

PRODUKTIVITI IMBUHAN *MEM-*, *MEMPER-* DAN *MEMPER-KAN*: PENDEKATAN LINGUISTIK KORPUS

*(Affix Productivity of Mem-, Memper- and Memper-kan:
A Corpus Linguistics Approach)*

Imran Ho bin Abdullah

imranho@ukm.edu.my

Pusat Pengajian Bahasa dan Linguistik,
Fakulti Sains Sosial dan Kemanusiaan,
Universiti Kebangsaan Malaysia,
40360 Bangi, Selangor, Malaysia.

Hazimah binti Yusof

hazimah@dbp.gov.my

Bahagian Penyelidikan Bahasa,
Dewan Bahasa dan Pustaka,
Jalan Dewan Bahasa,
50460 Kuala Lumpur, Malaysia.

Tel.: +603-89215555

Abstrak

Kajian ini bertujuan untuk mengukur dan membandingkan produktiviti morfologi imbuhan antara imbuhan awalan *mem-* dan *memper-* dan imbuhan apitan *memper-kan* dalam bahasa Melayu. Kedua-dua bentuk imbuhan berkenaan berfungsi untuk menerbitkan kata kerja dan kata adjektif (terbitan) daripada kata akar (KN, KK, KT, KAdj). Kajian ini juga melihat sama ada imbuhan *mem-* yang ditambah dengan morfem *per-* dan *-kan* menjadi imbuhan *memper-* dan *memper-kan*, mempengaruhi produktiviti morfologi bahasa Melayu.

Selain itu, dengan menggunakan dan membandingkan dua bentuk korpus, iaitu korpus Akhbar Sukan Dalam Talian yang berjumlah hampir 2.5 juta kata dan korpus UKM-DBP yang berjumlah lima juta kata turut dikaji untuk melihat sama ada saiz korpus mempengaruhi morfologi produktiviti imbuhan berkenaan. Dengan menggunakan perisian Wordsmith Tools 4.0, data korpus berkenaan dijana untuk mengisihkan bentuk imbuhan *mem-*, *memper-* dan *memper-kan* dengan imbuhan lain dan kata akar. Berdasarkan ukuran produktiviti Baayen (1989, 1992), produktiviti proses pengimbuhan tersebut telah diukur. Kajian ini mendapati bahawa imbuhan *memper-* mempunyai nilai produktiviti yang tinggi iaitu $P=0.009$ dan $P=0.030$ manakala imbuhan *mem-* mempunyai produktiviti yang paling rendah iaitu $P=0.002$ dan $P=0.001$ bagi kedua-dua korpus berkenaan. Penambahan morfem *per-* dan *-kan* daripada imbuhan *mem-* menjadi imbuhan *memper-* dan *memper-kan* tidak mempengaruhi sepenuhnya produktiviti sebaliknya dipengaruhi oleh *hapax*. Hal ini dapat dilihat imbuhan yang mempunyai morfem terbanyak, iaitu *memper-kan* hanya menduduki tangga kedua dengan nilai $P=0.004$ dan $P=0.008$ bagi kedua-dua korpus berkenaan. Namun begitu, saiz korpus ternyata mempengaruhi nilai produktiviti. Nilai produktiviti imbuhan *memper-* dalam korpus UKM-DBP ($P=0.030$) mempunyai julat perbezaan yang agak besar berbanding korpus Akhbar Sukan Dalam Talian ($P=0.009$), iaitu $P=0.021$.

Kata kunci: ukuran produktiviti, awalan, apitan, morfologi Melayu, data korpus sukan, data korpus umum, linguistik, komputer

Abstract

This study measures and compares the morphological productivity of the prefixes mem- and memper- and the circumfix memper-kan in Malay. The data used in this study are comprised of two corpora, that is Sports News Online corpus that contains nearly 2.5 million words and an UKM-DBP Corpus which totals five million words. This study also examined whether the size of the corpus has any effect on the morphology productivity of the affixes studied. Wordsmith Tools 4.0 was used to generate and sort the affixes mem-, memper- and memper-kan together with other affixes and root words. Using Baayen's productivity measure, the productivity of the affixes was measured. The results show that the productivity value of the affix memper- have a higher productivity index of $P=0.009$ and $P=0.030$ compared to the affix mem- which has a lower productivity index of $P=0.002$ and $P=0.001$ for both corpora respectively. Adding the

per- and -kan morpheme from mem- affix to memper- and memperkan affix does not fully affected productivity, instead is influenced by hapaxes. The affix which has the highest morpheme, that is memperkan, only ranked second with a value of $P=0.004$ and $P=0.008$ for both corpora concerned. However, the size of the corpus affected the productivity value. The productivity of morphological value memperin the UKM-DBP corpus ($P=0.030$) has a relatively large range of difference compared to the Sports News Online corpus ($P=0.009$), that is $P = 0.021$.

Keywords: productivity measure, prefix, circumfix, Malay morphology, sport corpus data, general corpus data, linguistics and computer

PENDAHULUAN

Bahasa Melayu seperti bahasa lain di dunia menghasilkan perkataan baharu yang membawa konsep baharu melalui proses morfologi, dan salah satu daripadanya menerusi penerbitan kata. Penerbitan kata ialah proses mewujudkan kata baharu dengan makna dan konsep baharu daripada kata dasar atau kata akar sedia ada. Kata dasar ialah kata asli yang wujud sebagai satu kata dan belum pernah melalui sebarang proses morfologi.

Salah satu proses penerbitan kata yang paling produktif dalam bahasa Melayu ialah pengimbuhan. Pengimbuhan kata dalam bahasa Melayu dibahagikan kepada empat jenis, iaitu awalan, akhiran, sisipan, dan apitan.

Dalam kajian ini, imbuhan awalan *mem-* dan *memper-* dan imbuhan apitan *memper-kan* diukur dengan menggunakan ukuran empirikal produktiviti. Dengan adanya ukuran empirikal, darjah produktiviti dua bentuk proses imbuhan, iaitu pengimbuhan awalan (*mem-* dan *memper-*) dan imbuhan apitan (*memper-kan*) dapat dinilai dan dibandingkan. Dalam kajian ini juga akan dilihat sama ada penambahan morfem *per-* dan *-kan* dari imbuhan *mem-* menjadi *memper-* dan *memper-kan* dalam pembentukan morfologi bahasa Melayu dan saiz data korpus mempengaruhi ukuran produktiviti morfologi bahasa Melayu.

Kajian ini menggunakan ukuran produktiviti Baayen (1989, 1992) untuk mendapatkan indeks produktiviti berdasarkan pendekatan korpus. Indeks produktiviti Baayen mengukur produktiviti sebagai kebarangkalian sesuatu kata terbitan baharu yang menggunakan sesuatu imbuhan akan ditemui dalam korpus berdasarkan kekerapan kata dalam korpus. Ukuran produktiviti berdasarkan korpus dapat menggambarkan cara

sesuatu imbuhan itu digunakan. Kajian linguistik berasaskan pendekatan korpus juga semakin penting seiring dengan perkembangan linguistik pengkomputeran dan keupayaan teknologi komputer dalam menganalisis data teks yang besar.

IMBUHAN AWALAN DAN APITAN BAHASA MELAYU

Imbuhan merupakan salah satu bentuk proses morfologi bahasa Melayu yang menghasilkan kata terbitan selain pemajmukan kata dan penggandaan kata yang bertujuan untuk menambahkan lagi perbendaharaan kosa kata bahasa Melayu. Imbuhan bahasa Melayu terbahagi kepada empat jenis, iaitu awalan, akhiran, sisipan dan apitan. Dalam kajian ini hanya imbuhan awalan *mem* dan *memper-* dan imbuhan apitan *memper-* sahaja yang dikaji untuk melihat produktiviti morfologinya sebagai satu kajian awal. Imbuhan *mem-* dan *memper-* merupakan dua bentuk varian imbuhan awalan selain *me-*, *men-* dan *meng-*. Imbuhan *mem-* hanya dapat bergabung dengan huruf [b] dan kata pinjaman [f], [p] dan [v] kata dasar golongan kata kerja dan kata adjektif menghasilkan kata kerja dan kata adjektif seperti contoh yang berikut:

(1) bakar		membakar
(2) bisu		membisu
(3) promosi		mempromosi
(4) pesona	→	mempersona
(5) fail		memfail
(6) fitnah		memfitnah
(7) veto		memveto

Jika diimbuhkan dengan kata nama yang bermula dengan huruf vokal dan huruf selain [b], [p], [f] dan [v], ternyata tidak sesuai dan salah dari segi tatabahasa. Contohnya:

(1) hamba	→	memhamba
(2) ekor	→	memekor

Bagi imbuhan awalan *memper-*, walaupun sepatutnya dirujuk sebagai imbuhan *per-* dalam bahasa Melayu, tetapi sebenarnya *per-* tidak boleh wujud dalam ayat tanpa awalan *meN-* kecuali dalam ayat perintah. Jadi, imbuhan *per-* perlu bergabung dengan imbuhan *meN-* menjadi *memper-*

untuk membentuk kata baharu. Imbuan awalan *memper-* dapat bergabung dengan kata dasar yang terdiri daripada kata nama, kata adjektif, kata tugas dan kata kerja (walaupun bersifat lewah tetapi penggunaannya meluas) menghasilkan bentuk kata kerja. Contoh kata terbitan imbuan *memper-* adalah seperti yang berikut:

(1) **kata nama**

alat	→	memperalat
daya		memperdaya

(2) **kata adjektif**

besar	→	memperbesar
luas		memperluas

(3) **kata tugas (kata sendi)**

oleh	→	memperoleh
------	---	------------

(4) **kata kerja (bersifat lewah tetapi penggunaan meluas)**

buat	→	memperbuat
------	---	------------

Bagi imbuan apitan *memper-kan*, golongan kata yang boleh bergabung dengan imbuan *memper-kan* ini ialah kata nama, kata kerja dan kata tugas. Walau bagaimanapun, perselisihan pendapat dalam kalangan ahli bahasa tempatan timbul apabila melibatkan kata akar daripada kumpulan kata adjektif. Contoh kata terbitan *memper-kan* adalah seperti yang berikut.

(1) **kata nama**

isteri	→	memperisterikan
hamba		memperhambakan

(2) **kata kerja**

juang	→	memperjuangkan
main		mempermainkan

(3) **kata tugas**

satu	→	mempersatukan
banyak		memperbanyakkan

Bagi imbuhan *memper-kan* yang bergabung dengan kata adjektif, telah menimbulkan perselisihan pendapat dalam kalangan ahli bahasa tempatan. Bagi Abdullah (1974), Arbak (1981) dan Asmah (1982), menyatakan bahawa imbuhan *memper-kan* seperti memperbesarkan boleh bergabung dengan kata adjektif membentuk kata kerja terbitan. Contohnya besar → *memperbesarkan*, dalam → *memperdalamkan*. Hal ini bertentangan dengan Nik Safiah *et al.* (1997) dan Asraf (1990) yang menyatakan imbuhan *memper-kan* tidak boleh bergabung dengan kata adjektif kerana dianggap lewah dalam bahasa Melayu. Namun, berdasarkan penggunaan produktiviti imbuhan ini yang tinggi dalam data korpus DBP, ternyata imbuhan *memper-kan* boleh digunakan walaupun bertentangan dengan prinsip kepadaan pemerhatian (Zaharani, 2007). Dalam kajian ini, imbuhan awalan *mem-* dan *memper-* serta imbuhan apitan *memper-kan* akan dibandingkan dan diukur menggunakan ukuran produktiviti morfologi dengan pendekatan linguistik korpus.

PRODUKTIVITI MORFOLOGI

Kajian berkenaan produktiviti morfologi yang menggunakan pendekatan linguistik korpus bukanlah satu bidang atau perkara baharu dalam kalangan ahli linguistik Barat. Hockett (1958) misalannya mentafsirkan “produktiviti” sebagai satu ciri morfologi yang membolehkan penutur membentuk gabungan baharu yang belum pernah digunakan sebelum itu. Bauer (2001:1) menyatakan bahawa istilah “produktiviti morfologi” membawa pelbagai pengertian dan beliau memberikan tafsiran produktiviti morfologi sebagai potensi sesuatu rumus morfologi menghasilkan kata terbitan baharu.

Pada tahun 2006, Baayen dalam bukunya yang bertajuk *Corpus Linguistics in Morphology: Morphological Productivity* mencadangkan kaedah linguistik korpus sebagai satu cara untuk mengkaji produktiviti morfologi. Penggunaan korpus pada hemat Baayen amatlah penting kerana sering kali penutur natif tidak sepakat tentang imbuhan yang mana lebih kerap dan produktif digunakan. Menurut Baayen, kategori morfologi yang menunjukkan peningkatan dalam keahliannya merupakan kategori yang produktif.

Penggunaan pendekatan linguistik korpus berbantuan teknologi maklumat dalam mengkaji morfologi produktiviti seperti yang disarankan oleh Baayen (2006) dapat menjadikan sesuatu kajian produktiviti morfologi bersifat autentik. Selain itu, cara ini dapat mengelakkan ketidaktepatan

kategori morfologi yang produktif dengan menggunakan kaedah konvensional, iaitu membuat semakan melalui kamus secara manual untuk mengetahui morfologi yang lebih produktif. Bagi penyusun kamus, imbuhan yang mengapit lebih banyak kata akar seterusnya dikatakan sebagai lebih produktif. Namun, cara ini mungkin tidak berkesan kerana hanya bentuk yang lazim akan tersenarai manakala bentuk yang jarang digunakan tetapi gramatis tidak akan tersenarai kerana kekangan teknologi maklumat.

Tambahan pula menurut Aronoff (1980), kajian produktiviti sesuatu proses morfologi menggunakan kamus adalah tidak tepat kerana bentuk yang tersenarai dalam kamus lazimnya sudah mengalami proses leksikalisasi. Oleh itu, kajian produktiviti morfologi yang menggunakan pendekatan linguistik korpus dapat membantu dalam menjelaskan ketetapan pembentukan kata terbitan atau kesediaan kata akar membentuk kata terbitan baharu.

Dalam kajian ini, dapat lihat bahawa walaupun imbuhan awalan *mem-* apabila ditambah dengan morfem *per + kan* menjadi imbuhan awalan *memper-* dan imbuhan apitan *memper-kan*, tidak semua kata akar dapat bergabung dengan ketiga-tiga imbuhan ini untuk membentuk kata terbitan yang baharu. Contohnya leksikal luas boleh bergabung dengan imbuhan *memper-* memperluas dan imbuhan *memper-kan* memperluaskan tetapi tidak boleh bergabung dengan imbuhan *mem-* memluas. Maka, kajian korpus seperti yang dilakukan di sini amatlah penting untuk mengetahui ciri empirikal kedua-dua bentuk imbuhan awalan *mem-* dan *memper-* dan imbuhan apitan *memper-kan* ini.

KAJIAN LEPAS

Kajian mengenai produktiviti morfologi dengan menggunakan pendekatan linguistik korpus telah banyak dilakukan di Barat. Hal ini kerana bidang kajian ini bukanlah merupakan suatu perkara baharu dalam dunia linguistik Barat. Antara kajian yang pernah dibuat oleh ahli linguistik Barat ialah Ingo Plag, Christiane Dalton-Puffer, Harald Baayen yang menulis tentang “Morphological Productivity Across Speech and Writing” dalam jurnal *English Language and Linguistics* 3.2 (November 1999) yang mengkaji produktiviti morfologi bahasa Inggeris dalam bentuk lisan dan tulisan dengan menggunakan data British National Corpus (BNC).

Selain itu, terdapat juga kajian dalam bahasa lain yang menggunakan ukuran formula produktiviti morfologi ini dalam mengkaji produktiviti

imbuan dan pembentukan kata sesuatu bahasa itu dan didapati berkesan. Contohnya kajian produktiviti morfologi bahasa Belanda (Baayen, 1989, 1992); dan bahasa Mandarin (Sproat dan Shih, 1996; Nishimoto, 2003).

Namun begitu, di Malaysia kajian mengenai produktiviti morfologi dengan menggunakan pendekatan linguistik korpus masih baharu dan masih tidak banyak kajian yang dilakukan. Kebanyakan kajian yang menggunakan pendekatan linguistik korpus, hanya mengkaji aspek imbuhan tertentu, sama ada menggunakan kaedah diakronik mahupun sinkronik tanpa mengaitkannya dengan pengiraan formula produktiviti morfologi ala Baayen. Zaharani Ahmad (2007) misalnya, dalam artikel jurnalnya yang bertajuk “Kekeliruan dalam Penggunaan Bahasa Melayu: Percanggahan antara Rumus Tatabahasa dengan Korpus Bahasa”, mengkaji tentang penggunaan imbuhan *memper-* dan *memper-kan* dengan menggunakan data korpus DBP. Dalam hal ini, didapati imbuhan *memper-kan* yang dianggap lewah dalam bahasa Melayu, ternyata mempunyai produktiviti yang tinggi penggunaannya dalam data korpus DBP. Dalam kajiannya ini, Zaharani melihat nilai produktiviti yang tinggi antara imbuhan *memper-* dan *memper-kan* berpandukan kekerapan kata imbuhan *memper-kan* yang tinggi dalam data korpus DBP.

