah atau bahasa asing. Lagipula,

masyarakat akan selalu berkembang seiring perkembangan jaman maka bahasa dengan sendirinya harus berkembang pula sesuai kebutuhan masyarakat pemakainya.

Saya lebih merisaukan masalah perambahan konstruksi bunyi asing ke dalam konstruksi bunyi bahasa Indonesia. Dalam hal ini, para ahli bahasa yang mempunyai wewenang harus bersikap menghadapi perusakan lingkungan konstruksi bunyi bahasa Indonesia.

Dalam konstruksi bunyi bahasa Indonesia yang dianggap asli tidak dikenal gugus konsonan. Presentase kemunculan konstruksi gugus konsonan yang paling tinggi hanyalah KKV (2,14%). Gugus konsonan yang sudah diakrabi oleh pemakai bahasa adalah gugus konsonan yang berasal dari kosa kata Jawa. Gugus konsonan itu pada umumnya hanya terdiri dari 2 konsonan dan konsonan kedua cenderung mengandung konsonan luncuran baik itu konsonan getar [r] maupun konsonan sampingan [l].

Adaptasi fonologis merupakan hal yang harus diperhitungkan dalam menerima kata baru. Saya menganggap bahwa adaptasi ini sebaiknya berlaku pada tataran fonologis maupun morfofonemik agar konstruksi bunyi bahasa Indonesia tidak kacau. Adaptasi fonologis bisa disejajarkan dengan sopan santun. Setiap masyarakat mempunyai adat kebiasaan yang berbeda. Sopan santun yang diharapkan dari seorang pendatang adalah usahanya menyesuaikan diri dengan adat setempat. Jika hal ini dilakukan, pada umumnya pendatang ini akan disenangi oleh penduduk setempat dan kehadirannya diterima sebagai salah seorang anggota keluarga masyarakat.

Jika sebuah kata asing masuk ke dalam khazanah kosa kata Indonesia, sopan santun yang diharapkan dari kata asing ini adalah penyesuaian konstruksi bunyinya dengan konstruksi bunyi bahasa Indonesia. Contoh:

*komputer, buku, meja, sepatu, lampu*

Selain itu jika bergabung dengan berbagai imbuhan, kata asing ini akan tunduk pada prosedur morfofonemik setempat. Jika hal ini terjadi, sudah barang tentu kita tidak merasa lagi bahwa kata itu adalah kata asing. Contoh:

*pikir berturunan pemikiran, terjemah berturunan menerjemahkan*

Konstruksi bunyi tidak dapat dipisahkan dari Pedoman Umum Pembentukan Istilah yang mengatur bagaimana kata-kata asing itu berintegrasi dengan konstruksi bunyi bahasa Indonesia. Mohon perhatikan kutipan mengenai Pedoman Umum Pembentukan Istilah (Edisi kedua berdasarkan Keputusan Menteri Pendidikan dan Kebudayaan Republik Indonesia, Nomor 0389/U/1988, tanggal 11 Agustus 1988 dan dicermatkan dalam Rapat Kerja Panitia Kerja Sama Kebahasaan, tanggal 16--20 Desember 1990), di bawah ini:

### 6.5 Penyesuaian Ejaan (KBBI 1993:1173):

Dalam perkembangannya, bahasa Indonesia menyerap pelbagai bahasa lain, baik dari bahasa daerah maupun bahasa asing, seperti Sanskerta, Arab, Portugis, Belanda, dan Inggris. Berdasarkan taraf integrasinya unsur serapan dalam bahasa Indonesia dapat dibagi atas tiga golongan besar

Pertama, unsur-unsur yang sudah lama terserap ke dalam bahasa Indonesia yang tidak perlu lagi diubah ejaannya. Misalnya, sirsak, iklan, otonomi, dongkarak, pikir, paham, aki.

Kedua, unsur asing yang belum sepenuhnya terserap ke dalam bahasa Indonesia, seperti shuttle cock, real estate. Unsur-unsur ini dipakai di dalam konteks bahasa Indonesia, tetapi pengucapannya masih mengikuti cara asing.

