



BUNGA RAMPAI

TATA BAHASA KONTEMPORER: MORFOLOGI





ARTIKEL TATA BAHASA KONTEMPORER: MORFOLOGI

Kontributor:

Prihantoro
Gede Primahadi Wijaya Rajeg
Karlina Denistia
Dewi Puspita
Dora Amalia
Adi Budiwiyanto
Raden Muhammad Arie Andhiko Ajie
David Moeljadi
Nazarudin

**Hak Cipta pada Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi
Republik Indonesia.
Dilindungi Undang-Undang.**

**BUNGA RAMPAI ARTIKEL TATA BAHASA KONTEMPORER INDONESIA:
MORFOLOGI
@2024**

Penulis

Prihantoro
Gede Primahadi Wijaya Rajeg
Karlina Denistia
Dewi Puspita
Dora Amalia
Adi Budiwiyanto
Raden Muhammad Arie Andhiko Ajie
David Moeljadi
Nazarudin

Penelaah

Sry Satriya Tjatur Wisnu Sasangka
Eva Tuckyta Sari Sujatna
Mohammad Umar Muslim
Umi Kulsum
Iwa Lukmana

Desain Sampul

Munafsin Azis

Pengatak

Nurjaman

Penerbit

Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi

Dikeluarkan oleh

Pusat Pengembangan dan Pelindungan Bahasa dan Sastra,
Badan Pengembangan dan Pembinaan Bahasa
Kawasan IPSC Jalan Tangkil, Km.4, Tangkil, Citeureup, Sukahati, Kabupaten Bogor, Jawa
Barat
<https://spiritpusbanglin.kemdikbud.go.id/portal.php>

Cetakan pertama. 2024
ISBN 978-623-504-583-2

KATA PENGANTAR

Bahasa akan selalu berkembang setiap saat sejalan dengan perkembangan penuturnya. Perkembangan tersebut juga tentu terjadi pada bahasa Indonesia. Selain perkembangan dari sisi penutur, kemajuan ilmu dan teknologi juga berpengaruh besar terhadap bahasa Indonesia. Mengingat perkembangan bahasa Indonesia yang tidak dapat dimungkiri itu, Badan Pengembangan dan Pembinaan Bahasa, melalui Pusat Pengembangan dan Pelindungan Bahasa dan Sastra memandang perlu disusunnya tata bahasa Indonesia yang menggambarkan penggunaan bahasa Indonesia saat ini melalui Tata Bahasa Kontemporer.

Penyusunan buku Tata Bahasa Indonesia Kontemporer dilakukan berdasarkan dua topik, yaitu morfologi dan sintaksis, yang masing-masing melalui tiga tahap, yaitu penulisan artikel oleh tim kontributor, reviu artikel oleh tim penyusun, dan penulisan bab oleh tim penyusun. Artikel yang telah direviu kemudian diseminarkan untuk mendapatkan balikan dan masukan dari para pengamat bahasa Indonesia.

Bunga rampai ini adalah kumpulan artikel yang ditulis oleh tim kontributor pada topik Morfologi. Artikel-artikel dalam bunga rampai ini sudah melalui tahap reviu, seminar, dan revisi pascaseminar. Selain dijadikan dasar untuk penulisan bab buku Tata Bahasa Kontemporer, hasil tulisan tim kontributor kami tayangkan utuh dalam bentuk bunga rampai agar gagasan aslinya juga dapat dibaca dan dijadikan rujukan oleh para peneliti, penulis, dan peminat bahasa Indonesia.

Bogor, Oktober 2023
Kepala Pusat Pengembangan dan
Pelindungan Bahasa dan Sastra



Imam Budi Utomo

DAFTAR ISI

Kata Pengantar	iii
Daftar Isi	iv
Tata Bahasa Indonesia Kontemporer Berbasis Korpus <i>Prihantoro</i>	1
Afiksasi Verba Dalam Bahasa Indonesia <i>Gede Primahadi Wijaya Rajeg dan Karlina Denistia</i>	28
Afiksasi Nomina Dalam Bahasa Indonesia <i>Karlina Denistia dan Gede Primahadi Wijaya Rajeg</i>	64
Afiks Pembentuk Adjektiva dalam Bahasa Indonesia <i>Dewi Puspita dan Dora Amalia</i>	94
Reduplikasi Utuh dan Sebagian dalam Bahasa Indonesia <i>Dora Amalia dan Dewi Puspita</i>	120
Pemajemukan dalam Bahasa Indonesia: Pendekatan Distribusional <i>Adi Budiwiyanto</i>	144
Pemajemukan Kata Dalam Bahasa Indonesia: Ancangan Fraseologis <i>M. Arie Andhiko Ajie dan Adi Budiwiyanto</i>	172
Reduplikasi Sebagian, Salin Suara, dan Trilingga dalam Bahasa Indonesia <i>David Moeljadi</i>	202
Akronim dalam Bahasa Indonesia <i>Nazarudin</i>	224

TATA BAHASA INDONESIA KONTEMPORER BERBASIS KORPUS

Prihantoro

Universitas Diponegoro

Abstrak

Artikel ini merupakan artikel pendahulu buku Tata Bahasa Indonesia Kontemporer (TBIK) berbasis yang analisisnya berbasiskan data korpus. Tujuan artikel ini adalah memberikan gambaran bagaimana linguistik korpus bisa berperan dalam analisis TBIK. Selain itu, artikel ini bertujuan memberikan pondasi untuk memahami artikel-artikel lain dalam buku TBIK yang juga ditulis menggunakan metode dan teknik penelitian linguistik korpus. Pada bagian pertama, akan dibahas definisi tata bahasa yang tidak membatasi deskripsi pada level morfosintaksis saja, tapi juga level analisis linguistik yang lain seperti fonetik, fonologi, dan morfologi. Bagian kedua membahas hubungan antara komputer, program, dan korpus (data linguistik) serta beberapa korpus Bahasa Indonesia (BI). Arsitektur korpus TBIK, korpus yang digunakan sebagai salah satu referensi utama dalam paper-paper di buku ini, dibahas pada bagian ke tiga. Sedangkan bagian keempat membahas tentang model analisis linguistik korpus yang menggunakan data korpus TBIK maupun korpus BI lain yang dibahas sebelumnya. Pada bagian terakhir kita akan membahas beberapa perspektif terkait model analisis linguistik korpus.

Kata kunci: korpus, kontemporer, tata bahasa, metadata

1. TATA BAHASA INDONESIA

1.1 Bahasa Indonesia

Dalam rumpun bahasa Austronesia, Bahasa Melayu (BM) merupakan bahasa yang sebarannya cukup luas di Asia Tenggara, seperti BM Patani di Patani, Thailand, BM Ambon di Ambon, Indonesia, BM Kelantan di Kelantan Malaysia. Di Indonesia sendiri, terdapat banyak varian BM yang tersebar di beberapa daerah (lihat Sneddon et al [1]). Karena sama-sama berasal dari rumpun BM, tidak mengherankan ada beberapa elemen varian-varian BM ini yang beririsan, misalnya dalam hal kosakata yang sama, digunakan di lebih dari satu varian BM.

Melalui proses standarisasi tertentu, beberapa varian BM diadopsi menjadi bahasa resmi di beberapa negara Asia tenggara, seperti di Indonesia, Malaysia, Singapura dan Brunei Darussalam. BM yang telah distandarisasi di Indonesia, dan digunakan sebagai bahasa negara resmi disebut sebagai Bahasa Indonesia (BI), yang kedudukannya secara administratif di Indonesia telah diatur di Undang-Undang No. 24 tahun 2009¹ (mengenai bendera, bahasa, lambang negara serta lagu kebangsaan).