Kajian “Imbuhan Kompleks Kata Nama *Ke-an* Bahasa Melayu: Analisis Data Korpus” yang dikaji oleh Zaharani Ahmad dan Shakira Khairudin (2010) pula, iaitu terhadap imbuhan kompleks *ke-an* contohnya keseorangan dan kependudukan, menggunakan teori morfologi leksikal dengan memanfaatkan data korpus lima juta UKM-DBP. Dapatan kajian ini menunjukkan rumus pembentukan kata imbuhan *ke-an* terdiri daripada dua strata atau tingkat dan boleh wujud dalam rangkap akhiran selain rangkap awalan dalam bahasa Melayu.

Selain itu, terdapat juga kajian yang dilakukan oleh Zaharani Ahmad, Shakira Khairudin dan Nor Hashimah Jalaludin (2012) yang mengkaji tentang “Perilaku Morfologi Awalan *Ber-* dalam Bahasa Melayu Klasik dan Bahasa Melayu Moden: Satu Kajian Perbandingan”. Dalam kajian ini, perbandingan imbuhan *ber-* dengan memanfaatkan data korpus klasik DBP dan korpus moden UKM-DBP yang bertujuan untuk melihat gejala pengekalan dan inovasi yang berlaku dalam sejarah perkembangan bahasa Melayu menggunakan teori morfologi generatif. Dapatan kajian menunjukkan imbuhan *ber-* mengalami pengekalan dan juga inovasi sama ada mengalami peluasan, pengecilan, pengambilalihan mahupun penghilangan.

Kajian lain, iaitu kajian Syed Zainal Ariff Syed Jamaluddin dan Noor Azaliza Yahya (2013), mengkaji tentang “Perbandingan Darjah Transitiviti Awal *Ber-* dan *Per-* dalam Bahasa Melayu” dengan menjadikan empat teks terpilih abad ke-17 dan 18 sebagai sumber kajian. Dapatan kajian ini menunjukkan imbuhan *ber-* mempunyai darjah transitiviti lebih tinggi daripada imbuhan *per-* dengan menggunakan teori Hopper dan Thompson (1980).

Berdasarkan kajian terdahulu ternyata tidak ada lagi kajian imbuhan dari sudut produktiviti morfologi imbuhan menggunakan pendekatan Baayen (1992, 1996). Setakat ini, hanya Imran Ho (2011) mengkaji imbuhan dari segi produktiviti morfologi. Dalam kajian beliau yang bertajuk “Produktiviti Imbuhan Apitan *Ke ... An* dan *Pen ... An*: Pendekatan Linguistik Korpus”, beliau mengkaji imbuhan apitan *ke-an* dan *peN-an* dengan menggunakan data korpus akhbar *Berita Harian* yang dijana daripada data korpus lima juta UKM-DBP. Dengan menggunakan ukuran Baayen, dapat dilihat bahawa imbuhan apitan *ke-...-an* lebih tinggi nilai produktiviti berbanding *peN-...-an*, namun imbuhan *peN-...-an* didapati mempunyai bilangan jenis (*type*) yang lebih tinggi daripada *ke-...-an* dengan menggunakan pendekatan sikronik.

Kajian yang dilakukan oleh Imran Ho ini hanya mencakupi imbuhan apitan *ke-an* dan *peN-an* dan dikaji menggunakan pendekatan sinkronik. Sehubungan dengan itu, kajian yang dilakukan ini merupakan kesinambungan daripada kajian yang telah dilakukan oleh Imran Ho Abdullah. Kajian ini akan meneliti perbandingan produktiviti imbuhan awalan *mem-* dan *memper-* dengan awalan apitan *memper-kan* dengan menggunakan data korpus Akhbar Sukan Dalam Talian dan data korpus UKM-DBP. Dengan adanya kajian ini sedikit sebanyak akan dapat mengisi kelompangan, dan membantu melengkapkan sebahagian lagi kajian imbuhan yang berkaitan dengan produktiviti morfologi dengan menggunakan ukuran Baayen (1992, 1996) dalam bahasa Melayu.

METODOLOGI KAJIAN

Dalam kajian ini, dua bentuk korpus digunakan, iaitu korpus sukan Akhbar Dalam Talian dan korpus UKM-DBP. Data korpus Sukan Akhbar Dalam Talian merupakan data korpus yang dikumpulkan oleh pelajar sarjana UKM yang mengambil subjek Linguistik Korpus sesi 2014/15. Dalam korpus ini terdapat tiga jenis akhbar dalam talian, iaitu akhbar *Utusan Online*, *Kosmo Online* dan *Metro Online*. Saiz korpus ini berjumlah

2 473 971 patah perkataan dengan 8027 artikel sukan yang dihimpunkan daripada teks *Utusan Online*, *Kosmo Online* dan *Metro Online* antara 1 Januari 2001 – 30 September 2014.

Data korpus UKM-DBP merupakan subkorpus Korpus 5 Juta UKM-DBP yang dibina untuk tujuan nahu praktis, iaitu nahu berdasarkan penggunaan. Tujuan kajian ini menggunakan dua bentuk korpus ini yang berbeza saiz data korpus adalah untuk melihat dan membuat perbandingan sama ada saiz korpus mempengaruhi produktiviti morfologi imbuhan yang dikaji. Statistik asas korpus sukan akhbar dalam talian dan korpus UKM-DBP yang digunakan adalah seperti dalam Jadual 1 dan 2 berikut.

Jadual 1 Statistik asas korpus akhbar sukan dalam talian.

Jumlah perkataan (token)	2 473 971
Jumlah jenis perkataan berbeza	35 949
Nisbah jenis - token	1.45

Jadual 2 Statistik asas korpus UKM-DBP.

Jumlah perkataan (token)	5 103 320
Jumlah jenis perkataan berbeza	102 602
Nisbah jenis - token	2.01

Jika dilihat daripada statistik asas kedua-dua data korpus ini, dalam Jadual 1 dan Jadual 2, didapati bahawa saiz korpus antara korpus Akhbar Sukan Dalam Talian dan data korpus UKM-DBP hanya berbeza nisbah *type-token* sebanyak 0.56 sahaja. Jika dibandingkan kedua-dua data korpus ini dengan saiz korpus bahasa lain yang digunakan dalam kajian morfologi ternyata masih relatif kecil. Namun begitu, disebabkan kajian ini merupakan kajian awal, kedua-dua data korpus yang kecil ini adalah sesuai.

Dengan menggunakan perisian WordSmith, satu senarai daftar kata awal (SK1) telah dijana daripada kedua-dua korpus ini. Contoh senarai daftar kata awal ini dapat dilihat dalam Rajah 1 dan 2.

Seterusnya, perkataan berimbuhan awalan *mem-* dan *memper-* dan imbuhan apitan *memper-kan* yang terdapat dalam korpus Akhbar Sukan Dalam Talian dan korpus UKM-DBP telah diasingkan daripada kata imbuhan dan kata akar serta angka yang tidak berkaitan dengan kajian ini. Hasilnya, senarai daftar kata yang telah dibersihkan seperti dalam contoh Rajah 3 dan 4.

Word	Freq	% Texts	% Lemmas	Set
PERLU	4,100	0.13	2,037	51.69
BERJAYA	4,067	0.13	2,004	50.85
BOLEH	3,981	0.12	2,032	51.56
HAJUN	3,981	0.12	2,070	52.52
PUSINGAN	3,938	0.12	1,403	37.88
NEGERI	3,935	0.12	1,421	36.06
DAPAT	3,896	0.12	2,054	52.12
MAHU	3,853	0.12	1,925	48.85
A	3,846	0.12	135	3.43
O	3,840	0.12	137	3.48
AKSI	3,800	0.12	1,747	44.33
SAINGAN	3,800	0.12	1,849	46.92
AHMAD	3,795	0.12	1,784	45.27
PERSATUAN	3,739	0.12	1,810	45.93
MUSIM	3,730	0.12	1,420	36.03
b	3,711	0.12	135	3.43
DEPAN	3,688	0.11	1,971	50.01
Z	3,672	0.11	240	6.09
Q	3,628	0.11	255	6.47
Z	3,612	0.11	133	3.37
PERLUMBAHAN	3,603	0.11	841	21.34
Q	3,580	0.11	135	3.45
MAMPU	3,540	0.11	1,964	49.84
SEPERTI	3,440	0.11	1,995	50.62
PULA	3,436	0.11	1,959	49.71
MEMBERI	3,405	0.11	1,352	39.53
BAIK	3,319	0.10	1,843	46.76
LATHAN	3,306	0.10	1,416	35.93
MEHERUSI	3,303	0.10	1,579	40.04
TERBUKA	3,280	0.10	1,273	32.30
KEPUTUSAN	3,243	0.10	1,720	43.64
SEMENTARA	3,239	0.10	2,132	54.10

Rajah 1 Senarai kata awal (SK1) yang dijana daripada korpus Akhbar Sukan Dalam Talian UKM-DBP.