Ketiga, unsur yang pengucapannya dan penulisannya disesuaikan dengan kaidah bahasa Indonesia. Dalam hal ini diusahakan agar ejaan bahasa asing hanya diubah seperlunya sehingga bentuk Indonesianya masih dapat dibandingkan dengan bentuk aslinya.

### 6.6 Penyesuaian Huruf Gugus Konsonan Asing (KBBI 1993:1174)

Huruf gugus konsonan pada istilah asing yang tidak diterjemahkan dan diterima ke dalam bahasa Indonesia, sedapat-dapatnya dipertahankan bentuk visualnya. Kaidah penyesuaian ejaan yang diuraikan pada pasal 6.5 tetap berlaku dalam pelambangan huruf gugus konsonan itu.

#### Huruf Gugus Konsonan di Awal atau di Tengah

bd- (0) bdelium	bl- (22) blastula	br- (64) bromida	cz -(0) czardas	dr- (20) drama
fl- (25) fleksibel	lr- (26) afrikat	ft- (0) ftalein	gl- (51) glotis	gn- (1) gnomon
gr- (48) gradasi	kl- (83) klorofil	kr- (213) kriket	kn- (1) knebelit	pl- (85) plastik



KONSTRUKSI BUNYI BAHASA INDONESIA

pn- (3) pneumonia	pr- (181) prefiks	ps- (33) psikologi	pl- (0) pteridologi	sk- (40) skema
sl- (5) slidometer	skl- (2) sklerosis	skr- (4) skripsi	sm- (1) smaragdit	sn- (1) snobisme
sp- (68) spageti	sf- (5) sferulit	spl- (0) splenektomi	st- (145) stabil	str- (27) strategi
sv- (0) svedberg	sw- (31) swasta	tn- (1) tnesis		

Huruf Gugus Konsonan Akhir

-ft (0) lift	-ks (59) konteks	-lk (1) talk	-ld (0) kobold	-lf (1) golf
-lm (3) ilm	-ls (1) wals	-lt (3) basalt	-mf (0) limfa	-nk (2) bank
-ns (34) ons	-ps (6) gips	-rb (1) rubarb	-rk (3) patriark	-rd (1) fvord
-rg (1) erg	-rk (0) mark	-rn (17) isotern	-rp (1) skizokarp	-rf (6) mort'
-rps (1) korps	-rs (5) mars	-rt (3) introvert	-rtz (2) hertz	-sk (2) obelisk

Angka yang muncul di dalam setiap kotak merupakan jumlah kosa kata yang berkonstruksi seperti yang tertera pada kotak itu dalam KBBI 1993 dari A hingga Z.

Pertama-tama, dalam pedoman penyesuaian ejaan kata asing ke dalam konstruksi bunyi bahasa Indonesia. Rumusan itu yang menyatakan bahwa unsur asing yang pengucapannya dan penulisannya disesuaikan dengan kaidah bahasa Indonesia diusahakan agar ejaan bahasa asing hanya diubah seperlunya sehingga bentuk Indonesianya masih dapat dibandingkan dengan bentuk aslinya. Mengapa pedomannya seperti ini? Bukankah ini memberi peluang pada konstruksi bunyi asing untuk menyusup ke dalam konstruksi bunyi bahasa Indonesia?

Yang masih dipertanyakan adalah masalah *ejaan bahasa asing hanya diubah seperlunya*. Seberapa jauh disebut seperlunya? Ditambah lagi dengan pernyataan *sehingga bentuk Indonesianya masih dapat dibandingkan dengan bentuk aslinya*. Untuk kepentingan siapakah rumusan ini dibuat? Apakah ini dirumuskan untuk kepentingan para ahli bahasa agar mudah melakukan pelacakan asal-usul kata? Saya rasa, rumusan ini perlu dipertimbangkan kembali karena nyata-nyata tidak sejalan dengan konstruksi bunyi bahasa Indonesia yang dominan.