Seiring waktu berjalan, proses-proses standarisasi BI menghasilkan beberapa penciri khusus dibandingkan varian BM yang lain, baik yang telah distandarisasi menjadi bahasa resmi negara maupun yang tidak. Salah satu penciri khusus BI adalah tata bahasa. Satu hal yang perlu disepakati sebelum kita membaca lebih lanjut adalah, tata bahasa di sini dimaknai secara makro, sehingga aspek linguistik yang dibahas tidak hanya pada level sintaksis atau morfologi saja tapi juga level linguistik yang lain seperti semantik, leksikografi, fonetik, fonologi maupun ortografi (termasuk ejaan dan penggunaan tanda baca).

Perbedaan tata bahasa tersebut secara singkat bisa kita lihat pada contoh berikut. Dalam bahasa Indonesia, kita biasa menyebut bulan ke delapan dengan nama bulan *Agustus*. Akan tetapi, dalam varian standar bahasa Melayu yang digunakan di Malaysia dan Brunei Darussalam, kata yang digunakan adalah *Ogos*, bukan Agustus.

Di sini kita lihat ada perbedaan ejaan meski Bahasa Indonesia dan Bahasa Malaysia dan Brunei masih merupakan satu rumpun melayu. Perbedaan lebih kecil terdapat pada preposisi BI, *karena* yang dalam BM Brunei dan Malaysia dieja *kerena*. Perbedaan yang sangat ekstrim mungkin bisa diilustrasikan pada kata *handuk* yang dalam BM Malaysia dan Brunei disebut *tuala*. Pada kasus ini, digunakan kosakata yang sangat berbeda, dan tidak memiliki kemiripan sama sekali dengan BI.

Istilah MABBIM				
Istilah Sumber	Istilah Indonesia	Istilah Brunei	Istilah Malaysia	Bidang
backhaus towel clamp	penjepit handuk Backhaus	penjepit tuala Backhaus	penjepit tuala Backhaus	Veterinar

Gambar 1 Perbedaan terjemahan *towel* dalam BI dan BM Malaysia dan Brunei, diambil dari Istilah MABBIM²

- <https://jdih.mkri.id/mg58ufsc89hrsg/193fa997c8319d8606f1747565e49cf2de73ddebe.pdf> (diakses tanggal 6 April 2023)
- <https://prpm.dbp.gov.my/Cari1?keyword=handuk&d=103448&#LIHATSINI> (diakses tanggal 6 April 2023)

Dari contoh sederhana ini bisa kita lihat bahwa salah satu fungsi tata bahasa adalah membedakan struktur satu bahasa dengan bahasa lain. Tentu saja masih banyak fungsi yang lain, yang tidak akan dibahas pada paper ini karena keterbatasan tempat. Pada bagian selanjutnya, kita akan secara singkat membahas mengenai buku tata bahasa Indonesia yang sudah ada.

1.2 Buku Tata Bahasa Baku Bahasa Indonesia dan Tata Bahasa Indonesia Kontemporer

Bukankah sudah ada buku Tata Bahasa Baku Bahasa Indonesia (TBBBI) [2]? Benar, Buku ini membahas konsep-konsep dasar tata bahasa³ BI secara komprehensif, mulai dari fonetik, fonologi, morfologi, semantik, sintaksis, hingga wacana. Sejauh ini, TBBBI (menurut hemat kami) adalah buku referensi tata bahasa (*reference grammar*) yang pembahasan aspek-aspek linguistiknya paling lengkap dibandingkan buku-buku sejenis, misalnya Indonesian Reference Grammar [1].

Buku TBBBI terbit pertama kali tahun 1988 untuk menyongsong kongres Bahasa Indonesia V, dan disempurnakan pada tahun-tahun berikutnya. TBBI edisi terbaru adalah TBBI edisi ke-4 yang terbit pada tahun 2017, yang hingga saat tulisan ini dibuat, dapat diakses secara daring pada laman kemendikbud berikut⁴. Versi daring buku ini sangat membantu karena dapat menjangkau masyarakat yang lebih luas.

Sekarang, bagaimana dengan TBIK? Buku TBIK juga dirancang untuk membahas tata bahasa BI dari berbagai aspek linguistik seperti TBBBI. Namun demikian, ada beberapa perbedaan penting yang patut diberikan sorotan. Pertama, data-data bahasa Indonesia yang disajikan dalam artikel-artikel di buku TBIK diambil dari data bahasa autentik, disusun secara sistematis, dan tersimpan secara digital. Data ini disebut *korpus*, yang akan kita bahas pada bagian 2.1. Oleh sebab itu, dapat dipastikan bahwa kata, frasa, kalimat yang ada dalam data-data tersebut memang benar digunakan oleh penutur BI, bukan rekaan dari penulis. Karena datanya bersifat digital dan tersusun secara sistematis, tentulah akan lebih mudah untuk melakukan replikasi dari analisis yang tersaji.

3 Ilustrasi analisis disajikan pada Bagian 4

4 <https://repositori.kemdikbud.go.id/16351/1/Tata%20Bahasa%20Baku%20Bahasa%20Indonesia%20edisi%20keempat.pdf> (diakses tanggal 6 April 2023)

Kedua, korpus referensi utama dari setiap artikel yang ada di buku ini adalah Korpus Tata Bahasa Indonesia Kontemporer (Korpus TBIK). Sebagai data-data pendukung, para penulis juga menggunakan korpus Bahasa Indonesia yang lain. Korpus TBIK dan beberapa korpus bahasa Indonesia akan dibahas pada bagian 2.2 dan bagian 2.3. Jika korpus yang digunakan tidak ada pada artikel ini, maka pada artikel terkait akan dibahas secara ringkas korpus apa saja yang digunakan. Ketiga, data-data yang digunakan dalam korpus ini adalah data kontemporer, yang diambil dari tahun 2011 hingga 2020. Keempat, pembaca nanti akan melihat beberapa sentuhan kuantitatif. Sentuhan kuantitatif bisa diilustrasikan salah satunya dengan frekuensi pemunculan unit linguistik.

No.	Query result	No. of occurrences	Percent
1	<i>praktik</i>	2557	72.15%
2	<i>praktek</i>	987	27.85%

Gambar 2 Frekuensi pemunculan *praktik* dan *praktek* di korpus TBIK

Misalnya, manakah yang lebih banyak digunakan? *Praktik* atau *praktek*? Dengan menggunakan program analisis korpus, kita bisa dengan yakin mengatakan bahwa *praktik* lebih banyak digunakan. Pertanyaan seperti ini secara introspektif bisa berbeda-beda jawabannya antara satu penutur dengan penutur lain. Namun dengan keberadaan teknologi korpus, konfirmasi dapat dilakukan dengan lebih mudah.

Secara ringkas, artikel ini memberikan jawaban untuk beberapa pertanyaan, yaitu: (1) apa peran komputer, program, dan data bahasa dalam analisis linguistik korpus? (2) seperti apakah arsitektur korpus TBIK yang digunakan? (3) model analisis linguistik korpus seperti apa yang bisa digunakan? (4) perspektif apa yang ditawarkan oleh artikel ini? Jawaban atas pertanyaan-pertanyaan ini dapat ditemukan secara urut pada bagian 2, 3, 4 dan 5.

