

Penterjemahan atau Penerjemahan: Satu Analisis Korpus Data Berkomputer

MUHAMMAD FAUZI JUMINGAN

Jabatan Bahasa Asing

Fakulti Bahasa Moden dan Komunikasi

Universiti Putra Malaysia, Serdang, Selangor

Kata kunci: Korpus data berkomputer, penterjemahan, penerjemahan

ABSTRAK

Leksikal *penterjemahan* dan *penerjemahan* sering diguna pakai secara silih berganti. Dari segi semantik, kedua-dua leksikal ini mempunyai makna yang sama. Tetapi dari segi morfologinya, adakah pembentukan kedua-dua leksikal ini masih sama atau adakah salah satunya mesti digugurkan? Ada pendapat mengatakan yang leksikal penterjemahan itu merupakan kata pinjaman bahasa Arab yang perlu dikekalkan huruf 't' nya. Pada masa yang sama, ada pendapat mengatakan bahawa kata pinjaman Arab itu sudah menjadi milik bahasa Melayu lantas huruf 't' nya mesti digugurkan. Justeru, yang manakah yang betul? Oleh sebab itu, makalah ini cuba merungkai masalah tersebut melalui satu analisis korpus data berkomputer yang dibina oleh Dewan Bahasa dan Pustaka (DBP) dengan kerjasama Universiti Sains Malaysia (USM). Dalam konteks perkembangan linguistik moden hari ini, penggunaan korpus linguistik sama yang berbentuk manual atau berkomputer, sudah mula diiktiraf sebagai salah satu data atau sampel linguistik yang sah dan boleh diguna pakai.

ABSTRACT

The lexical of "penterjemahan" and "penerjemahan" have been used by the translators in most of their works. Semantically, both of the particular lexical or words have the same meaning. But morphologically, is it the structure of both lexical still remain or either one must be deleted? There is an opinion that the lexical are borrowed from Arabic words and the alphabet of "T" must be dropped. At the same time, there is also another opinion about the lexical become the Malay word. Therefore, the "T" must be dropped. So, which is the correct one? According to this matter, this article aims to analyse the using of both lexical based on the corpus data that has been developed by the Dewan Bahasa dan Pustaka (DBP) in co-operation with Universiti Sains Malaysia (USM) since years ago. In the context of modern linguistics development, the use of linguistic corpus whether in manual form or computerised, is valid and consider as one of the real and original data or linguistic sample which can be used.

PENGENALAN

Leksikal *terjemah* merupakan satu perkataan yang diambil daripada bahasa asing, iaitu Arab. Leksikal ini sudah diterima pakai dan meresap dalam bahasa dan budaya Melayu sejak sekian lama. Berdasarkan leksikal ini muncullah kata-kata terbitan lain seperti *menterjemah*, *menterjemahkan*, *terjemahan*, *ketidakbolehterjemahan* dan *kebolehterjemahan*.

Namun begitu, berdasarkan kata nama apitan *penterjemahan*, lahir satu lagi leksikal yang berlainan bentuk morfologinya tetapi maksudnya masih sama, iaitu *penerjemahan*.

Ada sesetengah sarjana menggunakan *penerjemahan* (lihat Muhammad Bukhari Lubis dan rakan-rakan 1998) dan ada juga dalam kalangan mereka yang menggunakan *penterjemahan*. Malah, ada juga sesetengah institusi kebahasaan yang berwibawa seperti Dewan Bahasa dan Pustaka (DBP) yang menggunakan leksikal *penerjemahan*.

MASALAH DAN ISU KAJIAN

Menurut Kamus Dewan Edisi Ketiga (KD), apitan kata nama bagi leksikal *terjemah* itu

membawa maksud perihal (kerja, usaha, dsb.) diterjemahkan (KD, 1433). Dengan kata lain, leksikal tersebut membawa maksud proses diterjemah. Namun begitu, kata entri untuk penerjemahan tidak ada. Oleh itu, perkataan penerjemahan dianggap sebagai tidak gramatis, iaitu tidak mematuhi rumus nahu Melayu.