Berdasarkan hasil perhitungan pada semua lema di KBBI 1993, gugus konsonan yang berjumlah 40 itu, sebagian besar berfrekuensi sangat rendah kemunculannya dalam kamus. Gugus konsonan yang dominan ternyata {kr-} sebagaimana yang telah dikenal oleh para pemakai bahasa. Bahkan akhir-akhir ini gugus konsonan rangkap tiga juga diterima di kalangan masyarakat, pada umumnya gugus konsonan yang berkonstruksi konsonan geseran diikuti dengan konsonan hambat lalu diakhiri dengan konsonan getar, seperti {str-}.

Ada sejumlah ahli bahasa yang berpendapat bahwa pedoman ini dibuat khusus untuk kata-kata asing saja, jadi tidak perlu dirisaukan dampaknya terhadap kosa kata bahasa Indonesia. Lagi pula pedoman ini dianggapnya untuk memenuhi kebutuhan lapisan cendekia dalam masyarakat Indonesia. Saya justru beranggapan bahwa langkah ini kurang tepat. Perizinan semacam ini tanpa disadari akan berdampak luas dan merusak konstruksi bunyi bahasa Indonesia. Dalam perjalanan waktu, bahasa Indonesia akan mempunyai dua sistem bunyi. Sistem bunyi pertama untuk mengatur 96,08% kosa kata sedangkan sistem bunyi kedua untuk mengatur 3,92% kosa kata serapan asing.

Saya yakin bahwa ahli bahasa tidak akan mengalami kesulitan untuk menerapkan kedua sistem yang berdampingan itu. Bagaimana dengan para pemakai bahasa? Yakinkah kita bahwa masyarakat tidak akan mencampuradukkan kedua sistem bunyi itu? Bahkan ada kemungkinan, cepat atau lambat, sistem bunyi bahasa Indonesia akan menyerupai sistem bunyi bahasa asing. Apakah hal itu yang diinginkan? Bukankah kita juga perlu mempertahankan kepribadian konstruksi bunyi bahasa Indonesia?

Konstruksi gugus konsonan seharusnya dikaitkan dalam membahas pemenggalan kata. Hal ini penting untuk diketahui oleh para pemakai bahasa karena gugus konsonan itu tidak boleh diceraikan. Dengan demikian, komputer memerlukan daftar gugus konsonan agar dapat memenggal dengan baik. Pertanyaan yang muncul ke permukaan adalah apakah dalam kenyataan berbahasa sehari-hari konstruksi gugus konsonan dalam bahasa Indonesia memang berjumlah 40?

#### 4. Ejaan dan Pemenggalan Kata

Konstruksi bunyi bahasa Indonesia jarang sekali diutak-atik apalagi yang berkaitan dengan masalah pemenggalan kata. Kemungkinan, pemenggalan kata dianggap sesuatu yang tidak penting, otomatis dipahami dan tak bermasalah. Padahal proses pembuatan program komputerisasi pemenggalan kata dalam bahasa Indonesia menuntut pemahaman tentang konstruksi bunyi, baik konstruksi bunyi bahasa Indonesia maupun adaptasi konstruksi bunyi bahasa daerah atau asing yang masuk ke dalam himpunan kosa kata bahasa Indonesia.

Pemenggalan kata dalam bahasa Indonesia sangat erat kaitannya dengan sistem ejaan. Kedua hal itu boleh-boleh saja dianggap “sepele” namun keduanya pasti merupakan “batu sandungan” yang akan dihadapi oleh setiap pemakai bahasa Indonesia ragam tulis. Sebagai contoh:

*standar* berturunan *standardisasi* alih-alih *standarisasi*  
*efektif* berturunan *efektivitas* alih-alih *efektifitas*  
*teori* berturunan *teoretis* alih-alih *teoritis*

*kongres* harus dibaca *konggres*  
*linguistik* harus dibaca *lingguistik*

Contoh tersebut di atas dengan mudah dapat dijabarkan oleh seorang ahli bahasa, namun membingungkan para pemakai bahasa. Ejaan dan pemenggalan kata harus jelas kriterianya sehingga mudah dipelajari. Jika tidak, ada kemungkinan, konsentrasi penulisan banyak terserap pada masalah ejaan serta pemenggalan kata bukannya pada konstruksi kalimat. Saya menganggap, dalam hal ini, kepentingan pemakai bahasa harus diutamakan daripada kepentingan ahli bahasa. Lagipula masalah ejaan dan pemenggalan kata lebih merupakan perjanjian menulis daripada perjanjian tata bahasa.