2. KOMPUTER, PROGRAM, DAN DATA

Tidak dapat dipungkiri, teknologi telah memberikan kontribusi yang sangat besar dalam perkembangan ilmu pengetahuan termasuk linguistik. Misalnya, data linguistik yang sangat besar dan kalau dicetak bisa memenuhi ruang perpustakaan, sekarang hanya cukup disimpan dalam sebuah komputer

atau sebuah hard-disk portabel. Program yang ada dalam komputer tersebut memungkinkan kita untuk melakukan penelusuran, observasi dan analisis. Misalnya, kita bisa melakukan pengarsipan dan penelusuran data menggunakan Windows Explorer. Kita bisa membaca teks menggunakan program seperti MS Word. Kita dapat melakukan analisis teks secara statistik menggunakan program SPSS.

Pada intinya, ada tiga hal yang sangat penting untuk kita bedakan sebelum kita masuk ke bagian selanjutnya. Yang pertama adalah data yang kita miliki seperti teks, transkrip, data demografis yang relevan (misalnya jenis kelamin, asal, bahasa), maupun catatan-catatan mengenai data tersebut baik yang bersifat khusus maupun umum. Yang kedua adalah perangkat keras yang kita gunakan untuk melakukan akses dan penyimpanan data, dalam hal ini komputer. Yang ketiga adalah perangkat lunak untuk membaca, menganalisis, maupun memodifikasi (seperti menambah, memperbaiki, mengurangi, memberikan penekanan) data tersebut.

2.1 Pendekatan Linguistik Berbasis Teknologi Korpus

Data linguistik yang sifatnya alami, baik lisan atau tulisan, bisa tuturan maupun teks, dapat kita sebut sebagai korpus [3]. Inilah definisi *korpus* yang paling sederhana. Pada banyak buku teks linguistik korpus (baca [4], [5], dan [6]), ada beberapa definisi lanjutan mengenai korpus *modern* yang ditawarkan, namun satu aspek yang menonjol adalah data tersebut berbentuk digital. Tentu saja hal ini dimaksudkan agar data tersebut mudah diakses, disimpan, ditelusuri dan dianalisis menggunakan komputer dan program-program yang relevan, baik program yang berbasis website, maupun program-program yang harus diinstall terlebih dahulu ke komputer pengguna.

Paling tidak ada tiga jenis program yang biasa digunakan dalam bidang kajian linguistik korpus. Yang pertama adalah program-program yang berfungsi melakukan pemrosesan data korpus. Ada beberapa jenis pemrosesan data korpus (lihat [3]). Di sini akan dibahas dua jenis pemrosesan saja.

Pemrosesan data yang pertama adalah tokenisasi. Tokenisasi secara sederhana bisa kita definisikan sebagai proses ‘pemotongan’ teks menjadi unit analisis yang lebih kecil, yang disebut *token*. Unit ini bisa berada pada tataran linguistik tertentu, seperti morfem, kata, klausa, atau kalimat, tergantung dari kebutuhan kita.

Tokenisasi yang cukup banyak diaplikasikan adalah tokenisasi pada level kata. Misalnya, pada teks *kita bisa naik kereta api* di bawah, kita bisa menganggap bahwa kalimat ini mengandung 5 token pada level kata. Cara menentukan tokennya sederhana, yaitu hanya dengan menggunakan spasi sebagai pembatas antar token (*token delimiter*).

(1) *kita bisa naik kereta api*

Namun kita juga bisa menggunakan skema lain, misalnya dengan menganggap bahwa *kereta api* adalah satu token tersendiri, bukan dua token yang berbeda. Ini bisa didasarkan pada argumen bahwa *kereta api* dianggap sebagai satu kompositum tersendiri yang maknanya non-komposisional. Kereta yang berbasis diesel tidak mengeluarkan api, namun tetap disebut kereta api, atau kereta saja. Sehingga, meskipun ada spasi diantara *kereta* dan *api*, kedua kata tersebut hanya dianggap sebagai satu token saja.

Meski tokenisasi dilakukan pada level kata, kita juga bisa menetapkan aturan tertentu mengenai tanda baca. Misalnya saja, program yang digunakan untuk melakukan tokenisasi di LancsBox tidak menganggap tanda baca sebagai token karena bukan kata. Sedangkan program tokenisasi yang digunakan di TreeTager [7] menganggap tanda baca seperti koma, titik, tanda baca atau seru sebagai token. Sehingga, jika kalimat *kita bisa naik kereta api* disusul dengan tanda titik, maka titik itu sendiri juga akan dihitung sebagai token.

No.	Word	Frequency
1	.	1,516,202
2	,	1,509,515
3	yang	770,495
4	dan	653,062
5	di	365,949

Gambar 3 Daftar frekuensi (*frequency list*) yang memasukkan tanda baca sebagai token

Bagaimana dengan kata *uangku* pada contoh kalimat di bawah ini? Apakah *uangku* dianggap satu token saja, atau dianggap dua token yaitu *uang* dan *ku* (bentuk klitik dari *aku*)? Ini tentu saja tergantung dengan skema tokenisasi yang kita pakai, dengan argumen tertentu, baik yang sifatnya preskriptif, maupun pragmatis (baca: untuk kemudahan pemrosesan). Intinya,

semakin rumit skema tokenisasi yang kita pakai, semakin rumit juga desain program yang kita gunakan.

(2) *Aku ambil uangku*

Program yang kedua adalah program berjenis sistem anotasi. Program ini memberikan label analisis pada token yang ada dalam korpus. Misalnya, program POS tagger memberikan label kelas kata untuk setiap token. Misalnya, kalimat *kita pergi ke masjid* dapat dianotasi menggunakan program POS tagger, dan hasilnya kira-kira seperti ini.

(3) kita_PRP
pergi_VB
ke_IN
masjid_NN

Bisa kita lihat, ada label anotasi yang disisipkan di ujung kanan setiap token. Deskripsinya sebagai berikut, pronomina (PRP), verba (VB), preposisi (IN), nomina (NN). Desain labelnya ini bersifat manasuka, sehingga tidak masalah jika kita ingin menggunakan label yang berbeda (misalnya, menggunakan PRO, untuk pronomina) atau memilih mengikuti label anotasi yang sudah ada misal, label anotasi yang dipakai di [8] atau di [9].

Selain anotasi pada level kata, anotasi juga bisa dilakukan pada level morfem atau frasa/kalimat. Karena keterbatasan tempat, tidak semua akan dijelaskan di sini. Beberapa referensi ini mungkin dapat membantu. Untuk anotasi morfem, bisa membaca [10], [9], [11]. Sedangkan untuk anotasi frasa/kalimat, salah satunya bisa membaca [12].

Program yang ketiga adalah program indeksasi, penelusuran, dan analisis data korpus. Indeksasi bermakna seperti *membaca*. Misalnya, kita membutuhkan program MS Word untuk membaca file docx. Kita membutuhkan file Acrobat Reader untuk membaca file pdf. Nah, untuk membaca file korpus, kita bisa gunakan program baik yang berbasis web seperti CQPweb [13], Sketch Engine [14], dan program yang harus kita install dulu ke komputer lokal seperti WordSmith [15], LancsBox [16] atau AntConc [17]. Program-program yang disebutkan tadi sudah memiliki fitur yang bisa kita pakai untuk melakukan analisis seperti analisis konkordansi, kolokasi dan kata kunci (lihat lengkapnya di [10], atau [5]). Untuk bagian analisisnya, kita akan bahas lebih lanjut di bagian ke 4 dari artikel ini.

