

Kata Majmuk Bahasa Melayu: Isu dan Model Cadangan

Tuan Norhafizah Tuan Zakaria^[1], Mohd Juzaiddin Ab Aziz^[2], Mohd Rusmadi Mokhtar^[3] & Saadiyah Darus^[4]

^[1,2,3]*Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia, Selangor*
^[4]*Fakulti Sains Sosial dan Komunikasi, Universiti Kebangsaan Malaysia, Selangor*

tn_hafizah@yahoo.com

Abstrak. Pemajmukan merupakan satu proses yang merangkaikan dua kata dasar (atau lebih), membentuk satu kata baru yang mempunyai makna yang khusus. Secara amnya, kata majmuk mempunyai dua bahagian iaitu kepala (juga dipanggil inti) dan penerang. Kertas kerja ini akan membincangkan isu berkaitan dengan kata majmuk Bahasa Melayu dan mencadangkan satu model pembangunan bagi mengecam kata majmuk dalam ayat Bahasa Melayu. Model cadangan ini melibatkan sintaksis dan semantik Bahasa Melayu.

Keywords: kata majmuk; Bahasa Melayu; sintaksis; semantik; inti dan penerang

PENGENALAN

Menurut Tatabahasa Dewan [1], proses pemajmukan merupakan proses merangkaikan dua (atau lebih) kata dasar dan membawa makna yang khusus. Kata majmuk di eja terpisah dan bertindak sebagai satu unit, iaitu bentuknya tidak boleh menerima sebarang penyisipan unsur lain. Sebagai contoh, perkataan ‘kerusi malas’ merupakan satu kata majmuk kerana ia tidak boleh diselitkan dengan sebarang perkataan seperti ‘yang’. Kata majmuk biasanya dieja terpisah melainkan beberapa kata majmuk yang telah mantap. Jadual 1 menunjukkan contoh kata majmuk yang dieja terpisah dan kata majmuk yang telah mantap dan stabil.

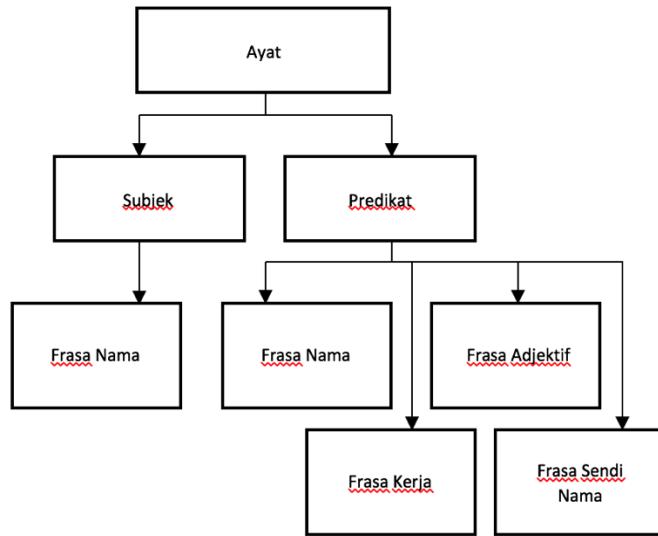
JADUAL (1). Bentuk Kata Majmuk

Bentuk Biasa	Bentuk Mantap
jalan raya	antarabangsa
tunjuk ajar	kakitangan
gempa bumi	tanggungjawab

Kata majmuk terbahagi kepada dua bahagian iaitu kepala (juga disebut inti) dan penerang. Kepala berada di bahagian hadapan (kiri) kata majmuk manakala penerang berada di bahagian kanan. Sebagai contoh, perkataan ‘meja’ pada kata majmuk ‘meja makan’ merupakan inti kata majmuk manakala ‘makan’ merupakan penerang, menerangkan fungsi meja tersebut sebagai tempat untuk makan.