Kaidah pemenggalan kata jika sudah dikomputerisasikan secara umum akan sangat membantu para pemakai bahasa. Khususnya, mengingat media cetak masa kini sudah menggunakan komputer. Kesalahan pemenggalan kata di berbagai surat kabar (kecuali halaman pertama) dan berbagai majalah dapat kita temukan dengan mudah.

Program pemenggalan kata yang tersedia tentunya untuk bahasa Inggris, sehingga para pekerja di media cetak harus bekerja ekstra keras mengejar batas waktu ditambah pemenggalan kata harus dilakukan satu per satu secara manual.

## 5. Pedoman Pemenggalan Kata

Berdasarkan ujicoba terdahulu, disimpulkan bahwa komputerisasi pemenggalan kata sulit untuk direalisasikan. Ada baiknya saya paparkan secara ringkas mengenai Pedoman Pemenggalan Kata. Kaidah pemenggalan kata yang kita punyai sudah mengalami beberapa kali perubahan<sup>3</sup>. Seingat saya sekurang-kurangnya ada tiga versi yaitu *Keputusan Presiden 1972* (lihat Buku EYD 1972), *Keputusan Menteri Pendidikan dan Kebudayaan 1987* (lihat KBBI 1993:1147--1164) serta *Keputusan Kerjasama Kebahasaan Majelis Bahasa Brunei Darussalam, Indonesia dan Malaysia 1991* (lihat KBBI 1993:1179--1181).

*Pedoman Pemenggalan Kata* (1991):

*"Pedoman ini merupakan penjabaran atas ketentuan tentang pemenggalan kata dalam Pedoman Umum Ejaan bahasa Indonesia Yang Disempurnakan. Pemenggalan kata ini bersangkutan dengan penulisan bukan pengucapan. Jadi, pemenggalan kata tidak sama dengan penyukuian kata (silubifikasi), yang lebih berhubungan dengan fonologi bahasa Indonesia. Prinsip yang dilaksanakan dalam pedoman pemenggalan kata ini ialah prinsip gramatikal dan prinsip ortografis. Prinsip fonologis dan prinsip etimologis sejauh mungkin dihindarkan".*

Komposisi Pedoman Pemenggalan Kata 1991 adalah sebagai berikut:

Pedoman Pemenggalan Kata (16 kaidah)	
2 kaidah berdasarkan prinsip gramatikal	14 kaidah berdasarkan prinsip ortografis

<sup>3</sup> Dalam makalah ini saya tidak membicarakan sejauh mana keputusan presiden bisa diganti oleh keputusan menteri atau keputusan kerjasama kebahasaan.

Ternyata, dari ke-16 kaidah itu, 2 rumus merupakan rumus yang dibuat berdasarkan prinsip gramatikal dan 14 rumus lainnya merupakan rumus yang dibuat berdasarkan prinsip ortografis. Jika melihat komposisi ini tentunya kita menduga bahwa bahasa Indonesia sistem pemenggalan katanya cenderung berdasarkan prinsip ortografis. Ternyata dugaan itu tidak benar, sebagaimana kita ketahui bersama, sistem pemenggalan kata bahasa Indonesia sangat gramatikal. Sebagai contoh:

*Kata makanan diharapkan terpenggal sebagai  
ma-kan-an alih-alih ma-ka-nan.*

Jika kosa kata bahasa Indonesia diharapkan terpenggal secara morfologis mengapa rumus berdasarkan prinsip ortografis jauh lebih banyak? Jika ditilik lebih jauh lagi, keperluan apakah yang mendorong lahirnya 14 rumus ortografis?