Lantas bagaimana dengan data-data yang hanya tidak bisa diakses publik? Bisakah data-data tersebut disebut korpus? Banyak sekali data-data linguistik kecil yang digunakan untuk melakukan penelitian misalnya skripsi, tesis, disertasi, atau menulis artikel. Tidak semuanya diunggah ke publik. Beberapa data linguistik dikumpulkan dengan perjanjian tertentu, sehingga hanya bisa diakses anggota tim atau orang-orang yang ditentukan. Ada pula data yang memang sifatnya komersil dan tidak bisa diakses secara langsung. Misalnya, korpus BNC (British National Corpus) [18] [19]. Kondisi-kondisi semacam ini tidak menghalangi data tersebut untuk disebut sebagai korpus. Bahkan ada juga data yang awalnya tertutup, namun seiring berjalannya waktu menjadi terbuka untuk publik. Salah satunya adalah data BNC.

2.2 Beberapa Korpus Bahasa Indonesia

Pada bagian ini kita akan membahas beberapa korpus BI yang bisa diakses secara bebas. Link ke korpus-korpus tersebut akan diberikan pada tabel 1. Yang pertama kita bahas adalah Korpus Indonesia (KOIN), yang dapat diakses menggunakan aplikasi web milik BPBB Kemendikbud Ristekdikti. Teks korpus diambil dari 5 bidang utama, yaitu sains, sastra, media massa, hukum, dan korespondensi. Korpus ini adalah korpus monolingual BI yang ukurannya sekitar 10 juta token dan masih akan berkembang.



The image shows a web interface for searching the KOIN (Korpus Indonesia) database. It features a search bar at the top labeled 'Cari'. Below it, there is a 'Kategori' (Category) section with a grid of checkboxes for various domains: Fisika, Sosial, Kesehatan, Hayati, Sastra, Opini, Editorial, Nasional, Internasional, Olahraga, Ekonomi dan Keuangan, Gaya hidup, Gadget, Travelling, Perempuan & Kecantikan, Otomotif, Film, Budaya, Biografi, Peraturan dan UU, and Surat Menyurat. All these checkboxes are currently checked. To the right of the categories is a blue search button with a magnifying glass icon. At the bottom left, there is a 'Regex' toggle switch which is currently turned off, and a 'MATI' button.

Gambar 4 Tampilan penelusuran KOIN berdasarkan domain korpus

Di aplikasi CQPweb Lancaster ada beberapa korpus BI yang dapat digunakan. Yang pertama adalah korpus LCC Indonesian yang datanya diambil dari Leipzig Corpora Collection. Teks korpus ini dikumpulkan secara otomatis dari Internet. Di CQPweb, korpus ini dianotasi menggunakan

IPOSTagger [20]. Korpus BPPT-PAN Indonesian dan UI-1M adalah dua korpus lain yang diindeks di CQPweb. Data kedua korpus ini diperoleh dari PAN Localisation Project [21]. Korpus BPPT-PAN Indonesian adalah korpus berbahasa Indonesia yang memiliki pasangan data paralel berbahasa Inggris, yaitu BPPT-PAN English yang juga diindeks di CQPweb. Dari segi variasi, teks di korpus ini diambil dari 4 domain, yaitu olahraga, sains, ekonomi, dan internasional. Seperti Korpus LCC Indonesian, Korpus BPPT-PAN juga dianotasi menggunakan IPOSTagger. Sedangkan korpus UI-1M adalah korpus yang sudah beranotasi. Dua korpus ini ukurannya tidak terlalu besar, yaitu 500K dan 900K, masing-masing.

pada 2005 menjadi Rp12,07 pada 2006. pada 2005 perseroan **mencetak** laba bersih Rp92,55 miliar atau melonjak 152,74
In 2005 , the company posted a net profit of Rp92.55 billion , up 152.74 percent from a year earlier .
orasi untuk membentuk agen asuransi profesional . Kami ingin **mencetak** agen-agen asuransi jiwa nasional yang berkompetensi
ant to have competent life insurance agents who have international disciplines , AAJI chairwoman Evelina Pietruschka said
ikan Selasa . Sepanjang semester pertama 2007 Holcim hanya **mencetak** laba bersih sebesar Rp5,76 miliar , padahal pada
. Holcim gained a net profit of Rp5.76 billion only , a sharp decrease from last year 's net profit in the same period last year

Gambar 5 Hasil penelusuran konkordansi *mencetak* pada korpus BPPT-PAN Indonesian dan segmen paralelnya dalam korpus BPPT-PAN English

Di Sketch Engine, ada dua korpus BI, yaitu Indonesian Web as a Corpus (WaC) dan Indonesian Open Subtitle. Ke-dua korpus ini ukurannya cukup besar, yaitu 90M dan 70M token. Dari segi variasi teks, data korpus WaC diambil dari website berbahasa Indonesia, sedangkan korpus Indonesian Open Subtitle diambil dari terjemahan transkrip film.



Gambar 6 Kolokasi *bali* pada korpus Indonesia WaC

Tabel di bawah meringkas poin-poin penting dari korpus-korpus yang telah didiskusikan dengan singkat sebelumnya. Pada bagian selanjutnya, kita akan membahas korpus TBIK, korpus yang digunakan sebagai referensi utama artikel-artikel yang ada di buku ini.

Tabel 1 Beberapa korpus BI yang bisa diakses publik

Korpus	Domain	Bahasa	Anotasi	Ukuran	Sumber
KOIN ⁵	Sains Sastra Media masa Hukum Korespondensi	Indonesia, Monolingual	POS	15M	BPPB
LCC Indonesian CQPweb ⁶	Tahun	Indonesia, Monolingual	POS	232M	Leipzig Corpora Collection, raw

5 <https://korpusindonesia.kemdikbud.go.id/index.php?r=site/home>

6 <https://cqweb.lancs.ac.uk/lccindonesian2/> (diakses tanggal 6 April 2023)

BPPT-PAN Indonesian CQPweb ⁷	Ekonomi Olahraga Sains Internasional	Paralel dengan BPPT-PAN Bhs Inggris ⁸	POS	0.55M	BPPT_PAN Localisation project, mentah
UI-1M ⁹	TA ¹⁰	Indonesia, Monolingual	POS	0.95M	BPPT_PAN Localisation project, beranotasi
Indonesian WaC Sketch Engine ¹¹	TA	Indonesian, Monolingual	POS	90M	Web, beranotasi
IndoPhone ¹²	TA	Bahasa Indonesia dan daerah	TA	TA	The Language Archive ¹³
Indonesian Open Subtitle Sketch Engine ¹⁴	TA	Indonesian, Parallel	POS	77M	Subtitle

3. KORPUS TBIK

Berdasarkan struktur data korpus dan fungsinya, korpus dapat dikategorikan menjadi beberapa jenis, salah satunya adalah korpus umum. BNC, Longman Corpus dan COCA adalah beberapa contoh korpus umum untuk bahasa Inggris. Data korpus ini cukup bervariasi mulai dari domain, periodisasi/tahun penerbitan, jenis kelamin penulis dan sebagainya. Ini berbeda dengan korpus spesialis, yang datanya biasanya diperoleh dari domain yang sangat spesifik. Salah satu contohnya adalah SCOTUS, yaitu korpus bahasa Inggris bidang hukum yang datanya diperoleh dari keputusan pengadilan di Amerika Serikat.

Korpus umum dapat difungsikan sebagai referensi untuk membuat dokumentasi tata bahasa. Misalnya saja, buku Longman Grammar of Written and Spoken English dibuat menggunakan referensi korpus Longman.