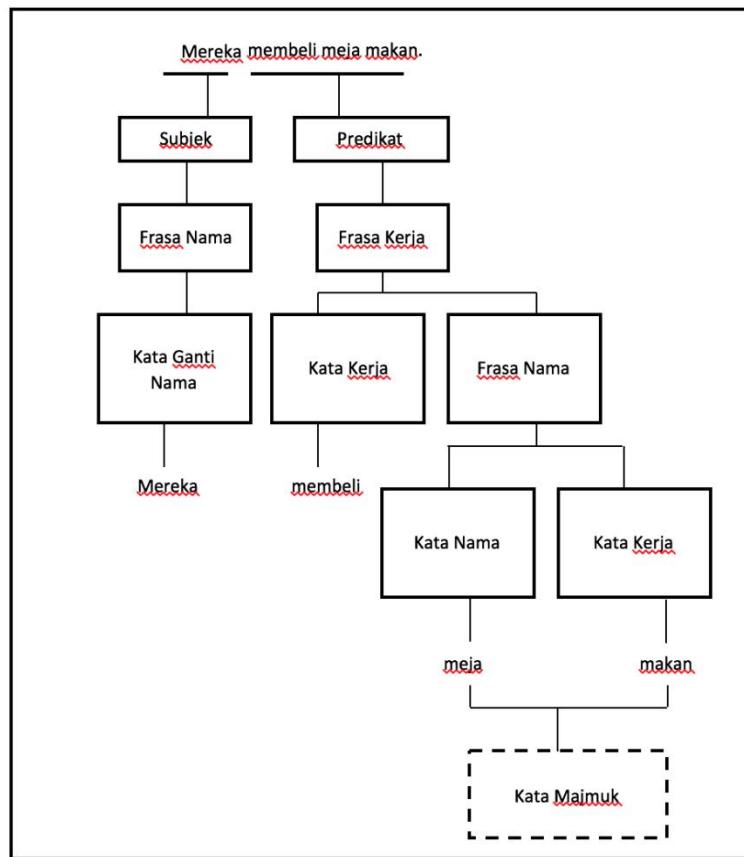
Ayat Bahasa Melayu boleh dibahagikan kepada dua bahagian utama iaitu Subjek dan Predikat. Subjek biasanya dibentuk daripada satu jenis frasa sahaja, iaitu Frasa Nama.

Manakala predikat ayat boleh dibentuk daripada empat jenis frasa iaitu Frasa Nama, Frasa Kerja, Frasa Adjektif atau Frasa Sendi Nama. Frasa pula terbentuk daripada gabungan beberapa perkataan yang boleh dikategorikan mengikut jenis kata seperti Kata Nama, Kata Kerja, Kata Adjektif dan sebagainya. Rajah 1 menunjukkan hirarki ayat Bahasa Melayu.



RAJAH 1. Hirarki Ayat Bahasa Melayu

Nik Safiah et.al.[1] menjelaskan bahawa satu kata majmuk Bahasa Melayu sebagai satu kata yang dibina hasil gabungan dua atau lebih kata dasar. Rajah 2 menunjukkan contoh kata majmuk dalam satu ayat Bahasa Melayu.



RAJAH 2. Contoh Kata Majmuk

Kata majmuk biasanya wujud pada frasa ayat. Untuk menjelaki kata majmuk dalam satu ayat, ayat perlu dihursti daripada bahagian atas ayat, melalui setiap aras ayat. Dalam Rajah 2, ‘meja makan’ merupakan kata majmuk yang berada pada frasa nama pada bahagian predikat ayat.

ISU KATA MAJMUK

Terdapat beberapa kajian berkisar tentang kata majmuk bahasa Melayu telah dijalankan oleh para pengkaji. Zaidi [2] telah menjalankan satu kajian berdasarkan korpus untuk mengekstrak kata majmuk bahasa Melayu. Korpus tersebut dikumpul daripada Sistem Korpus yang dibina oleh Dewan Bahasa & Pustaka (DBP). Beliau menggunakan satu sistem konkordans yang dibangunkan oleh DBP dan Universiti Sains Malaysia (USM) untuk mendapatkan kata majmuk. Beliau menyenaraikan beberapa kata kunci sebagai input, di mana sistem konkordans ini akan menyenaraikan output dalam bentuk jujukan konkordans. Pengecaman kata majmuk daripada output sistem dibuat secara manual.