## 6. Ujicoba Komputerisasi

Secara prinsip, pembuatan program komputer untuk pemenggalan kata bukanlah hal yang sulit. Asalkan kaidah-kaidah pemenggalan kata dalam bahasa Indonesia sudah dirumuskan oleh ahli bahasa. Berdasarkan rumusan tersebut ahli komputer akan membuat sejumlah alur-perintah agar komputer dapat melakukan pemenggalan kata secara otomatis. Dalam hal ini ahli komputer bertugas sebagai penerjemah yaitu menerjemahkan rumus-rumus pemenggalan kata ke dalam komputer dalam bentuk perintah.

Komputer sangat efektif untuk melaksanakan hal-hal yang berpola dan berulang. Sebagaimana telah saya sampaikan di atas, komputerisasi pemenggalan kata bukanlah hal sulit asalkan kaidah-kaidah pemenggalan katanya sudah tersedia. Dari sudut pandang seorang ahli komputer, tentunya, kaidah pemenggalan kata yang baik adalah kaidah yang kriterianya jelas agar mesin dapat mengenali bentuk serta dapat melaksanakan pemenggalan secara tepat. Selain itu, semakin sedikit pengecualian yang muncul semakin baik kaidah pemenggalan katanya. Sedikitnya jumlah pengecualian berakibat langsung pada pemakaian

Ternyata, dari ke-16 kaidah itu, 2 rumus merupakan rumus yang dibuat berdasarkan prinsip gramatikal dan 14 rumus lainnya merupakan rumus yang dibuat berdasarkan prinsip ortografis. Jika melihat komposisi ini tentunya kita menduga bahwa bahasa Indonesia sistem pemenggalan katanya cenderung berdasarkan prinsip ortografis. Ternyata dugaan itu tidak benar, sebagaimana kita ketahui bersama, sistem pemenggalan kata bahasa Indonesia sangat gramatikal. Sebagai contoh:

*Kata makanan diharapkan terpenggal sebagai  
ma-kan-an alih-alih ma-ka-nan.*

Jika kosa kata bahasa Indonesia diharapkan terpenggal secara morfologis mengapa rumus berdasarkan prinsip ortografis jauh lebih banyak?. Jika ditilik lebih jauh lagi, keperluan apakah yang mendorong lahirnya 14 rumus ortografis?

## **6. Ujicoba Komputerisasi**

Secara prinsip, pembuatan program komputer untuk pemenggalan kata bukanlah hal yang sulit. Asalkan kaidah-kaidah pemenggalan kata dalam bahasa Indonesia sudah dirumuskan oleh ahli bahasa. Berdasarkan rumusan tersebut ahli komputer akan membuat sejumlah alur-perintah agar komputer dapat melakukan pemenggalan kata secara otomatis. Dalam hal ini ahli komputer bertugas sebagai penerjemah yaitu menerjemahkan rumus-rumus pemenggalan kata ke dalam komputer dalam bentuk perintah.

Komputer sangat efektif untuk melaksanakan hal-hal yang berpola dan berulang. Sebagaimana telah saya sampaikan di atas, komputerisasi pemenggalan kata bukanlah hal sulit asalkan kaidah-kaidah pemenggalan katanya sudah tersedia. Dari sudut pandang seorang ahli komputer, tentunya, kaidah pemenggalan kata yang baik adalah kaidah yang kriterianya jelas agar mesin dapat mengenali bentuk serta dapat melaksanakan pemenggalan secara tepat. Selain itu, semakin sedikit pengecualian yang muncul semakin baik kaidah pemenggalan katanya. Sedikitnya jumlah pengecualian berakibat langsung pada pemakaian

memori pada komputer dan kecepatan komputer melaksanakan pemenggalan kata.