OBJEKTIF KAJIAN

- a) melihat wujudnya penggunaan leksikal penerjemahan dalam wacana bahasa Melayu;
- b) mengeksploitasikan penggunaan data korpus berkomputer yang dibangunkan oleh DBP; dan
- c) membetulkan persepsi kalangan pengguna bahasa Melayu bahawa leksikal penerjemahan adalah betul dan gramatis.

DEFINISI KAJIAN

Data Korpus Berkomputer

Sejarah data korpus bermula pada tahun 1964 apabila Francis dan Kucera membangunkan Korpus Universiti Brown (Rusli dan rakan-rakan 2004: 1). Penggunaan istilah korpus dan korpus berkomputer silih bertukar ganti.

Menurut Sinclair (dalam Rusli 2004: 4), korpus diertikan sebagai kumpulan cebisan bahasa (atau teks lengkap) yang dipilih dan disusun mengikut kriteria linguistik yang eksplisit dan digunakan sebagai sampel sesuatu bahasa.

Manakala korpus komputer pula ialah korpus yang diberi penanda, kod, dan diformatkan secara piawai serta dapat dicapai dan diproses dengan komputer.

Di bawah istilah korpus pula ada istilah subkorpus yang merupakan cabang atau mini korpus yang lebih kecil. Data korpus, lazimnya, mengandungi data yang berbentuk perkataan atau frasa dalam kuantiti yang sangat besar, iaitu jutaan perkataan. Misalnya, korpus *Birmingham Collection of English Texts (BCET)* yang membesar daripada 7.3 juta kata (pada pertengahan 1970an) kepada 20 juta kata menjelang 1985.

Pangkalan Data Korpus Bahasa Melayu DBP

Data korpus bahasa Melayu mula dibangunkan di Malaysia oleh DBP pada tahun 1983 yang diberi nama Projek Analisis Teks Secara Komputer (Rusli dan rakan-rakan: 2). Projek ini menyasarkan data teks sebesar dua juta kata

melalui teknik ala korpus Brown. Data ini diambil daripada pelbagai sumber seperti teks klasik, buku, majalah, akhlak, dan efemeral (seperti brosur dan risalah kecil).

Berdasarkan sasaran Kerja Utama DBP 2001-2005, projek ini menyasarkan pengumpulan data sebanyak 30 juta kata lalu menjadikan jumlah keseluruhan pada tahun 2005 nanti sebanyak 120 juta kata. Apabila siap sepenuhnya nanti, projek ini akan menggantikan sistem yang sedia ada yang dibuat dengan kerja sama USM yang bermula pada tahun 1994.

Data korpus DBP ini dikategorikan berdasarkan tipologi atau jenis teks tertentu. Tipologi ini kenali juga sebagai subpangkalan data yang mengandungi kira-kira 100 juta kata (sehingga Mac 2004). Berikut ini ialah jadual subpangkalan data yang diklasifikasikan berdasarkan tipologi teks:

Imbuhan (Apitan) Leksikal "Terjemah"

Ada empat jenis apitan yang membentuk kata nama terbitan, iaitu 1) *peN...-an*; 2) *pe-...-an*; 3) *peR-...-an*; 4) *ke-...-an*. Untuk apitan 1, *peN...-an*, ada mempunyai lima kategori lagi. Namun yang berkaitan dengan leksikal "penterjemahan" hanyalah kategori yang ketiga (TDEB 1994: 120-121).

Dalam kategori ketiga ini, *peN...-an* akan menjadi *pen-...-an* apabila digabungkan dengan kata dasar yang bermula dengan huruf *d, c, j, sy*, dan *z* serta kata pinjaman yang kata dasarnya bermula dengan huruf *t* dan *s* (TDEB: 121).