Imran Ho et. al. [3] telah menjalankan kajian berdasarkan korpus menggunakan rujukan Hikayat Abdullah yang mengandungi 106605 perkataan sebagai data input bagi mengkaji perkataan-perkataan yang berkolokasi dengan perkataan ‘Melayu’. Dalam kajian ini, mereka menggunakan satu alat konkordans iaitu WordSmith bagi mendapatkan perkatan-perkataan yang berkolokasi dengan perkataan ‘Melayu’. Mereka menggunakan satu pengukuran statistik, Mutual Information untuk menyusun kedudukan output perkataan yang berkolokasi dengan perkataan ‘Melayu’. Antara output yang diperolehi adalah ‘Bahasa Melayu’, ‘orang Melayu’ dan ‘kitab Melayu’.

Kajian yang dilakukan oleh Suhaimi [4] adalah bertujuan untuk mengenalpasti kedudukan inti dan penerang yang hadir dalam setiap frasa nama bahasa Melayu. Beliau telah membina satu senarai petua heuristik yang digunakan untuk mengecam kedudukan inti dan penerang dalam frasa nama bahasa Melayu. Sebanyak 3100 ayat bahasa Melayu telah dianalisis bagi membina kerangka asas frasa nama (KAFN) dan kerangka baru frasa nama (KBFN). Kajian kerangka ini melibatkan empat jenis frasa iaitu frasa nama, frasa kerja, frasa adjektif dan frasa sendi nama. Kemudian, beliau membina rumus yang diimplementasi ke dalam bentuk linguistik komputer dengan menggunakan petua heuristik. Sebanyak 15 jenis petua heuristik telah dibina yang berfungsi untuk mengecam inti dan penerang yang wujud dalam frasa nama bahasa Melayu. Beliau menggunakan undang-undang DM (Diterangkan – Menerangkan) bagi mengecam inti dan penerang dalam satu ayat. Eksperimen yang dijalankan oleh beliau memberikan hasil kejituhan sebanyak 80.4% dan hasil dapatan semula sebanyak 55.1%. Selain aspek sintaksis, kajian ini turut melibatkan aspek semantik. Bagi aspek semantik, tujuan kajian tersebut adalah untuk mentafsirkan hubungan semantik yang wujud antara inti dan penerang frasa nama. Namun demikian, kajian ini hanya memfokuskan kepada ayat dasar frasa nama bahasa Melayu sahaja. Ayat yang dikumpul sebagai input kajian juga tertumpu kepada kemahiran yang dimiliki oleh para pelajar sekolah rendah daripada tahun satu hingga tahun lima sahaja. Kata majmuk yang dicam oleh beliau tertumpu kepada inti kata nama sahaja. Beliau tidak mengkaji kata kerja majmuk (contohnya urus niaga), kata adjektif majmuk (biru tua) dan kata majmuk setara (ibu bapa).