Komputerisasi pemenggalan kata menemui jalan buntu. Kebuntuan ini bukannya karena belum ada kaidahnya tetapi karena kriteria dasar pemenggalannya berdasarkan prinsip gramatikal dan ortografis. Secara sederhana saja, apabila komputer harus memenggal sebuah kata bilakah kata itu dipenggal secara gramatikal atau secara ortografis? Para pemakai bahasa saja bingung apalagi komputer.

Sebagai ilustrasi, beberapa bahasa asing lainnya hanya menggunakan satu kriteria dasar. Ada yang gramatikal ada pula yang ortografis akan tetapi tidak kedua-duanya. Bahasa Inggris Amerika menggunakan kriteria ortografis, sedangkan bahasa Inggris Britania menggunakan kriteria gramatikal. Bahasa Perancis milih kriteria ortografis, sedangkan bahasa Belanda menggunakan kriteria gramatikal. Dan masih banyak bahasa lainnya untuk memperpanjang daftar, yang penting semua bahasa itu memilih satu kriteria saja. Bagaimana dengan bahasa Indonesia?

Saya rasa kita harus bersikap, jika tidak para pemakai akan terus mengeluh dan bingung karena tidak ada kejelasan. Sesuatu yang jelas bagi hali bahasa belum tentu jelas bagi pemakai bahasa. Menghadapi ulang tahun ke-50 republik ini, rasanya bahasa Indonesia harus menunjukkan kepribadiannya. Dalam hal ini, diperlukan keberanian mengambil sikap demi perkembangan bahasa Indonesia itu sendiri.

Tidak mudah untuk menentukan pilihan. Apakah sebaiknya bahasa Indonesia menggunakan prinsip gramatikal untuk pemenggalan kata atau lebih baik memilih prinsip ortografis? Untuk menjawab hal itu, di bawah ini akan ditampilkan hasil ujicoba.

#### **Ujicoba Prinsip Ortografis**

Pada ujicoba ini pemenggalan kata bahasa Indonesia hanya dilakukan berdasarkan prinsip ortografis. Dari keseluruhan lema yang ada dalam KBBI 1993 masih terdapat 3000 lema yang bermasalah. Masalahnya terletak pada kesalahan memenggal pada kata-kata yang mengandung diftong.

Distribusi diftong dan bukan diftong dalam bahasa Indonesia tidak sederhana. Lihat bagan di bawah ini:

**Distribusi Deret Vokal**

Deret huruf	Awal kata	Tengah kata	Akhir kata
ai	<i>ai-nul-ya-kin (3)</i>	<i>kai-sar (54)</i>	<i>ba-lai (402)</i>
a-i	<i>a-ib (3)</i>	<i>a-ja-ib (11)</i>	<i>nya-i (43)</i>
ae	<i>aes-te-li-ka (21)</i>	<i>maes-tro (2)</i>	- (0)
a-e	- (0)	<i>da-e-rah (9)</i>	<i>ho-na-e (1)</i>
au	<i>au-top-si (52)</i>	<i>sau-na (101)</i>	<i>be-li-au (236)</i>
a-u	<i>a-um (5)</i>	<i>sa-uh (22)</i>	<i>ma-u (6)</i>
ei	<i>ei-ko-san (3)</i>	<i>plei-doi (11)</i>	<i>ar-bei (6)</i>
e-i	- (0)	<i>a-te-is (6)</i>	<i>me-i (2)</i>
eu	<i>eu-fo-ni (18)</i>	<i>seu-da-ti (42)</i>	- (0)
e-u	- (0)	<i>pas-te-ur (5)</i>	- (0)
oi	<i>oi-ku-me-ne (1)</i>	<i>troi-ka (17)</i>	<i>a-soi (23)</i>
o-i	- (0)	<i>e-go-is (12)</i>	- (0)
ui	- (0)	<i>kuin-tal (9)</i>	<i>ci-hui (14)</i>
u-i	<i>u-ih (5)</i>	<i>ku-il (31)</i>	<i>un-cu-i (13)</i>

Angka yang muncul di dalam setiap kotak merupakan jumlah kosa kata yang berkonstruksi seperti yang tertera pada kotak itu dalam KBBI 1993 dari A hingga Z.