Contohnya:

dapat	-	pendapatan
cari	-	pencarian
jajah	-	penjajahan
syarikat	-	pensyarikatan
zalim	-	penzaliman

Kata pinjaman:

tadbir	-	pentadbiran
tafsir	-	pentafsiran
terjemah	-	penterjemahan
sintesis	-	pensintesisan
stabil	-	penstabilan
swasta	-	penswastaaan

Berdasarkan fakta di atas, sememangnya *t* dalam kata pinjaman akan kekal apabila mengalami apitan kata nama terbitan seperti "penterjemahan" dan "pentafsiran". Hal ini berbeza dengan huruf *t* yang bukan berbentuk

JADUAL 1
Data akhbar

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Akhbar	data 94,95,96, dan 98	10,111,504
Akhbar97	data 97	3,443,849
Akhbar99	data akhbar NSTP Online tahun 1999	6,055,096
Akhbar00	data akhbar NSTP Online tahun 2000	6,800,502
Akhbar01	data akhbar NSTP Online tahun 2001	4,825,314
Akhbar01-Ekonomi	data akhbar NSTP Online tahun 2001-ekonomi	147,924
Akhbar01-Hiburan	data akhbar NSTP Online tahun 2001-hiburan	239,035
Akhbar01-Sukan	data akhbar NSTP Online tahun 2001-sukan	926,910
Akhbar02	data akhbar NSTP Online tahun 2002	4,586,869
Akhbar02-Ekonomi	data akhbar NSTP Online tahun 2002-ekonomi	227,605
Akhbar02-Hiburan	data akhbar NSTP Online tahun 2002-hiburan	420,438
Akhbar02-Sukan	data akhbar NSTP Online tahun 2002-sukan	1,101,196
Akhbar03	data akhbar NSTP Online tahun 2003	5,114,146
Akhbar03-Ekonomi	data akhbar NSTP Online tahun 2003-ekonomi	74,690
Akhbar03-Hiburan	data akhbar NSTP Online tahun 2003-hiburan	676,615
Akhbar03-Sukan	data akhbar NSTP Online tahun 2003-sukan	1,163,734
JUMLAH NSTP		45,915,427
UTUSAN	data Utusan Online	6,448,577
HARAKAH	data Harakah Edisi Internet	624,699
Jumlah Data Akhbar		25,444,819

JADUAL 2
Data buku

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
DB3	data buku 70-an ke atas	11,137,717
DB2	data buku 70-an ke atas	9,739,899
DB1	data buku 60-an ke bawah	2,759,585
DB4	data buku 70-an ke atas	1,807,618
Jumlah Data Buku		25,444,819

JADUAL 3
Data majalah

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Majalah	data majalah	4,861,827
Majalah1	data majalah tambahan	3,361,029
Majalah Ilmiah	data majalah ilmiah	1,887,516
Majalah bukan Ilmiah	data majalah bukan ilmiah	2,119,001
Jumlah Data Majalah		12,229,373

JADUAL 4
Data teks Melayu lama/tradisional

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Klasik	Data teks Melayu lama atau teks tradisional	2,440,258
	Jumlah Data Teks Melayu Lama	2,440,258

JADUAL 5
Data teks terjemahan

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Terjemah	Data terjemahan ke dalam bahasa Melayu	1,886,106
	Jumlah Data Terjemahan	1,886,106

JADUAL 6
Data teks Sabah dan Sarawak

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Sukuan	Data bahasa Melayu terbitan Sabah & Sarawak	1,038,250
	Jumlah Data Bahasa Melayu Sabah dan Sarawak	1,038,250

JADUAL 7
Data buku teks

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Buku teks	Data buku teks sekolah	1,095,726
	Jumlah Data Buku Teks	1,095,726

JADUAL 8
Data teks drama

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Drama	Data drama	215,867
	Jumlah Data Drama	215,867

JADUAL 9
Data efemeral

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Efemeral	Data efemeral	173,131
	Jumlah Data efemeral	173,131

JADUAL 10
Data puisi

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Puisi	data puisi	2,348
	Jumlah Data Puisi	2,348

JADUAL 11
Data kad bahan

Subpangkalan	Keterangan	Kata (sehingga Mac 2004)
Kad bahan	Rekod kad bahan (kutipan kad bahan untuk penyusunan kamus)	3,130,641
	Jumlah Data Kad Bahan	3,130,641