Word	Freq	% Texts	% Lemmas	Set
POLIS	2,797	0.05	5	100.00
BUKU	2,787	0.05	5	100.00
DIBUAT	2,780	0.05	5	100.00
MUDAH	2,779	0.05	5	100.00
SALAH	2,777	0.05	5	100.00
HAMA	2,775	0.05	5	100.00
MEMBARTU	2,769	0.05	5	100.00
TEKNOLOGI	2,767	0.05	5	100.00
AGAMA	2,766	0.05	5	100.00
DEMIKAN	2,758	0.05	5	100.00
ASING	2,752	0.05	5	100.00
PERNIAGAAN	2,740	0.05	5	100.00
MEMBAWA	2,731	0.05	5	100.00
AHWAL	2,723	0.05	5	100.00
WDS	2,717	0.05	4	80.00
BELUM	2,714	0.05	5	100.00
DATANG	2,696	0.05	5	100.00
SEMAKIN	2,686	0.05	5	100.00
TERPAKSA	2,671	0.05	5	100.00
MENARIK	2,652	0.05	5	100.00
PENGUNIA	2,650	0.05	5	100.00
KEMUDIAN	2,642	0.05	5	100.00
MOHD	2,629	0.05	5	100.00
JENIS	2,622	0.05	5	100.00
KEPUTUSAN	2,583	0.05	5	100.00
PROSES	2,579	0.05	5	100.00
BAHAN	2,575	0.05	5	100.00
SUKANI	2,569	0.05	5	100.00
KALANGAN	2,560	0.05	5	100.00
MENYEBABKAN	2,559	0.05	5	100.00
TIBA	2,533	0.05	5	100.00
SUATU	2,528	0.05	5	100.00

Rajah 2 Senarai kata awal (SK1) yang dijana daripada korpus UKM-DBP.

Seterusnya, produktiviti imbuhan *mem-*, *memper-* dan *memper-*kan dianalisis secara kuantitatif. Pendekatan analisis untuk produktiviti yang

The figure consists of three side-by-side screenshots of Microsoft Excel spreadsheets, each displaying a list of words and their frequency statistics. The first spreadsheet is titled 'Imbuhan MEM-HY-BAHARU.xls', the second 'Imbuhan MEMPER-HY-BAHARU.xls', and the third 'Imbuhan MEMPERKAN-HY-BAHARU.xls'. Each spreadsheet has columns for 'N', 'Word', 'Freq', '%', 'Texts', and '% Lembar'.

N	Word	Freq	%	Texts	% Lembar
1	MEMER	3,406	0.11	1,952	49.53
2	MEMERUT	1,900	0.06	1,269	32.17
3	MEMERUTU	1,767	0.06	1,208	30.65
4	MEMERUTU	1,386	0.04	996	25.27
5	MEMERURU	722	0.02	580	14.72
6	MEMERUTAHU	709	0.02	592	15.02
7	MEMERUKA	697	0.02	600	15.22
8	MEMERUTUK	299	0.01	249	6.29
9	MEMERIBING	279	0.01	194	4.92
10	MEMERELASAH	267	0.01	241	6.12
11	MEMERIBAWI	265	0.01	229	5.81
12	MEMERAMER	212	0.01	182	4.62
13	MEMERIBALAS	192	0.01	165	4.19
14	MEMERIBAYAR	132	0.00	99	2.51
15	MEMERIBELU	130	0.00	104	2.64
16	MEMERIBIWA	128	0.00	104	2.64
17	MEMERIBAVAR	51	0.00	46	1.17
18	MEMERIBARA	51	0.00	48	1.22
19	MEMERIBANTAH	46	0.00	40	1.01
20	MEMERIBACA	44	0.00	42	1.07
21	MEMERIBALING	43	0.00	29	0.74
22	MEMERIPROMOSI	34	0.00	28	0.71
23	MEMERIBELINGGU	33	0.00	31	0.79
24	MEMERIBOLOT	28	0.00	28	0.71
25	MEMERIBUANG	27	0.00	27	0.69

N	Word	Freq	%	Texts	% Lembar
1	MEMPEROLEH	836	0.03	594	15.07
2	MEMPERSIJA	18	0.00	18	0.46
3	MEMPERTAHAN	15	0.00	15	0.38
4	MEMPERKUAT	5	0.00	5	0.13
5	MEMPERDAYA	4	0.00	4	0.10
6	MEMPERKEMAS	4	0.00	4	0.10
7	MEMPERTAKAI	4	0.00	3	0.08
8	MEMPERTINGKAT	3	0.00	3	0.08
9	MEMPERVENJAL	2	0.00	2	0.05
10	MEMPERLUHAT	2	0.00	2	0.05
11	MEMPERSIAP	2	0.00	2	0.05
12	MEMPERBESAR	1	0.00	1	0.03
13	MEMPERCEPAT	1	0.00	1	0.03
14	MEMPERKETAT	1	0.00	1	0.03
15	MEMPERKUKUH	1	0.00	1	0.03
16	MEMPERLUAS	1	0.00	1	0.03
17	MEMPERSOAL	1	0.00	1	0.03
18	MEMPERTANDING	1	0.00	1	0.03
19	MEMPERTIMBAING	1	0.00	1	0.03

N	Word	Freq	%	Texts	% Lembar
1	MEMPERTAHANKAN	895	0.03	688	17.46
2	MEMPERKEMALKAN	135	0.00	119	3.02
3	MEMPERKUKATKAN	93	0.00	89	2.26
4	MEMPERTARUHKAN	90	0.00	87	2.21
5	MEMPERSIAPKAN	84	0.00	76	1.93
6	MEMPERTIMBAINGKAN	81	0.00	71	1.80
7	MEMPERTINGKATKAN	63	0.00	60	1.52
8	MEMPERLUHATKAN	62	0.00	59	1.50
9	MEMPERTANDINGKAN	61	0.00	57	1.45
10	MEMPERKEMASKAN	54	0.00	52	1.32
11	MEMPERTAKAIKAN	46	0.00	40	1.01
12	MEMPERKUKUHKAN	45	0.00	42	1.07
13	MEMPERSOALKAN	39	0.00	38	0.96
14	MEMPERUNTUKKAN	38	0.00	37	0.94
15	MEMPERSEIMBANGKAN	30	0.00	29	0.74
16	MEMPERDAYAKAN	24	0.00	24	0.61
17	MEMPERTAJUHKAN	23	0.00	21	0.53
18	MEMPERBETULKAN	14	0.00	14	0.36
19	MEMPERCEPATKAN	14	0.00	14	0.36
20	MEMPERKASAKAN	14	0.00	13	0.33
21	MEMPERTAKATKAN	13	0.00	13	0.33
22	MEMPERLAHANKAN	13	0.00	13	0.33
23	MEMPERBAWUHKAN	12	0.00	11	0.28
24	MEMPERHEBATKAN	12	0.00	12	0.30
25	MEMPERKECOLKAN	12	0.00	12	0.30

Rajah 3 Senarai kata bersih imbuhan mem-, memper- dan memper-kan daripada korpus Akhbar Sukan Dalam Talian.

diterapkan dalam kajian ini ialah pendekatan Baayen (1989, 1992, 1994) yang mencadangkan kaedah pengukuran produktiviti sesuatu proses pembentukan kata berdasarkan data korpus. Formula Baayen adalah seperti yang berikut:

$$P = \frac{n_c}{N}$$

Berdasarkan formula ini, indeks produktiviti (p) bagi sesuatu imbuhan C ialah nisbah antara bilangan *hapax legomena* (nc) (perkataan yang berlaku hanya sekali dalam korpus) dengan jumlah keberlakuan kata terbitan (*sum of word tokens*) dengan imbuhan itu dalam korpus (N). Bilangan