Salah satu kaidah pemenggalan menyatakan bahwa suku kata yang mengandung gugus vokal *au*, *ai*, *oi*, *ae*, *ei*, *eu*, dan *ui*, baik dalam kata-kata Indonesia maupun dalam kata-kata serapan, diperlakukan sebagai satu suku. Contohnya sebagai berikut:

*au-la; pu-lau; san-tai; am-boi; bai-tul-ha-ram;*  
*sur-vei; ae-ro-bik; eu-fe-mis-me; dan ka-sui*

Bandingkan dengan contoh di bawah ini:

*ka-in; la-uk; da-ur; dan da-un*

Akan tetapi perhatikan pemenggalan kata-kata berikut:

*Me-i; pre-i; da-i; dan sa-i*

Masalah yang muncul untuk komputer adalah bilamana deret huruf *au*, *ai*, *oi*, *ae*, *ei*, *eu*, dan *ui*, harus diperlakukan sebagai satu kesatuan yang tidak boleh dipisahkan dan kapan pula deretan vokal itu harus dipisahkan. Berdasarkan perhitungan secara umum, deretan huruf *au*, *ai*, *oi*, *ae*, *ei*, *eu*, dan *ui* yang disatukan 1091 lema sedangkan deretan huruf *au*, *ai*, *oi*, *ae*, *ei*, *eu*, dan *ui* yang dipisahkan 605 lema (dengan rincian



frekuensi konstruksi seperti dalam tabel). Melihat perbandingan jumlah tersebut, tampaknya tak ada gunanya untuk menyatakan mana yang merupakan pedoman dan mana yang merupakan pengecualian. Jika yang 605 lema dijadikan pengecualian, hal ini sama sekali tidak efektif.

Jika kita menerima pemenggalan secara ortografis maka praktis program komputerisasi pemenggalan kata hampir terwujud. Kesulitannya hanyalah pada diftong ini. Sebagian besar kosa kata dapat ditangani secara otomatis oleh komputer hanya saja konsekuensinya kita harus mau menerima hasil pemenggalan seperti:

*mengemban* terpenggal sebagai *me-ngem-ban*  
*mengembang* terpenggal sebagai *me-ngem-bang*  
*aturan* terpenggal sebagai *a-tu-ran*  
*angkatan* terpenggal sebagai *ang-ka-tan*

#### **Ujicoba Prinsip Gramatikal**

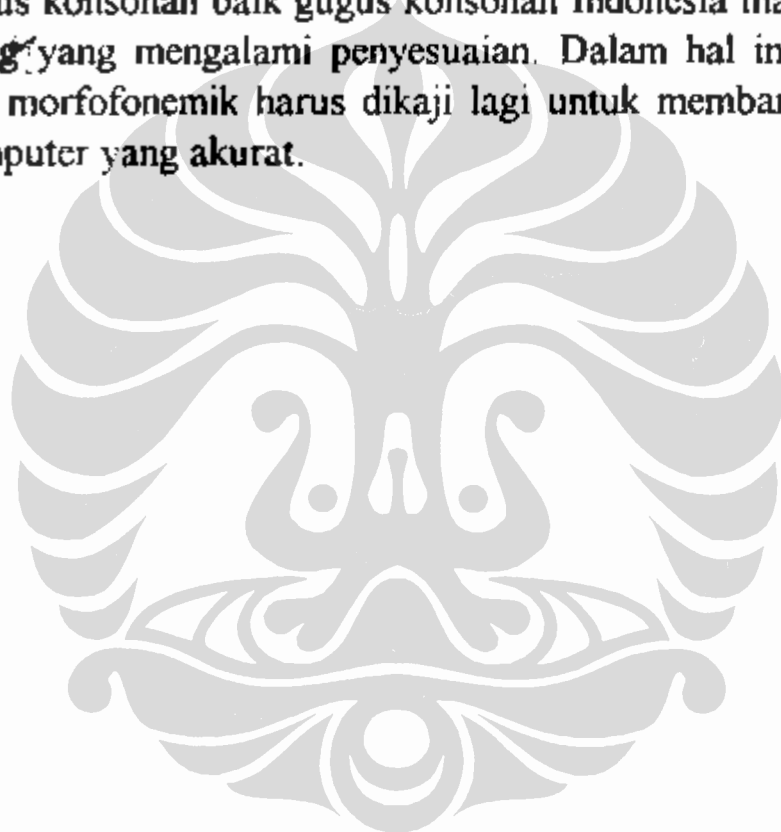
Pada ujicoba ini pemenggalan kata bahasa Indonesia hanya dilakukan berdasarkan prinsip gramatikal. Dari keseluruhan lema yang ada dalam KBBI 1993 hanya sebagian kecil yang tertangani secara otomatis. Sebagian besar bermasalah. Masalahnya terletak pada kesalahan memenggal pada kata-kata yang berimbuhan terutama yang kata dasarnya luluh di samping diftong. Proses morfofonemik bahasa Indonesia sulit dideteksi oleh komputer.