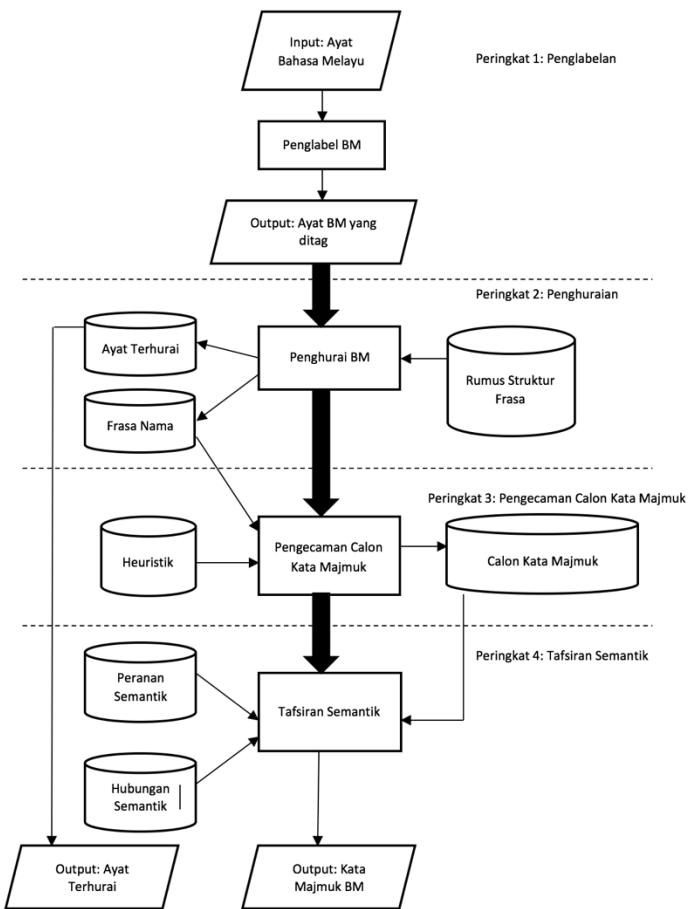
Kata majmuk boleh dicam menerusi binaan frasa dalam ayat. Namun demikian, untuk mengecam secara aplikasi komputer bukanlah satu tugas yang mudah. Setiap ayat perlu dipecahkan kepada setiap hirarki ayat. Bagi melatih aplikasi computer untuk mengecam frasa-frasa ini, petua heuristik perlu dibangunkan seterusnya membina algoritma komputer. Kebanyakan pakar bahasa menggunakan Rumus Struktur Frasa untuk mempersempit binaan frasa sesuatu ayat. Kajian kata majmuk bahasa Melayu

ini dilakukan secara meluas oleh ahli bahasa namun kajian secara rumus komputer untuk bahasa Melayu masih terhad dan boleh dikembangkan lagi.

Hubungan antara dua kata dasar dapat dipetakan menggunakan tafsiran semantik. Menurut Nik Safiah et. al. [1], frasa nama merupakan konstituen atau unsur yang menjadi pengisi frasa nama yang terdiri daripada kata nama sebagai unsur intinya. Unsur inti boleh wujud secara sendiri atau dengan sederetan penerang yang hadir sebelum frasa nama atau sesuadahnya, atau sebelum dan sesudah frasa nama. Ringkasnya, frasa nama terdiri daripada inti+inti dan inti+penerang. Ia mematuhi hukum D-M (Diterangkan – Menerangkan) dalam bahasa Melayu. D-M membawa maksud diterangkan (D) dan menerangkan (M). Dalam peraturan ini, unsur yang diterangkan merupakan unsur inti yang hadir pada kedudukan pertama dan unsur menerangkan merupakan unsur penerang yang hadir pada kedudukan kedua. Keadaan ini berbeza dengan bahasa Inggeris di mana kedudukan inti berada pada kedudukan kedua manakala kedudukan penerang (menerangkan) berada pada kedudukan pertama, contohnya *blue bird*.

MODEL CADANGAN

Rajah 3 mempersebahkan model cadangan bagi pengecaman kata majmuk dalam ayat Bahasa Melayu. Model cadangan ini mengandungi empat peringkat utama iaitu proses penglabelan jenis kata, penghuraian ayat kepada frasa dan jenis kata, pengecaman calon kata majmuk dan tafsiran semantik.



RAJAH 3. Model Cadangan

Empat peringkat utama dicadangkan pada model bagi mengecam kata majmuk dalam ayat Bahasa Melayu iaitu:

Peringkat 1: Praproses (Penglabelan Jenis Kata)

Satu penglabel jenis kata telah dibangunkan untuk kegunaan kajian ini. Penglabel ringkas ini dibangunkan menggunakan teknik berdasarkan kamus. Satu korpus telah dikumpul daripada sebuah buku bertajuk "Word Class in Malay: A Corpus-Based Approach" yang ditulis oleh Gerald and Zuraidah [5]. Korpus ini melibatkan empat novel yang telah dilabel dengan jenis kata Bahasa Melayu. 71 tagset telah digunakan mengikut Dewan Bahasa & Pustaka (DBP). 120,000 perkataan (15000 perkataan berbeza) telah dilabel dalam korpus tersebut. Jadual 1 menunjukkan contoh tagset oleh Gerald and Zuraidah [5]:

TABLE 1. Tagset

Bil	Set Label	Golongan Kata
1	#G	Nombor: 'Group' of n kumpulan
2	#K	Nombor: Kardinal
3	#M	Nombor: MeN- prefix
4	#O	Nombor: ordinal
5	#P	Nombor: Pecahan
6	#R	Nombor: Rumi
7	\$#	Matawang: RM

Tagset yang dikumpul diubah kepada tagset baru yang lebih ringkas bagi menyesuaikan dengan kajian ini. Tagset yang baru adalah mengikut Tagset Bahasa Melayu daripada Kamus Dwibahasa Malay-English [6]. Jadual 2 menunjukkan tagset baru yang digunakan dalam kajian kami.

JADUAL 2. Tagset Baru

Bil	Tagset	Tagset Lama	Tagset Baru
1	Kata Bilangan	#G, #K, #M, #O, #P, #R	KBIL
2	Matawang:RM	\$#	SYM
3	Adverbs	AA, AC, AF, AI, AL, AP, AT, AW	ADV
4	Bilangan (Kuantiti)	BK	KBIL
5	Kata Ganti Nama	DA, GD, GI, GK, GL	KGN
6	Kata Hubung	HB, HC, HR, HS	KH
7	Kata Sendi Nama	IA, IL	KSN
8	Kata Kerja	KA, KC, KE, KH, KI, KK, KP, KS, KT, KX	KK
9	Kata Nama	NA, NB, NG, NH, NL, NP, NS, NW	KN
10	Kata Nama Khas	NK	KNK
11	Kata Bantu	PA, PI, PM, PW	KB
12	Kata Adjektif	SA, SG, SK, SP, SS	ADJ
13	Kata Tanya	TA, TB, TL, TN	KT
14	Penanda Wacana	WI, WP, WS	PW
15	Kata Nafi	XA, XN	KNF

Peringkat 2: Penghuraian

Peringkat kedua kajian ini adalah menghuraikan ayat kepada frasa-frasa. Bagi proses penghuraian ayat Bahasa Melayu, satu rumus yang dipanggil sebagai ‘Rumus Struktur Frasa (RSF)’ akan digunakan. Amnya, ayat Bahasa Melayu dibina berdasarkan rumus tersebut. Rajah 4 menunjukkan rumus tersebut.

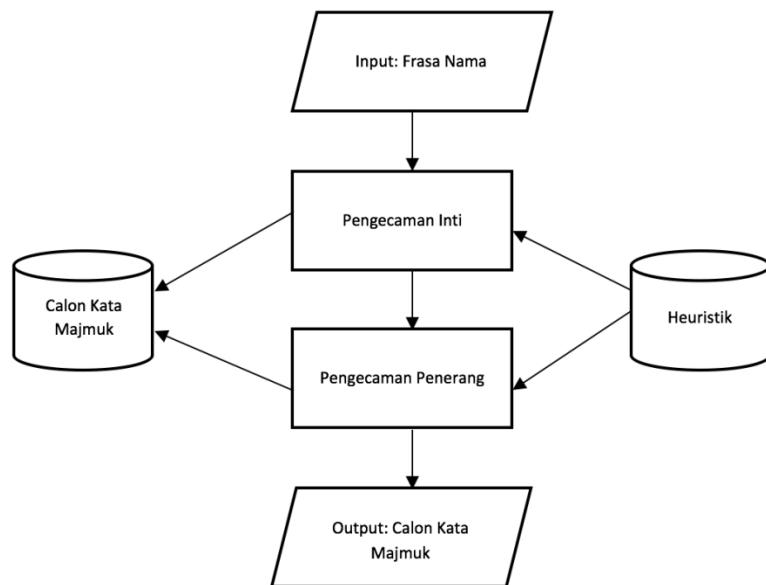
A	→	S + P
S	→	FN
P	→	[FN, FK, FA, FS]
FN	→	(KBil) + (PenBil) + (Gel) + KN + [KN] + (Pen) + PenBel
FK	→	(KB) + KKTTr + (Pel) + (Ket)
	→	(KB) + KKTr + KN + (Ket)
FA	→	(KB) + (KPeng) + KA + [KA] + (KPeng)
FS	→	KSN + (KN) + FN + (Ket)