7 <https://cqweb.lancs.ac.uk/bpptpanid/> (diakses tanggal 6 April 2023)

8 <https://cqweb.lancs.ac.uk/bpptpanen/> (diakses tanggal 6 April 2023)

9 <https://cqweb.lancs.ac.uk/ui1m/> (diakses tanggal 6 April 2023)

10 TA = tidak ada

11 <https://www.sketchengine.eu/indonesianwac-corpus/> (diakses tanggal 6 April 2023)

12 <https://cqweb.lancs.ac.uk/indophone/> (diakses tanggal 6 April 2023)

13 <https://archive.mpi.nl/tla/> (diakses tanggal 6 April 2023)

14 <https://www.sketchengine.eu/opensubtitles-parallel-corpora/> (diakses tanggal 6 April 2023)

Bagaimana dengan korpus TBIK? Korpus TBIK berdasarkan jenisnya juga termasuk korpus umum. Korpus ini bukanlah korpus spesialis yang datanya hanya diperoleh dari satu domain saja, namun data korpus TBIK diperoleh dari berbagai domain seperti koran, cerpen, perundangan, dan sebagainya yang akan kita bahas pada bagian selanjutnya. Korpus TBIK, seperti juga Longman Corpus, juga digunakan sebagai sumber data buku TBIK ini.

Korpus Referensi TBIK v1.3: powered by CQPweb	
Metadata for Korpus Referensi TBIK v1.3	
Corpus ID number	1388 (base 36 code: 00012k)
Date of installation on system	2022-Sep-26 08:33
Corpus title	Korpus Referensi TBIK v1.3
CQPweb's short handles for this corpus	tbik3 / TBIK3
Total number of texts in corpus	17,277
Total word tokens in all corpus texts	29,987,513
Word types in the corpus	445,500
Standardised type:token ratio (1,000-token basis)	0.3969 types per token
Non-standardised type:token ratio	0.0149 types per token
Indexed to CQPweb by	Prihantoro (prihantoro@live.undip.ac.id; prihantoro.rf.gd)

Gambar 7 Preview singkat data korpus TBIK

Gambar di atas memberikan pratinjau mengenai korpus ini. Bisa kita lihat bahwa ukuran korpus ini hampir mencapai 30 juta token, terdiri dari lebih dari 17 ribu teks lebih, dengan 445.500 kata unik. Teks ini sudah dianotasi pada level kata dengan dua jenis anotasi, yaitu kelas kata dan kata dasar. Pada bagian selanjutnya, kita akan membahas tentang arsitektur dan proses penciptaan korpus ini dengan lebih detail.

3.1 Metadata

Secara teknis sederhana, suatu korpus dapat terdiri dari kumpulan teks. Informasi mengenai teks inilah yang kita sebut metadata seperti peroidesasi, domain teks, jenis kelamin penulis, tingkat kesulitan dan sebagainya (lihat BNC 2014 [19]).

Korpus TBIK memiliki dua jenis metadata, yaitu tahun terbit dan jenis teks. Tahun terbit adalah tahun saat teks tersebut diterbitkan. Jangka waktu yang digunakan adalah tahun 2011 hingga 2020 (10 tahun). Metadata yang lain adalah jenis teks berdasarkan domain atau jenis teks sumber, yang saat ini jumlahnya adalah dua belas domain (Koran, Majalah, Cerpen, Novel, Buku Teks, Jurnal, Disertasi Tesis Skripsi, Biografi, Populer, Perundangan, Laman Resmi dan Surat Resmi).

Tabel di bawah ini menampilkan sampel file metadata. Dapat kita lihat bahwa kolom pertama menampilkan kode unik untuk setiap file. Kolom kedua menampilkan klasifikasi berdasarkan domain, dan kolom ketiga berdasarkan tahun terbit. Setiap baris harus unik. Misalnya meskipun dua baris pertama memiliki kesamaan dalam hal domain dan tahun, namun kode teksnya berbeda. Pada baris ke tiga sampai ke lima, bisa kita lihat bahwa semua informasinya berbeda, mulai dari kode unik, domain dan tahun terbit.

Tabel 2 Sampel file metadata korpus TBIK dengan informasi tahun dan domain teks

Nama file	Domain	Tahun
A1ZZ11099	A_koran	2011
A1ZZ11100	A_koran	2011
A1ZZ12001	A_koran	2012
B2I14066	B_majalah	2014
J1E19002	J_perundangan	2019

Kita bahas sedikit mengenai ukuran korpus berdasarkan pembobotan metadata. Bisa kita lihat pada tabel di bawah ini, bahwa jika dilihat dari tahun terbit, komposisi data setiap tahun (selama 10 tahun) berkisar dari 2833293(paling sedikit di tahun 2011) hingga 3228066 (paling banyak di tahun 2020). Dari domain, kita bisa juga melihat komposisinya dengan paling sedikit pada data jurnal dengan 2410112 token, dan paling banyak 2703612 pada data Disertasi, Tesis dan Skripsi.

Tabel 3 Kompisisi korpus TBIK berdasarkan tahun dan domain

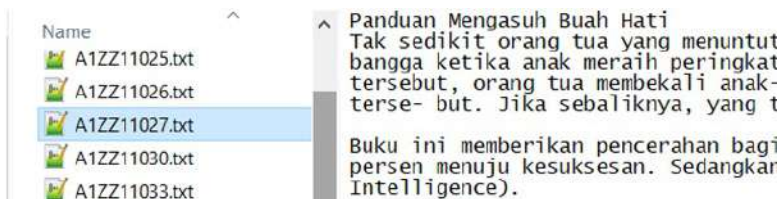
Tahun	Ukuran (token)	Domain	Ukuran (token)
2011	2833293	Koran	2701024
2012	2879347	Majalah	2439621
2013	3001535	Cerpen	2504297

2014	2893034	Novel	2622758
2015	2958978	Buku Teks	2467943
2016	2950811	Jurnal	2410112
2017	2951334	Disertasi Tesis Skripsi	2703612
2018	3205346	Biografi	2450463
2019	3070041	Populer	2450463
2020	3228066	Perundangan	2500251
		Laman Resmi	2440866
		Surat Resmi	2440607
Total	29987513	Total	29987513

3.2 Pemrosesan

Pemrosesan korpus dilakukan dengan menyesuaikan format yang diterima di CQPweb, program korpus yang digunakan untuk mengindeks korpus TBIK. Hal ini penting untuk ditetapkan lebih awal, karena dengan sistem yang berbeda, maka bisa jadi dibutuhkan format yang berbeda pula.

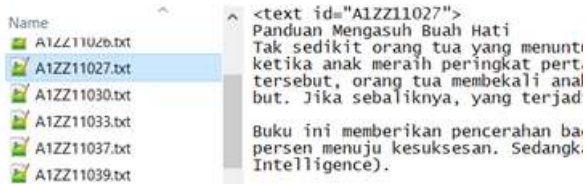
Sebagai catatan, untuk teks yang sifatnya tidak terlalu kompleks, pengguna biasa dapat mengupload teks secara langsung ke CQPweb dan pemrosesan akan dilakukan secara otomatis menggunakan resource yang tersedia¹⁵. Namun, untuk data yang ukurannya besar, dengan metadata yang kompleks, seperti TBIK, maka harus dilakukan oleh admin yang proses otomasinya dilakukan secara bertahap. Proses diawali dengan data korpus yang masih mentah. Data mentah ini berupa kumpulan file berekstensi .txt yang dikumpulkan dari sumber-sumber yang berasal dari 12 kategori domain, dan terbit di tahun 2011-2012. Setiap teks, atau bagian teks disalin dari sumber aslinya dan disimpan ke dalam file teks berbentuk .txt. Teks masih tersaji secara normal.