Jika, ingin menghasilkan pemenggalan otomatis yang menggunakan prinsip gramatikal, masih diperlukan penelitian-penelitian kecil untuk mendukung pemahaman konstruksi morfofonemik bahasa Indonesia. Salah satu kemungkinan yang terpikir saat ini yaitu membekali komputer dengan semua kata dasar dari A hingga Z. Setiap melaksanakan pemenggalan, komputer harus memeriksa terlebih dahulu dengan daftar konstruksi bunyi kata dasar. Hal ini memungkinkan akan tetapi terlalu memakan banyak memori serta pengoperasiannya memakan waktu. Sehingga lebih cepat melakukan pemenggalan secara manual dari pada secara otomatis.

## 7. Kesimpulan

Pertama, Komputerisasi pemenggalan kata akan dapat terwujud asalkan kita menentukan sikap. Prinsip mana yang akan dijadikan dasar kriteria pemenggalan. Prinsip gramatikal atau prinsip ortografis? Sebaiknya diusahakan membuat pedoman pemenggalan kata dengan satu prinsip saja supaya konsisten.

Kedua, pemenggalan kata hendaknya ditinjau secara bersamaan gugus konsonan baik gugus konsonan Indonesia maupun gugus konsonan asing yang mengalami penyesuaian. Dalam hal ini konstruksi fonologis dan morfofonemik harus dikaji lagi untuk membantu membuat program komputer yang akurat.



**Bahan Bacaan**

**Alieva, N. F. et al**

1991. *Bahasa Indonesia: Deskripsi dan teori*. Jakarta: Penerbit Kanisius

**Burton, D. M.**

1982. "Automated Concordances and Word-Indexes: Machine Decisions and Editorial Revisions," *Computers and the Humanities*. Vol. 16. Hlm. 195--218

**Butler, Christopher**

1985. *Computers in Linguistics*. New York: Basil Blackwell

**Halim, Amran**

1974. *Intonation: in Relation to Syntax in Bahasa Indonesia*. Jakarta: Lembaga Bahasa Nasional Departemen Pendidikan dan Kebudayaan.

**Hunt, R. and J. Shelley**

1983. *Computers and Commonsense*. 3rd edition. London: Prentice-Hall International.

**Maegaard, J. G.**

1976. "The Recognition of Finite Verb in French Texts," *ALLC Bulletin*, Vol. 4. Hlm 49--52.

**Oakman, R. L.**

1980. *Computer Methods for Literary Research*. Columbia: University of South Carolina Press.

MULTAMIA RMT LAUDER

**Microsoft Corporation**

1992. *Windows Resource Kit: Complete Technical Information for the Support Professional*. USA: Microsoft Corporation

**Tim Penyusun Kamus Pusat Pembinaan dan Pengembangan Bahasa**

1993. *Kamus Besar Bahasa Indonesia*. Edisi Kedua. Jakarta: Balai Pustaka