N	Word	Freq.	%	Texts	% Lemmas	Set
1	MEMBER	6,220	0.12	5	100.00	
2	MEMBUAT	5,148	0.10	5	100.00	
3	MEMBANTU	2,788	0.05	5	100.00	
4	MEMBAWA	2,731	0.05	5	100.00	
5	MEMBELI	1,717	0.03	5	100.00	
6	MEMBERTAHU	1,483	0.03	5	100.00	
7	MEMBINA	1,138	0.02	5	100.00	
8	MEMBACA	1,089	0.02	5	100.00	
9	MEMBUKA	967	0.02	5	100.00	
10	MEMBEUTUK	816	0.02	5	100.00	
11	MEMBAYAR	742	0.01	5	100.00	
12	MEMBUNUH	410	0.01	5	100.00	
13	MEMBANGUN	408	0.01	5	100.00	
14	MEMBAJANG	344	0.01	5	100.00	
15	MEMBELA	288	0.01	5	100.00	
16	MEMBANTAH	241	0.01	5	100.00	
17	MEMBESAR	225	0.01	4	80.00	
18	MEMBALAS	142	0.01	5	100.00	
19	MEMBAYAR	106	0.01	5	100.00	
20	MEMBENDUNG	104	0.01	5	100.00	
21	MEMPROSES	122	0.01	5	100.00	
22	MEMBURU	118	0.01	5	100.00	
23	MEMBANTERAS	117	0.01	5	100.00	
24	MEMBINGUNG	107	0.01	5	100.00	
25	MEMBASUH	89	0.01	4	80.00	

N	Word	Freq.	%	Texts	% Lemmas	Set
1	MEMPEROLEH	226	0.01	5	100.00	
2	MEMPERKUKUH	57	0.01	5	100.00	
3	MEMPERKAYA	42	0.01	2	40.00	
4	MEMPERLUAS	28	0.01	4	80.00	
5	MEMPERCEPAT	25	0.01	5	100.00	
6	MEMPERTINGKAT	24	0.01	4	80.00	
7	MEMPERBAK	22	0.01	3	60.00	
8	MEMPERTAHAN	18	0.01	5	100.00	
9	MEMPERKEJAL	17	0.01	4	80.00	
10	MEMPERKEJOL	10	0.01	4	80.00	
11	MEMPERKUAT	10	0.01	4	80.00	
12	MEMPERLUHAT	10	0.01	4	80.00	
13	MEMPERSOAL	10	0.01	5	100.00	
14	MEMPERJUANG	9	0.01	3	60.00	
15	MEMPERMUDAH	9	0.01	4	80.00	
16	MEMPERSENDA	8	0.01	4	80.00	
17	MEMPERBANYAK	7	0.01	2	40.00	
18	MEMPERBESAR	7	0.01	3	60.00	
19	MEMPERDAYA	7	0.01	3	60.00	
20	MEMPERLELAS	7	0.01	2	40.00	
21	MEMPERJAIN	7	0.01	3	60.00	
22	MEMPERTAJAM	7	0.01	3	60.00	
23	MEMPERKEMAS	6	0.01	4	80.00	
24	MEMPERKEMBANG	6	0.01	3	60.00	
25	MEMPERBETUL	5	0.01	2	40.00	

N	Word	Freq.	%	Texts	% Lemmas	Set
1	MEMPERTAHKAN	636	0.01	5	100.00	
2	MEMPERTAHKAN	612	0.01	5	100.00	
3	MEMPERTAJUKAN	521	0.01	5	100.00	
4	MEMPERTINGKATKAN	232	0.01	5	100.00	
5	MEMPERTAKAN	195	0.01	5	100.00	
6	MEMPERTUJUKAN	183	0.01	5	100.00	
7	MEMPERCEPATKAN	168	0.01	5	100.00	
8	MEMPERTUKKAN	162	0.01	5	100.00	
9	MEMPERSOALKAN	154	0.01	5	100.00	
10	MEMPERTUJUKAN	152	0.01	4	80.00	
11	MEMPERKUKHKAN	119	0.01	5	100.00	
12	MEMPERTAKKAN	89	0.01	5	100.00	
13	MEMPERTAMBAHKAN	87	0.01	5	100.00	
14	MEMPERBANYAKKAN	44	0.01	5	100.00	
15	MEMPERSEBAHKAN	40	0.01	4	80.00	
16	MEMPERTUKKAN	38	0.01	5	100.00	
17	MEMPERKEMBANGKAN	37	0.01	5	100.00	
18	MEMPERSIAPKAN	36	0.01	5	100.00	
19	MEMPERTAKKAN	34	0.01	4	80.00	
20	MEMPERTUKKAN	29	0.01	5	100.00	
21	MEMPERTUKKAN	29	0.01	5	100.00	
22	MEMPERHAYAKAN	28	0.01	4	80.00	
23	MEMPERLUASKAN	27	0.01	5	100.00	
24	MEMPERKEJOLKAN	27	0.01	5	100.00	
25	MEMPERTUJUKAN	26	0.01	3	60.00	

Rajah 4 Senarai kata imbuhan *mem-*, *memper-* dan *memper-*kan bersih daripada korpus UKM-DBP.

hapax legomena adalah sama dengan bilangan jenis kata terbitan berbeza berimbuhan *C* (*word types*) yang muncul sekali sahaja dalam korpus.

Dalam kajian korpus, istilah jenis (*type*) merujuk jenis kata berbeza dan token merujuk kekerapan keberlakuan. Contohnya, jika perkataan *saya* ditemui 20 kali dalam korpus, maka kekerapan token (keberlakuan) ialah 20 dan kekerapan *type* (jenis) ialah satu. Dalam hal imbuhan, lazimnya, sekiranya korpus mempunyai keberlakuan {memfitnah, membara, memfitnah, membara, membara}, kekerapan token ialah lima dan kekerapan *type* ialah dua. Formula Baayen ini dikenali sebagai satu ukuran produktiviti berasaskan *hapax* kerana kebergantungan pengukuran kepada bilangan *hapax* dalam sesuatu korpus serta kekerapan token dalam menilai kebarangkalian munculnya kata terbitan baharu (neologisma). Pada pendapat Nishimoto (2003) formula Baayen ini mengguna pakai kaedah

Imbuan MEM data akhbar dalam talian			Imbuan MEMPER akhbar dalam talian			Imbuan MEMPERTAHAN akhbar dalam talian		
N	Word	Freq.	N	Word	Freq.	N	Word	Freq.
1	MEMBERI	3406	1	MEMPEROLEH	837	1	MEMPERTAHANKAN	897
2	MEMBUAT	1901	2	MEMPERSA	18	2	MEMPERKENALKAN	136
3	MEMBAWA	1767	3	MEMPERTAHAN	16	3	MEMPERKUATKAN	93
4	MEMBANTU	1387	4	MEMPERKUAT	6	4	MEMPERTARUHKAN	90
5	MEMBURU	722	5	MEMPERDAYA	4	5	MEMPERSIAPKAN	84
6	MEMBERITAHU	709	6	MEMPERKEMAS	4	6	MEMPERTIMBANGKAN	81
7	MEMBUKA	697	7	MEMPERTIKAI	4	7	MEMPERTINGKATKAN	64
8	MEMBEUTUK	298	8	MEMPERTINGKAT	3	8	MEMPERLUHATKAN	62
9	MEMBIMBING	279	9	MEMPERKEJAL	2	9	MEMPERTANDINGKAN	61
10	MEMBELASAH	268	10	MEMPERLIHAT	2	10	MEMPERKEMASKAN	54
11	MEMBEJAM	265	11	MEMPERSIAP	2	11	MEMPERTIKAIKAN	46
12	MEMPAMER	212	12	MEMPERBESAR	1	12	MEMPERKUKUHKAN	45
61	MEMPORAK	2	13	MEMPERCEPAT	1	28	MEMPERKUJUKKAN	9
62	MEMBAHAGI	1	14	MEMPERKETAT	1	29	MEMPERGUNAKAN	8
63	MEMBAHANG	1	15	MEMPERKLUJAH	1	30	MEMPERTEMUKAN	7
64	MEMBAJA	1	16	MEMPERLUAS	1	31	MEMPERLUASKAN	6
65	MEMBANGKIT	1	17	MEMPERSOAL	1	32	MEMPERBESARKAN	5
66	MEMBAYANG	1	18	MEMPERTANDING	1	33	MEMPERINJAKAN	5
67	MEMBEKAL	1	19	MEMPERTIMBANG	1	34	MEMPERSIJAKAN	5
68	MEMBELAH	1			905	35	MEMPERJUDIKAN	4
69	MEMBELANJA	1	V		19	36	MEMPERELASKAN	3
70	MEMBENTANG	1	N		905	37	MEMPERSEALAHKAN	3
71	MEMBERANGSANG	1	n (hapax)		8	38	MEMPERUNTUKAN	3
72	MEMBIAS	1	P		0.009	39	MEMPERKETATKAN	2
73	MEMBIASA	1				40	MEMPERBAGAIKAN	1
74	MEMBIKINI	1				41	MEMPERGIATKAN	1
75	MEMBOCENG	1				42	MEMPERINDAHKAN	1
76	MEMBREK	1				43	MEMPERKEMBANGKAN	1
77	MEMBURUK	1				44	MEMPERKENAIKAN	1
78	MEMPERAGA	1				45	MEMPERMAINKAN	1
79	MEMPERKASA	1				46	MEMPERSENDIKAN	1
80	MEMPROSES	1				47	MEMPERTURUNKAN	1
81	MEMPROTES	1						2060
		13109						
						V		47
V		81				N		2060
N		13109				n (hapax)		8
n (hapax)		20				P		0.004
P		0.002						

hapax (n) merujuk kekerapan kata yang mempunyai kekerapan hanya sekali sahaja
Formula $P=n/N$