Gambar 8 Tampilan salah satu teks mentah korpus TBIK

¹⁵ Lihat tutorialnya di sini <https://www.youtube.com/watch?v=qarwSQPg3n4&t=213s> (diakses tanggal 6 April 2023)

Kemudian, setiap teks diberi penanda berupa tag XML pembuka di bagian atas yang isinya adalah *text id* atau judul teks (tanpa ekstensi txt), misalnya `<text id="A1ZZ11027">`. Kode ini harus sama persis dengan kode yang ada di file metadata yang telah dijelaskan sebelumnya. Misalnya jika teks berjudul A1ZZ11027.txt, maka kode yang ada di metadata dan text id haruslah juga sama. Di akhir teks, haruslah diberi tag penutup `</text>`.



Gambar 9 Tampilan teks setelah diberi text id

Kemudian, diaplikasikanlah program tokeniser untuk memotong teks tersebut menjadi token-token yang nantinya akan dianalisis. Program yang digunakan adalah skrip Perl tokeniser yang ada di dalam TreeTagger. Program ini mengubah format teks. Token disajikan secara vertikal. Satu baris terdiri dari satu token. Setelah itu, dilakukanlah pemeriksaan dan penyesuaian data. Misalnya, ada beberapa data yang belum bertokenisasi secara sempurna. Perhatikan contoh kalimat di bawah ini.

- (4) *Sikap Indonesia dalam masalah ini --sebagaimana Saudi-- akan menjadi bandul*

Kata *sebagaimana* dan *Saudi* tidak berhasil ditokenisasi dengan sempurna, karena ada tanda dua strip -- yang menempel. Kesalahan-kesalahan seperti ini diperbaiki sehingga dua kata tersebut bertokenisasi dengan baik, seperti contoh di bawah ini:

Tabel 4 Contoh perbaikan tokenisasi

Sebelum	Sesudah
--sebagaimana Saudi--	-- Sebagaimana Saudi --

Selanjutnya, data korpus ini diberi kode unik metadata seperti yang telah dibahas pada bagian sebelumnya dengan menggunakan tag XML sebagai penanda kode teks. Kemudian, dilakukan proses tokenisasi untuk setiap file. Proses ini dilakukan menggunakan program tokeniser yang ada program TreeTager. Hasilnya adalah satu baris terdiri dari satu token. Pada tahap ini sebetulnya korpus sudah bisa diindeks karena sudah memenuhi syarat untuk diindeks sebagai korpus mentah.

```
masalah
ini
--
sebagaimana
Saudi
--
akan
menjadi
bandul
bagi
dukungan
```

Gambar 10 Teks setelah ditokenisasi

Setelah tokenisasi, diaplikasikanlah program TreeTagger¹⁶, modul BI. Program ini menambahkan label anotasi kelas kata di kolom selanjutnya, lalu kata dasar di kolom sebelahnya. Sehingga, satu baris terdiri dari tiga komponen: token, label anotasi, dan kata dasar.

```
masalah>NN→masalah
ini>PR→ini
-->Z→--
sebagaimana>SC→sebagaimana
Saudi→NNP→saudi
-->Z→--
akan→MD→akan
menjadi>VB→jadi
bandul→NN→bandul
```

Gambar 11 Teks setelah ditokenisasi dan dianotasi

Hingga saat tulisan ini dibuat, ada dua jenis anotasi yang melekat pada TBIK, yaitu anotasi kelas kata dan kata dasar, seperti yang telah dijelaskan di

¹⁶ <https://www.cis.lmu.de/~schmid/tools/TreeTagger/> (diakses tanggal 6 April 2023)

atas. Namun bukan tidak mungkin, di masa depan, akan ada anotasi-anotasi lainnya seperti anotasi semantik, sintaksis, bahkan prgramatik! Namun untuk saat ini, kita cukupkan dulu anotasi pada bagian kelas kata dan kata dasar.

3.3 Indeksasi dan Lisensi

Ada dua jenis file yang perlu diindeks. Yang pertama adalah file-file teks korpus (seperti disebutkan pada bagian 3.2) dan yang kedua adalah file metadata (seperti dijelaskan pada bagian 3.1). Setelah kedua file ini terindeks dengan sempurna, maka ditambahkan file-file pendukung, misalnya *Corpus Manual* yang menjelaskan tentang cara penggunaan korpus ini secara singkat dan *POS Tagset* yaitu daftar label kelas kata yang bisa digunakan.

Menu	Korpus Referensi TBIK v1.3: powered by CQPw																																									
Corpus queries	Restricted Query																																									
Standard query	<div style="border: 1px solid black; height: 80px; margin-bottom: 10px;"></div> <div> <div>Query mode:</div> <div>Simple query (ignore case) ▼</div> <div>Simple query language</div> </div> <div> <div>Number of hits per page:</div> <div>50 ▼</div> </div> <div> <div>Match strategy:</div> <div>Standard ▼</div> <div>Start query Reset query</div> </div>																																									
Restricted query																																										
Word lookup																																										
Frequency lists																																										
Keywords																																										
Analyse corpus																																										
Export corpus																																										
Saved query data				<div>Select the text-type restrictions for your query</div> <table border="1"> <thead> <tr> <th>sumber</th> <th>tahun</th> <th></th> </tr> </thead> <tbody> <tr><td><input type="checkbox"/> A_koran</td><td><input type="checkbox"/> 2011</td><td></td></tr> <tr><td><input type="checkbox"/> B_majalah</td><td><input type="checkbox"/> 2012</td><td></td></tr> <tr><td><input type="checkbox"/> C_cerpen</td><td><input type="checkbox"/> 2013</td><td></td></tr> <tr><td><input type="checkbox"/> D_novel</td><td><input type="checkbox"/> 2014</td><td></td></tr> <tr><td><input type="checkbox"/> E_buku_teks</td><td><input type="checkbox"/> 2015</td><td></td></tr> <tr><td><input type="checkbox"/> F_jurnal</td><td><input type="checkbox"/> 2016</td><td></td></tr> <tr><td><input type="checkbox"/> G_disertasi_tesis_skripsi</td><td><input type="checkbox"/> 2017</td><td></td></tr> <tr><td><input type="checkbox"/> H_biografi</td><td><input type="checkbox"/> 2018</td><td></td></tr> <tr><td><input type="checkbox"/> I_populer</td><td><input type="checkbox"/> 2019</td><td></td></tr> <tr><td><input type="checkbox"/> J_perundangan</td><td><input type="checkbox"/> 2020</td><td></td></tr> <tr><td><input type="checkbox"/> K_laman_resmi</td><td></td><td></td></tr> <tr><td><input type="checkbox"/> L_surat_resmi</td><td></td><td></td></tr> </tbody> </table>			sumber	tahun		<input type="checkbox"/> A_koran	<input type="checkbox"/> 2011		<input type="checkbox"/> B_majalah	<input type="checkbox"/> 2012		<input type="checkbox"/> C_cerpen	<input type="checkbox"/> 2013		<input type="checkbox"/> D_novel	<input type="checkbox"/> 2014		<input type="checkbox"/> E_buku_teks	<input type="checkbox"/> 2015		<input type="checkbox"/> F_jurnal	<input type="checkbox"/> 2016		<input type="checkbox"/> G_disertasi_tesis_skripsi	<input type="checkbox"/> 2017		<input type="checkbox"/> H_biografi	<input type="checkbox"/> 2018		<input type="checkbox"/> I_populer	<input type="checkbox"/> 2019		<input type="checkbox"/> J_perundangan	<input type="checkbox"/> 2020		<input type="checkbox"/> K_laman_resmi		
sumber	tahun																																									
<input type="checkbox"/> A_koran	<input type="checkbox"/> 2011																																									
<input type="checkbox"/> B_majalah	<input type="checkbox"/> 2012																																									
<input type="checkbox"/> C_cerpen	<input type="checkbox"/> 2013																																									
<input type="checkbox"/> D_novel	<input type="checkbox"/> 2014																																									
<input type="checkbox"/> E_buku_teks	<input type="checkbox"/> 2015																																									
<input type="checkbox"/> F_jurnal	<input type="checkbox"/> 2016																																									
<input type="checkbox"/> G_disertasi_tesis_skripsi	<input type="checkbox"/> 2017																																									
<input type="checkbox"/> H_biografi	<input type="checkbox"/> 2018																																									
<input type="checkbox"/> I_populer	<input type="checkbox"/> 2019																																									
<input type="checkbox"/> J_perundangan	<input type="checkbox"/> 2020																																									
<input type="checkbox"/> K_laman_resmi																																										
<input type="checkbox"/> L_surat_resmi																																										
Corpus info																																										
View corpus metadata																																										
Corpus manual ↗																																										
CS UI POS tagset ↗																																										
Indonesian TreeTagger ↗																																										
Admin tools																																										
Go to admin control panel																																										
Corpus settings																																										
Manage access																																										