JADUAL 12
Jumlah kumulatif data teks (sehingga Mac 2004)

Subpangkalan	Jumlah kata	Peratus
Akhbar	52,988,703	52.65%
Buku	27,797,010	27.62%
Majalah	12,229,373	12.15%
Teks tradisional	2,440,258	2.43%
Lain-lain	3,303,772	3.28%
Terjemahan	1,886,106	1.87%
Jumlah	100,645,222	100.00%

Sumber: Semua data di atas diambil daripada makalah Rusli Abdul Ghani dan rakan-rakan (2004: 8-12)

pinjaman yang akan luluh t-nya dan digantikan dengan bentuk sengau yang dilafazkan pada tempat yang sama daerah artikulasinya (TDEB: 121). Misalnya seperti:

tendang - penendang
tentu - penentuan
tamat - penamat

CONTOH DATA KORPUS (KONKORDONS)

Dalam subtopik ini, contoh data korpus atau konkordans yang membuktikan wujudnya penggunaan leksikal "penerjemahan" akan ditunjukkan. Subkorpus yang mengandungi leksikal "penerjemahan" ialah data akhbar (akhbar 94,95,96,98), data db2 (data buku 70an ke atas), data majalah ilmiah, data majalah (majalah bukan ilmiah), dan data majalah 1

(data majalah tambahan).

Antara sampel-sampel tersebut adalah seperti:

FILE : akhbar (akhbar 94,95,96,98)

1 i~D8endinginkannya kembali****?
Penerjem****ahan Nurhalim ini juga mampu
m

FILE : db2 (buku 70an ke atas)

1 ku - Pelajaran dan Bahagian
Penerjemahan diwujudkan untuk menerbi
2 Bonn kepada usaha-usaha ***penerjemahan***,
dan Passau kepada kajian
3 i. Tetapi penerbitan dan ***penerjemahan***
buku-buku ilmu pengetahu
4 ns dan teknologi. Dengan ***penerjemahan***,
mereka dapat membawa mas

JADUAL 12
Kekerapan leksikal “penerjemahan”

Subkorpus	Jumlah bilangan kata	Kekerapan Leksikal “Penerjemahan”
Akhbar (94,95,96,98)	10,111,504	1
DBP2 (buku 70-an ke atas)	9,739,899	18
Majalah Ilmiah	1,887,516	2
Majalah 1 (majalah tambahan)	3,361,029	4
Majalah(majalah bukan ilmiah)	4,861,827	2
Jumlah Kekerapan		27

5 il kira keperluan seperti ***penerjemahan*** melalui komputer, pempro 2 a hendak didasarkan pada ***penerjemahan*** karya tersebut \$\$20 ke d

FILE : majalah-ilmiah

1 arangnya dengan melakukan ***penerjemahan*** cerpen-cerpen yang diha
2 mahan. Rupa-rupanya kerja ***penerjemahan*** itu dapat membentuk land

FILE : majalah 1 (data majalah tambahan)

1 namaan khas, dan panduan ***penerjemahan*** umum); dan pelaksanaan
2 dan teknologi; penulisan/***penerjemahan*** dan penerbitan buku (umum
3 yediakan kemudahan sistem ***penerjemahan*** bahasa Melayu kepada bah
4 lajaran, penulisan buku, ***penerjemahan***, penyelidikan) dan apl

FILE : majalah (bukan ilmiah)

1 umbangan terbarunya ialah ***Penerjemahan*** Arab-Melayu: Peraturan d

RUMUSAN DAN KESIMPULAN

- a) Tidak salah pengguna bahasa Melayu menggunakan leksikal “penerjemahan” kerana ia memang wujud dalam korpus data. Berdasarkan pendapat ahli linguistik moden, analisis terhadap sesuatu bahasa haruslah berdasarkan penggunaan bahasa itu dalam konteks yang “sebenar”. Oleh sebab korpus data merupakan himpunan kata yang berbentuk penggunaan “sebenar” maka sewajarnya leksikal “penerjemahan” boleh diterima dan digunakan.
- b) Leksikal “penerjemahan” merupakan perkataan yang gramatis dari segi nahunya. Meskipun ia berbentuk kata pinjaman tetapi apabila ia telah masuk dan meresap ke dalam bahasa Melayu maka ia menjadi milik bahasa Melayu. Oleh itu, pembentukan morfologinya haruslah mengikut rumus nahu Melayu. Sama seperti perkataan lain yang mengalami luluhan huruf *t* dengan