Rajah 5 Hasil pengiraan imbuan *mem-*, *memper-* dan *memper-*kan berdasarkan Formula Baayen (1989, 1992) korpus akhbar sukan dalam talian dalam format excel.

anggaran Good-Turing yang memberikan anggaran secara matematik tentang kebarangkalian menemui perkataan baharu dalam sesuatu korpus berdasarkan *hapax* yang wujud dalam korpus itu.

Imbunan MEM data korpus UKM-DBP		
N	Word	Freq.
1	MEMBERI	6231
2	MEMBUAT	5152
3	MEMBANTU	2768
155	MEMBAHAYA	1
156	MEMBAHAYA	1
157	MEMBAHU	1
158	MEMBALUN	1
159	MEMBANIR	1
160	MEMBAPTIS	1
161	MEMBAUR	1
162	MEMBAYANG	1
163	MEMBEDAK	1
164	MEMBEGAL	1
165	MEMBEKAS	1
166	MEMBELOT	1
167	MEMBELUDAK	1
168	MEMBETING	1
169	MEMBIAS	1
170	MEMBIAYA	1
171	MEMBIDAN	1
172	MEMBUI	1
173	MEMBIBANG	1
174	MEMBINASA	1
175	MEMBOLOS	1
176	MEMBONDONG	1
177	MEMBONJOL	1
178	MEMBORONG	1
179	MEMBUAH	1
180	MEMBUAS	1
181	MEMBULUNG	1
182	MEMBUKUT	1
183	MEMBUNCAH	1
184	MEMBUNCI	1
185	MEMBUNGKAS	1
186	MEMBUSAR	1
187	MEMBUSUNG	1
188	MEMBUSUT	1
189	MEMPARODI	1
190	MEMPELOPOR	1
191	MEMPELLUHU	1
192	MEMPERDULI	1
193	MEMPIAWAI	1
194	MEMPLAGIAT	1
195	MEMPRAKTIK	1
196	MEMPREBIU	1
197	MEMPROVOKASI	1
198	MEMPUITIS	1
		29579
V		198
N		29579
n (hapax)		44
P		0.001

Imbunan MEMPER data korpus UKM-DBP		
N	Word	Freq.
1	MEMPEROLEH	226
2	MEMPERKULUH	57
3	MEMPERKAYA	42
4	MEMPERLUAS	28
5	MEMPERCEPAT	25
42	MEMPERBUAT	1
43	MEMPERBUDAK	1
44	MEMPERDENGAR	1
45	MEMPERGUNA	1
46	MEMPERINGAT	1
47	MEMPERKETAT	1
48	MEMPERLICIN	1
49	MEMPERMODAL	1
50	MEMPERMUDA	1
51	MEMPERLOK	1
52	MEMPERLUANG	1
53	MEMPERALAH	1
54	MEMPERSIINGKAT	1
55	MEMPERSUJUH	1
56	MEMPERTARUH	1
57	MEMPERTEGAS	1
58	MEMPERTEMU	1
59	MEMPERTINGGI	1
60	MEMPERTUDUH	1
61	MEMPERTURUN	1
		657
V		61
N		657
n (hapax)		20
P		0.03

Imbunan MEMPERKAN data korpus UKM-DBP		
N	Word	Freq.
1	MEMPERLHATKAN	637
2	MEMPERTAHANKAN	614
3	MEMPERKENALKAN	521
4	MEMPERTINGKATKAN	234
5	MEMPERKATAKAN	195
6	MEMPERLUANGKAN	185
79	MEMPERBAGIKAN	1
80	MEMPERBAIKKAN	1
81	MEMPERBAINDINGKAN	1
82	MEMPERBINANGKAN	1
83	MEMPERBOLEHKAN	1
84	MEMPERBUDIAKAN	1
85	MEMPERCAKAPKAN	1
86	MEMPERCAITKAN	1
87	MEMPERCATIKAN	1
88	MEMPERLOKKAN	1
88	MEMPERDAHKAN	1
88	MEMPERLOHOKAN	1
90	MEMPERJODOHKAN	1
91	MEMPERLANGGAKAN	1
92	MEMPERLABAGAKAN	1
93	MEMPERLECEHKAN	1
94	MEMPERLOINCAN	1
95	MEMPERMARTABAKAN	1
96	MEMPERMULAKAN	1
97	MEMPERNAGAKAN	1
98	MEMPERLOLOKKAN	1
99	MEMPERPADATKAN	1
100	MEMPERSAIDARKAN	1
101	MEMPERSEDIAKAN	1
102	MEMPERSEMPITKAN	1
103	MEMPERTEGANGKAN	1
104	MEMPERTUMBUNGAN	1
105	MEMPERTUJUKAN	1
106	MEMPERTUHANKAN	1
107	MEMPERTUNLUKKAN	1
108	MEMPERTURUNKAN	1
109	MEMPERTURUTKAN	1
110	MEMPERTUTURKAN	1
111	MEMPERUNDANGKAN	1
		4185
V		111
N		4185
n (hapax)		33
P		0.008

Rajah 6 Hasil Pengiraan imbunan mem-, memper- dan memper-kan berdasarkan Formula Baayen (1989, 1992) korpus UKM-DBP dalam format Excel.

Indeks produktiviti (p) menyatakan kebarangkalian satu perkataan baharu dengan imbunan itu akan muncul dalam korpus selepas N token imbunan berkenaan telah disampel. Oleh itu, nilai p adalah berasaskan

token (dan bukan *type*) dan amat bergantung pada kekerapan keberlakuan *word-token* dalam korpus. Jumlah jenis kata (*sum of word types*) dengan imbuhan berkaitan dalam korpus (V) tidak diambil kira secara langsung dalam ukuran darjah produktiviti imbuhan tersebut.

Dalam kajian ini, imbuhan *mem-*, *memper-* dan *memper-kan* korpus Akhbar Sukan Dalam Talian dan korpus UKM-DBP yang dibersihkan dalam helaian eletronik perisian WordSmith telah disimpan dalam format Excel untuk memudahkan pengiraan produktiviti morfologi mengikut formula yang digunakan oleh Baayen (1989, 1992). Contoh senarai daftar kata imbuhan *mem-*, *memper-* dan *memper-kan* dan hasil pengiraannya dalam format Excel adalah seperti dalam Rajah 5 dan 6.

DAPATAN DAN PERBINCANGAN

Dapatan analisis produktiviti berdasarkan ukuran *hapax* dalam korpus Akhbar Sukan Dalam Talian dan korpus UKM-DBP dapat dilihat dalam Jadual 3 dan Jadual 4. Berdasarkan semua keberlakuan kata terbitan dengan imbuhan *mem-*, *memper-* dan *memper-kan*, V ialah bilangan jenis kata terbitan berbeza (*types*), N adalah jumlah keberlakuan kata terbitan dengan apitan tersebut (token) dan n1 ialah bilangan *hapax* untuk apitan tersebut. P ialah indeks produktiviti berdasarkan formula Baayen.