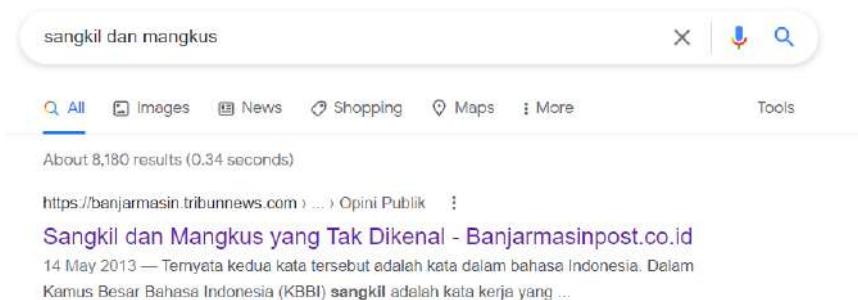
Gambar 12 Halaman depan korpus TBIK (*restricted search*)

Apakah korpus ini terbuka untuk diakses publik? Hingga saat tulisan ini dibuat, tidak. Korpus ini hanya bisa diakses oleh pengguna yang terdaftar sebagai kontributor buku TBIK. Akan tetapi, kita tidak tahu tentang kebijakan masa depan. Korpus BNC saja, yang pada awalnya adalah korpus berbayar, sekarang dapat diakses secara gratis di berbagai aplikasi web korpus seperti BNCweb, CQPweb, Sketch Engine, English-Corpora. Mudah-mudahan korpus ini juga nantinya demikian!

4. BEBERAPA MODEL ANALISIS TATA BAHASA BERBASIS KORPUS

4.1 Otentisitas

Salah satu kegunaan korpus adalah untuk memastikan bahwa analisis yang kita buat berdasarkan data otentik, yang benar-benar digunakan, dan bukan rekaan. Perhatikan hasil penelusuran Google berikut. Website yang paling atas menjelaskan bahwa kata *sangkil* dan *mangkus* tidak begitu dikenal, dan ekivalensinya yang sering digunakan adalah *efektif* dan *efisien*. Apakah ini benar? Mari kita lakukan penelusuran data korpus TBIK.



Gambar 13 Hasil penelusuran *sangkil* dan *mangkus* di Google

Hasil penelusuran TBIK menggunakan kueri *sangkil* dan *mangkus* menunjukkan bahwa ungkapan ini digunakan dalam korpus TBIK. Akan tetapi, frekuensinya sangat rendah. Dari keseluruhan data (dari hampir 30 juta token), ungkapan ini muncul hanya satu kali. Bandingkan dengan *efektif* dan *efisien* yang muncul sebanyak 349 kali pada 239 teks yang berbeda, seperti bisa kita lihat pada gambar di bawah ini.

larutan Dithane M-45 0,2 % dengan **sangkil dan mangkus** . Cara lain dengan menyemprotkan

Gambar 14 Hasil penelusuran *sangkil dan mangkus* di korpus TBIK

pemerintahan yang baik adalah yang **efektif dan efisien** . Ahli hukum tata negara dari Univers
a menanamkan gaya berenang yang **efektif dan efisien** kepada calon-calon atlet yang masih b
in untuk menciptakan suasana kerja **efektif dan efisien** serta menghindarkan dari jebakan-jeba

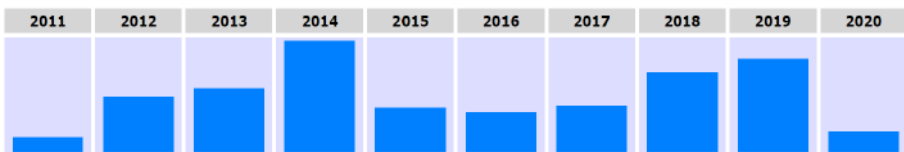
Gambar 15 Hasil penelusuran *efektif dan efisien* di korpus TBIK (3 hasil teratas)

Dengan hasil penelusuran seperti ini, kita bisa dengan aman mengatakan bahwa berdasarkan data korpus TBIK, ekspresi *efektif dan efisien* lebih sering digunakan pada bahasa Indonesia kontemporer.

4.2 Informasi Metadata

Pada bagian sebelumnya, kita telah melihat file sampel metadata yang menampilkkan data TBIK berdasarkan sumber domain maupun tahun terbit artikel. Keberadaan metadata ini akan sangat membantu dalam hal analisis data. Perhatikan hasil penelusuran menggunakan kueri KPU (singkatan dari Komite Pemilihan Umum) yang dianalisis menggunakan fitur *Distribution* di CQPweb.

Your query "linguistik" returned 560 matches in 79 texts]; frequency: 18.674 instances p



Gambar 16 Distribusi token *KPU* pada korpus TBIK (Restriksi: A_Koran)

Ada tiga interpretasi yang paling tidak bisa kita tawarkan dari hasil analisis ini. Frekuensi pemunculan *KPU* sangat tinggi pada tahun 2014 dan 2019. Hal ini sangat mungkin sesuai dengan kinerja KPU yang menjadi sorotan di tahun-tahun pemilu, dalam data ini di tahun 2014 dan 2019. Sorotan direfleksikan dalam korpus sehingga dapat muncul pada gambar di atas. Kedua, setelah pemilu, pemunculan KPU akan menurun drastis. Ketika, ada kecenderungan bahwa frekuensi pemunculan KPU akan semakin tinggi dengan semakin dekatnya pemilu berikutnya. Di Indonesia, Pemilu terjadi setiap 5 tahun sekali. Sehingga, terjadi peningkatan yang stabil dalam kurun waktu 5 tahun. Kita bisa memprediksi bahwa setelah 2020, akan terjadi peningkatan frekuensi hingga mencapai akan mencapai puncaknya pada tahun 2024, saat pemilu akan terjadi. Semua analisis ini dimungkinkan karena adanya metadata tahun terbit dalam korpus TBIK.

4.3 Frekuensi

Dalam sebuah analisis, kadang kita membandingkan pemunculan satu kata yang sama, dalam korpus yang berbeda. Tujuannya adalah membandingkan, di korpus mana kata tersebut muncul lebih banyak. Namun jika tidak hati-hati, maka kita bisa membuat klaim yang salah. Misalnya, pada korpus TBIK pemunculan token *linguistik* adalah sebesar 560 kali, sedangkan pada korpus LCC Indonesian adalah 778 kali. Frekuensi seperti ini disebut frekuensi absolut. Sekilas terlihat bahwa kata *linguistik* lebih banyak digunakan di korpus LCC Indonesian. Apakah ini kesimpulan yang tepat?