JADUAL 13
Kekerapan leksikal “penterjemahan” dalam korpus data bahasa Melayu DBP

Subkorpus	Jumlah Bilangan Kata	Kekerapan Leksikal “Penterjemahan”
Akhbar (94,95,96,98)	10,111,504	85
DBP2 (buku 70-an ke atas)	9,739,899	63
Majalah Ilmiah	1,887,516	63
Majalah 1 (majalah tambahan)	3,361,029	197
Majalah (majalah bukan ilmiah)	4,861,827	=
Jumlah Kekerapan		408

JADUAL 14
Kekerapan peratusan leksikal “penterjemahan” dan “penerjemahan”

Kata Leksikal	Kekerapan	Peratusan
Penterjemahan	27	6.20%
Penerjemahan	408	93.79%
Jumlah	435	99.99% = 100%

dibunyikan sengau maka leksikal “penerjemahan” harus bersifat demikian.

- c) Sesuatu kajian yang mantap dan tuntas haruslah berdasarkan penggunaan “sebenar” sesuatu bahasa itu oleh penutur jati bahasa tersebut. Penggunaan “sebenar” ini ada dihindarkan dalam korpus data. Oleh itu, penggunaan korpus data sebagai data analisis bahasa amatlah dialu-alukan. Sebagai pengkaji bahasa juga, kita tidak boleh mencipta sampel bahasa secara imaginasi sahaja tanpa didasarkan pada penggunaan sebenar. Itulah sebabnya, ramai ahli linguistik moden hari ini lebih menumpukan kajian bahasa dan linguistik secara deskriptif dan bukan perskriptif. Kajian deskriptif lebih ideal dan “berpijak di bumi nyata” berbanding kajian perskriptif yang lebih bersifat “khayalan dan awang-awangan”. Justeru, deskriptif lebih popular daripada perskriptif.

RUJUKAN

- ASMAH HJ. OMAR. 1993. *Nahu Melayu Mutakhir Edisi Keempat*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- _____. 2004. *Data Korpus Bahasa Melayu DBP*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- _____. 2000. *Kamus Dewan Edisi Ketiga*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

MUHAMMAD BUKHARI LUBIS, MOHD FAUZI JUMINGAN, dan KHAIRUNNAWA MOHAMED. 1998. *Penerjemahan Arab-Melayu-Arab: Peraturan dan Dasar*. Serdang: Penerbit UPM.

NIK SAFIAH KARIM dan RAKAN-RAKAN. 1994. *Tatabahasa Dewan Edisi Baharu*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

NOR HASHIMAH JALALUDDIN. 2004. Kata ganti nama ia dan ianya dalam bahasa Melayu: pendekatan rangka rujuk silang. Kertas Kerja *Seminar Sehari Linguistik*, anjuran Program Linguistik UKM.

NORWATI Md. YUSUF. 2001. Penggunaan korpus dalam kajian bahasa dan terjemahan berlandaskan terjemahan mesin. Dalam *Prosiding Seminar Sehari Jabatan Bahasa Melayu dan Terjemahan*, anjuran UKM dan DBP.

RUSLI ABDUL GHANI, NORHAFIZAH MOHAMED HUSIN, dan CHIN LEE YIM. 2004. Pangkalan data korpus DBP: perancangan, pembinaan, dan pemanfaatan. Kertas Kerja *Seminar Sehari Linguistik*, anjuran Program Linguistik UKM.

ZAHARANI AHMAD. 2004. Pengimbuhan memperkan dengan kata adjektif: analisis data korpus berkomputer. Kertas Kerja *Seminar Sehari Linguistik*, anjuran Program Linguistik UKM.

(Received: 21 December 2004)