Dalam Jadual 3, dapat dilihat bahawa berdasarkan analisis yang dilakukan, didapati imbuhan *memper-* dalam korpus Akhbar Sukan Dalam Talian mempunyai indeks produktiviti yang lebih tinggi, iaitu 0.009 berbanding imbuhan *mem-* (0.002) dan *memper-kan* (0.004). Namun demikian, perbezaan nilai P antara ketiga-tiga imbuhan ini tidaklah begitu besar, iaitu berbeza antara 0.005 (*memper-kan*) dan 0.007 (*mem-*). Berdasarkan Jadual 3, sungguhpun lebih banyak jenis kata terbitan berbeza (V) dibentuk oleh imbuhan *mem-* (V = 81) berbanding dengan imbuhan *memper-* (V = 19), hakikatnya, imbuhan *mem-* bukan imbuhan yang paling produktif.

Bagi Jadual 4 pula, dapat dilihat bahawa berdasarkan analisis yang dilakukan, didapati imbuhan *memper-* dalam korpus UKM-DBP mempunyai indeks produktiviti yang lebih tinggi, iaitu 0.030 berbanding *mem-* (0.001) dan *memper-kan* (0.008). Perbezaan nilai P antara ketiga-tiga imbuhan ini agak begitu besar, iaitu berbeza antara 0.022 (*memper-kan*) dan 0.029 (*mem-*). Berdasarkan Jadual 4, sungguhpun lebih banyak jenis kata terbitan berbeza (V) dibentuk oleh imbuhan *mem-* (V = 198) berbanding dengan imbuhan *memper-* (V = 61), hakikatnya, imbuhan

Jadual 3 Hasil pengiraan imbuhan *mem-*, *memper-* dan *memper-kan* data akhbar sukan dalam talian.

Imbuhan	V	N	n1 (<i>hapax</i>)	P
mem-	81	13109	20	0.002
memper-	19	905	8	0.009
memper...kan	47	2060	8	0.004

Jadual 4 Hasil pengiraan imbuhan *mem-*, *memper-* dan *memper-kan* data UKM-DBP.

Imbuhan	V	N	n1 (<i>hapax</i>)	P
mem-	81	13109	20	0.001
memper-	19	905	8	0.030
memper...kan	47	2060	8	0.008

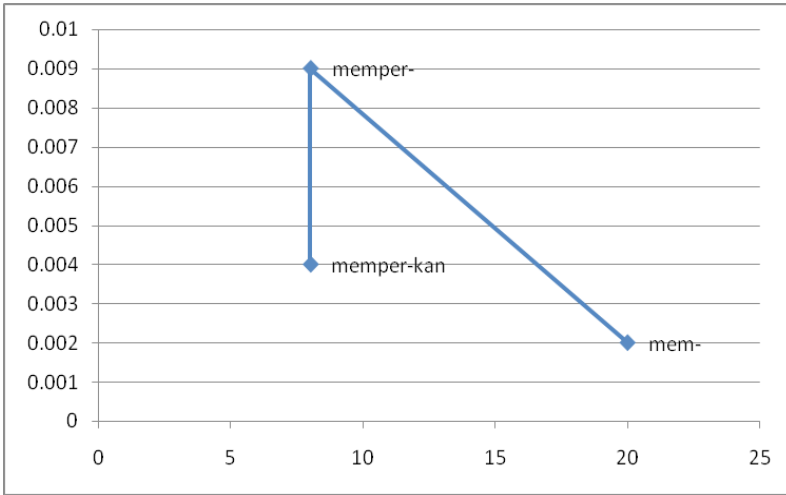
mem- bukanlah imbuhan yang paling produktif. Jika ditakrifkan dan diukur “produktiviti” dari segi bilangan kata akar berbeza (V) yang boleh mengambil imbuhan apitan, maka kesimpulan yang dibuat ialah *mem-* lebih “produktif” berbanding imbuhan *memper-*. Bilangan *type* yang tinggi bagi imbuhan *mem-* juga tercermin dalam kekerapan keberlakuan (frekuensi token) (N) untuk imbuhan tersebut yang jauh mengatasi imbuhan *memper-* dan *memper-kan*.

Sebaliknya, jika kita mengukur produktiviti berdasarkan formula Baayen, urutan produktiviti tiga imbuhan ini seolah-olah terbalik. Imbuhan *memper-* mempunyai nilai P tertinggi, berbanding *memper-kan* dan *mem-*. Tafsiran produktiviti Baayen sebenarnya merujuk kebarangkalian satu kata terbitan baharu atau neologisme dibentuk oleh imbuhan tersebut. Pengiraan produktiviti Baayen yang berdasarkan *hapax* (*hapax-based measure*), membuat andaian bahawa kekerapan token yang tinggi boleh dikaitkan

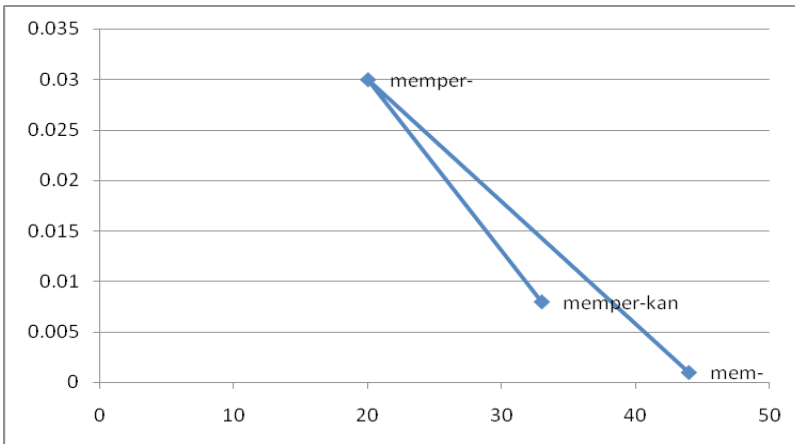
dengan darjah leksikalisasi yang tinggi. Darjah leksikalisasi merujuk sama ada sesuatu bentuk itu tersimpan dalam leksikon sebagai satu unit penuh, dan semakin tinggi darjah leksikalisasi sesuatu perkataan semakin rendah darjah produktiviti (Baayen, 1989, 1992).

Dapatan kajian ini setelah dianalisis imbuhan *mem-*, *memper-* dan *memper-kan* seperti dalam Jadual 3 dan Jadual 4, mendapati bahawa imbuhan *memper-* mempunyai produktiviti yang paling tinggi dalam kedua-dua korpus berkenaan. Bagi korpus Akhbar Sukan Dalam Talian, imbuhan *memper-* mencatatkan nilai $P=0.009$, manakala dalam korpus UKM-DBP, imbuhan *memper-* mencatatkan nilai $P=0.030$. Imbuhan yang mempunyai produktiviti yang paling rendah antara ketiga-tiga imbuhan ini dalam kedua-dua korpus ini ialah imbuhan *mem-*. Imbuhan *mem-* dalam korpus Akhbar Sukan Dalam Talian mencatatkan nilai $P=0.002$, manakala dalam korpus UKM-DBP mencatatkan nilai $P=0.001$. Kesimpulannya, pola produktiviti imbuhan yang paling tinggi, iaitu *memper-* dan imbuhan paling rendah, iaitu *mem-* dalam korpus Akhbar Sukan Dalam Talian dan korpus UKM-DBP adalah sama walaupun nilai P kedua-dua korpus berbeza.

Selain itu, berdasarkan pengiraan produktiviti ala Baayen ini, didapati bahawa penambahan morfem *per-* dan *-kan* daripada imbuhan awalan *mem-* kepada imbuhan *memper-* dan *memper-kan* (imbuhan apitan) ternyata tidak mempengaruhi sepenuhnya produktiviti morfologi sesuatu imbuhan itu. Faktor utama yang mempengaruhi produktiviti morfologi sesuatu imbuhan itu banyak dipengaruhi oleh *hapax* yang menjadi asas kepada konsep produktiviti ala Baayen. Dalam kedua-dua jadual tersebut dapat dilihat bahawa imbuhan *memper-* mencatatkan produktiviti tertinggi, iaitu 0.009 dan 0.030 walaupun secara intuitivanya sepatutnya imbuhan *memper-kan* yang lebih produktif kerana terdapat penambahan morfem *per-* dan *-kan*. Namun, dalam jadual di atas, dapat dilihat nilai P bagi imbuhan *memper-kan* dalam kedua-dua korpus merupakan kedua terendah selepas imbuhan *memper-*, iaitu masing-masing mencatatkan 0.004 dan 0.008, diikuti dengan imbuhan *mem-* yang mencatatkan nilai P yang paling rendah bagi kedua-dua korpus, iaitu masing-masing mencatatkan 0.002 dan 0.001. Di samping itu juga, saiz korpus sesuatu data juga turut mempengaruhi produktiviti morfologi sesuatu imbuhan itu. Berdasarkan perbandingan jadual hasil pengiraan imbuhan *mem-*, *memper-* dan *memper-kan* antara korpus Akhbar Sukan Dalam Talian yang berjumlah hampir 2.5 juta kata dengan korpus UKM-DBP yang berjumlah lima juta kata, dapat



Rajah 7 Bilangan hapax dan darjah produktiviti imbuhan *mem-*, *memper-* dan *memper-kan* korpus Akhbar Sukan Dalam Talian.