RAJAH 4. Rumus Struktur Frasa

Berdasarkan kepada rumus pada Rajah 4, ayat (A) mengandungi S dan P (ayat A mengandungi subjek dan predikat). Subjek (S) hanya terdiri daripada FN (frasa nama) manakala bagi P (predikat), terdapat empat jenis frasa yang mungkin membentuk P iaitu FN (frasa nama), FK (frasa kerja), FA (frasa adjektif) or FS (frasa sendi nama). Setiap frasa mempunyai elemen tersendiri yang bergabung membentuk satu ayat Bahasa Melayu yang tepat.

Peringkat 3: Pengecaman Calon Kata Majmuk

Bagi proses pengecaman calon kata majmuk dalam ayat Bahasa Melayu, beberapa heuristik akan dibina berdasarkan sintaks. Heuristik yang dibina adalah berdasarkan kepada kategori linguistik (jenis kata). Kajian ini akan memfokuskan kepada tiga jenis kata yang utama iaitu Kata Nama, Kata Kerja dan Kata Adjektif. Rajah 5 menunjukkan carta alir bagi proses pengecaman calon kata majmuk. Input bagi proses ini adalah Frasa Nama. Inti dan penerang akan dicam berdasarkan heuristik yang dibina. Skop kajian ini akan memfokuskan kepada kata majmuk yang terdiri daripada dua perkataan yang terpisah sahaja sebagai calon kata majmuk.



RAJAH 5. Carta Alir Proses Pengecaman Calon Kata Majmuk

Peringkat 4: Tafsiran Semantik

Peringkat keempat kajian melibatkan kajian semantik bahasa Melayu. Input bagi peringkat ini adalah calon kata majmuk yang diperolehi daripada penghuraian ayat bahasa Melayu. Pada peringkat ini, skop kajian adalah hubungan inti dan penerang yang bersifat endosentrik sahaja. Tiga kategori sifat utama kata majmuk endosentrik iaitu koordinatif, atributif dan kumulatif akan dikaji dengan mendalam. Hubungan semantik antara inti dan penerang akan dipetakan.

KESIMPULAN

Kertas kerja ini membincangkan isu kata majmuk Bahasa Melayu dan kemudian mencadangkan model cadangan bagi mengecam kata majmuk dalam ayat Bahasa Melayu berdasarkan kepada kaedah sintaktik dan semantik. Model cadangan melibatkan empat proses utama iaitu penglabelan jenis kata, penghuraian ayat, pengecaman calon kata majmuk berdasarkan heuristik yang akan dibina serta tafsiran semantik antara inti dan penerang.

REFERENCES

1. Nik Safiah, K. Farid, M. O., Hashim, M. & Abdul Hamid, M. (2013). *Tatabahasa Dewan Edisi Ketiga*. Dewan Bahasa dan Pustaka.
2. Zaidi, I. (2008). *Kata Majmuk Bahasa Melayu: Bentuk Berdasarkan Data Korpus*. Dewan Bahasa dan Pustaka.
3. Imran, H. A., Ruzy, S. H. & Norhafizah, M. H. (2010). *Lexical Associations in Hikayat Abdullah: A Collocational Analysis*. Research Journal of Applied Sciences. 5(6). 429-433.
4. Suhaimi, A. R., Nazlia, O. & Mohd Juzaiddin, A. A. 2011. *Transformation of Malay Head Modifier Noun Phrase into a Thematic Relation Structure*. 2011 International Conference on Electrical Engineering and Informatics.
5. Gerald, O. K. & Zuraidah, M. D. (2006). *Word Class in Malay: A Corpus-Based Approach*. Dewan Bahasa dan Pustaka.
6. Hock, O. Y. (2009). *Kamus Dwibahasa Edisi Kedua*. Pearson Longman.