Rajah 8 Bilangan hapax dan darjah produktiviti imbuhan *mem-*, *memper-* dan *memper-kan* korpus UKM-DBP.

dilihat bahawa jumlah produktiviti data korpus UKM-DBP lebih tinggi terutamanya bagi imbuhan *memper-* yang berbeza nilai $P=0.021$ hasil produktiviti morfologi, diikuti dengan imbuhan *memper-kan* yang berbeza

nilai $P=0.004$, dan seterusnya imbuhan yang terendah, iaitu imbuhan *mem-* yang mencatatkan perbezaan nilai $P=0.001$.

Rajah 7 dan Rajah 8 meringkaskan produktiviti imbuhan *mem-*, *memper-* dan *memper-kan* bagi korpus Akhbar Sukan Dalam Talian dan korpus UKM-DBP. Paksi mendatar (x) mewakili bilangan *hapax* sementara paksi menegak (y) mewakili indeks produktiviti. Dalam kedua-dua rajah tersebut didapati imbuhan *memper-* mencatatkan nilai produktiviti tertinggi, diikuti dengan imbuhan *memper-kan* dan imbuhan *mem-* mempunyai nilai P yang terendah. Oleh itu, imbuhan *memper-* dikatakan lebih mempengaruhi kadar pertumbuhan perbendaharaan kata dengan kebarangkalian pembentukan kata terbitan baharu (neologisma) daripada imbuhan *memper-kan* dan *mem-*. Berdasarkan data, ketiga-tiga imbuhan boleh dikatakan produktif, namun dengan pengiraan produktiviti bersyaratkan *hapax* (*hapax conditioned*), dapat dikemukakan data empirikal berhubung tahap sebenarnya produktiviti berdasarkan pengiraan *hapax*. Berdasarkan formula ini juga, rumus *memper-kan* dan *mem-* dijangka akan menghasilkan neologisma yang kurang daripada *memper-*. Kajian secara kuantitatif terhadap produktiviti morfologi Melayu yang dilakukan di sini juga dapat memberikan rumus morfologi imbuhan *mem-*, *memper-* dan *memper-kan* yang lebih tepat berbanding dengan kaedah kualitatif. Kombinasi kaedah kuantitatif dengan kaedah kualitatif dalam mengkaji pola morfologi sesuatu bahasa itu menjadi amat penting untuk maklumat produktiviti morfologi mengikut ciri penggunaan. Kajian produktiviti morfologi secara kuantitatif dapat membantu mengenal pasti rumus yang lebih utama dalam sistem morfologi sesuatu bahasa. Maklumat seperti ini amat diperlukan dan dapat menyumbang dalam kajian linguistik dalam pengkomputeran dan pemprosesan bahasa tabii. Selain itu, kaedah ini juga membantu penyusun kamus dan pengkaji bahasa terutama dalam membuat penandaan morfologi bahasa terutamanya bahasa Melayu.

KESIMPULAN

Secara keseluruhannya, dapatlah disimpulkan bahawa penggunaan pendekatan ukuran produktiviti berdasarkan *hapax* Baayen (1992, 1996) dalam kajian ini telah berjaya mencapai objektifnya dalam memperlihatkan pola morfologi imbuhan *mem-*, *memper-* dan *memper-kan* bahasa Melayu dengan menggunakan dua bentuk korpus, iaitu Korpus Akhbar Sukan Dalam Talian dan korpus UKM-DBP. Berdasarkan ukuran *hapax*, darjah

produktiviti sesuatu imbuhan dinyatakan sebagai nisbah bilangan *hapax* dengan bilangan kata terbitan berbeza yang dibentuk oleh imbuhan itu.

Dalam kajian ini, indeks produktiviti Baayen menunjukkan imbuhan *memper-* adalah lebih produktif daripada imbuhan *memper-kan* dan *mem-* dalam kedua-dua korpus berkenaan. Produktiviti morfologi sesuatu imbuhan itu bergantung pada *hapax* dan saiz korpus. Dapatan kajian ini telah membantu menghuraikan morfologi satu lagi pembentukan kata secara kuantitatif yang menjadikan kajian ini lebih empiris dan autentik. Penggunaan pendekatan Baayen ini amat baik diterapkan juga dalam mengkaji imbuhan lain dengan berbantuan linguistik korpus dan teknologi komputer agar satu bentuk rumus morfologi terutamanya morfologi bahasa Melayu dapat didokumentasikan untuk kajian dan rujukan pengkaji bahasa dan penyusun kamus.

RUJUKAN

- Abdullah Hassan, 1974. *The Morphology of Malay*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Asmah Omar, 1982. *Nahu Melayu Mutakhir*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Asraf, 1990. *Petunjuk Bahasa Malaysia Baku KBSM*. Petaling Jaya: Sasbadi Sdn. Bhd.
- Arbak Othman, 1981. *Tatabahasa Bahasa Malaysia*. Kuala Lumpur: Sarjana Enterprise.
- Aronoff, M., 1980. "The Relevance of Productivity in a Synchronic Description of Word Formation" dlm. J. Fisiak (ed.). *Historical Morphology*. The Hague: Mouton.
- Baayen, R.H., 1989. "*A Corpus-Based Study of Morphological Productivity: Statistical Analysis and Psychological Interpretation*". Disertasi Ph.d. Free University, Amsterdam.
- Baayen, R.H. dan Lieber, R. "Productivity and English Word-Formation: A Corpus-Based Study" dlm. *Linguistics* 29, hlm. 801–43, 1991.
- Baayen, R.H. dan Renouf, A., "Chronicling the Times: Productive Lexical Innovations in an English Newspaper" dlm. *Language* 72, hlm. 69–96, 1996.
- Baayen, R.H., 2006. "Corpus Linguistics in Morphology: Morphological Productivity" dlm. Lüdeling, A. and M. Kytö (ed.). *Corpus Linguistics. An International Handbook*. New York: Mouton De Gruyter.
- Bauer, L., 2001. *Morphological Productivity*. Cambridge, UK: Cambridge University Press.

- Hockett, C., 1958. *A Course in Modern Linguistics*. New York: Macmillan.
- Plag, I., Dalton-Puffer, C., dan Baayen H., “Morphological Productivity across Speech and Writing” dlm. *Jurnal To Appear In English Language And Linguistics* 3.2, November 1999.
- Imran Ho Abdullah, “Produktiviti Imbuhan Apitan *Ke ... An* dan *Pen ... An*: Pendekatan Linguistik Korpus” dlm. *Jurnal Linguistik* 12, Edisi Khas, hlm. 1-11, Mac 2011.
- Nik Safiah Karim, Farid M. Onn, Hashim Musa dan Abdul Hamid Mahmood, 1997. *Tatabahasa Dewan*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Sproat, R., dan Shih, “A Corpus-Based Analysis of Mandarin Nominal Root Compound.” dlm. *Journal of East Asian Linguistics* 5, hlm. 49–71.8. C. 1996.
- Syed Zainal Ariff Syed Jamaluddin dan Noor Azaliza Yahya, “Perbandingan Darjah Transitiviti Awal *Ber-* dan *Per-* dalam Bahasa Melayu” dlm. *Jurnal Jurnal Bahasa* 13:1, hlm. 64-103, Jun 2013.
- Zaharani Ahmad, “Kekeliruan dalam Penggunaan Bahasa Melayu: Percanggahan Antara Rumus Tatabahasa dengan Korpus Bahasa” dlm. *Jurnal Persatuan Linguistik* 8, hlm. 191-230, 2007.
- Zaharani Ahmad dan Shakira Khairudin, “Kajian Imbuhan Kompleks Kata Nama *Ke-an* Bahasa Melayu: Analisis Data Korpus” dlm. *Jurnal Bahasa* 10:2, Disember 2010.
- Zaharani Ahmad, Shakira Khairudin dan Nor Hashimah Jalaluddin, “Perilaku Morfologi Awalan *Ber-* dalam Bahasa Melayu Klasik dan Bahasa Melayu Moden: Satu Kajian Perbandingan” dlm. *Jurnal Bahasa* 12:2, hlm. 181-203, Disember 2012.