Mari sejenak kita beralih ke sebuah analogi. Misal, kita memberikan makanan sebanyak 300 kalori ke dua orang yang berbeda. Orang pertama hanya mampu menghabiskan 100 kalori, sedangkan orang kedua menghabiskan 200 kalori. Siapa yang paling banyak makan? Tentu kita akan menjawab orang kedua. Namun, kita mendapatkan informasi tambahan. Orang pertama berumur 4 tahun, sedang orang kedua berumur 24 tahun. Jika pertanyaan yang sama diajukan, siapa yang lebih banyak makan? Mungkin kita akan mengubah jawaban kita menjadi orang pertama.

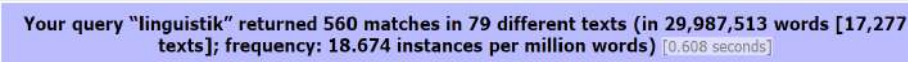
Kembali ke persoalan frekuensi tadi. Saat kita membandingkan frekuensi pemunculan pada dua korpus yang berbeda, maka kita harus terlebih dahulu menanyakan apakah perbedaan ukuran korpus tersebut terlalu jauh? Atau mirip? Jika mirip atau sama persis, tidak masalah. Tapi jika berbeda jauh,

maka kita sebaiknya menggunakan konsep frekuensi relatif, bukan absolut. Perhatikan dua gambar di bawah ini. Dalam CQPweb, satuan frekuensi absolut adalah *matches* sedangkan satuan frekuensi relatif adalah *instances per million words*.



Your query "linguistik" returned 778 matches in 14 different texts (in 232,853,664 words [14 texts]); frequency: 3.341 instances per million words [1.652 seconds]

Gambar 17 Frekuensi absolut dan relatif *linguistik* pada korpus LCC Indonesian



Your query "linguistik" returned 560 matches in 79 different texts (in 29,987,513 words [17,277 texts]); frequency: 18.674 instances per million words [0.608 seconds]

Gambar 18 Frekuensi absolut dan relatif *linguistik* pada korpus TBIK

Kita bisa melihat, meskipun frekuensi absolut *linguistik* lebih tinggi pada korpus LCC Indonesian (778) dibandingkan dengan korpus TBIK (560), frekuensi relatif untuk korpus LCC Indonesian lebih kecil (3,341) dibandingkan dengan korpus TBIK (18,674). Hal ini diakibatkan ukuran korpus yang berbeda. Ukuran korpus TBIK hampir mencapai 30 juta token, sedangkan ukuran korpus LCC Indonesian jauh lebih besar, yaitu sekitar 232 juta token!

Frekuensi relatif ini menunjukkan sealiknya, bahwa kata *linguistik* lebih tinggi pemunculannya di korpus TBIK dibandingkan dengan LCC Indonesian. Oleh karena itu, jika membandingkan data korpus atau subkorpus yang ukurannya jauh berbeda, gunakanlah frekuensi relatif.

4.4 Pola

Pendekatan linguistik korpus dapat digunakan untuk mengkonformasi intuisi mengenai pola konstruksi. Pada bagian ini kita akan membahas satu ilustrasi saja, yang akan kita awali dengan sebuah pertanyaan. Preposisi apa yang paling sering berada tepat disamping verba pasif berawalan *di-*? Secara intuitif, beberapa orang mungkin mengatakan *oleh* untuk menunjukan pelaku dari verba tersebut. Namun apakah ini berlaku untuk semua verba pasif tersebut? Hal ini bisa kita konformasi menggunakan pendekatan linguistik korpus. Yang pertama dilakukan adalah mendesain kueri. Kueri yang kita gunakan adalah *dipandang _IN*. Kueri ini memfokuskan pada preposisi yang tepat ada di samping verba pasif *dipandang*. Label *_IN* mengacu pada preposisi.

No.	Query result	No. of occurrences	Percent
1	dipandang sebagai	490	84.19%
2	dipandang dari	40	6.87%
3	dipandang oleh	21	3.61%

Gambar 19 Hasil Frequency breakdown *dipandang dari* korpus TBIK

Pada gambar di atas yang merupakan hasil penelusuran menggunakan teknik *Frequency breakdown* di korpus TBIK, kita bisa melihat bahwa ternyata, preposisi *oleh* hanya berada para urutan ketiga dengan proporsi 3,61%. Preposisi yang berada di posisi kedua, *dari*, juga tidak terpaut jauh, hanya 6,87%. Preposisi yang terbanyak digunakan adalah *sebagai*, dengan proporsi sebanyak 84,19%, sangat tinggi dan jauh berbeda dengan preposisi *dari*, *oleh* dan preposisi yang ada di urutan setelahnya. Kita dapat melakukan pemeriksaan tambahan dengan memasukkan kueri yang sama di korpus BI lain, dalam hal ini LCC Indonesian yang jauh lebih besar. Rangkingnya ternyata sama persis! Preposisi *sebagai* berada pada urutan pertama, disusul dengan *dari*, dan *oleh* ada urutan kedua dan ketiga. Pembahasan mengenai makna literal dan metaforis verba dengan kata dasar *pandang* dapat dilihat salah satunya di [22]¹⁷.

No.	Query result	No. of occurrences	Percent
1	dipandang sebagai	2715	76.59%
2	dipandang dari	434	12.24%
3	dipandang oleh	160	4.51%

Gambar 20 Hasil *frequency breakdown* dipandang dari korpus LCC Indonesian

4.4 Daftar Kata

Daftar kata memiliki banyak fungsi, baik dalam ranah linguistik deskriptif, pedagogis, leksikografi, maupun cabang ilmu linguistik lain. Misanya, daftar kata dapat digunakan untuk melakukan profiling secara cepat. Perhatikan dua daftar kata di bawah yang diambil dari korpus IndoPhone (korpus fonetik bahasa Indonesia dan daerah), yang membandingkan bahasa

17 <https://rinarxiv.lipi.go.id/lipi/preprint/download/638/686/586> (diakses tanggal 6 April 2023)

Jawa varian Semarangan dan Banyumasan berdasarkan frekuensi tertinggi (10 besar). Bisa kita lihat kosakata khas setiap ragam yang masuk dalam 10 besar seperti [həʔə] pada ragam Semarangan dan [kuwe] pada ragam Banyumasan.

Tabel 5 Sampel file metadata korpus TBIK dengan informasi tahun dan domain teks

Rangking/ Bahasa	Semarangn	Banyumasan
1	ya	ya
2	kan	apa
3	tu	kuwe
4	ngaʔ	ora
5	ndaʔ	siŋ
6	tə	si
7	yaŋ	nəŋ
8	həʔə	kae
9	itu	wis
10	ði	iya

Untuk TBIK, perbandingan daftar kata antar domain dapat menunjukkan kekhasan domain masing-masing. Perhatikan perbedaan 10 besar kata dengan frekuensi tertinggi pada domain cerpen dan perundangan. Pada domain cerpen, muncul pronomina *aku* dan bentuk kasual dari *tidak*, yaitu *tak*. Dua kata ini tidak muncul di 10 besar kosakata tersering di domain perundangan. Yang muncul adalah kosakata domain hukum seperti *pasal* dan *ayat*.

Tabel 6 Sampel file metadata korpus TBIK dengan informasi tahun dan domain teks

Rangking/ Domain	Cerpen	Perundangan
1	yang	dan
2	di	yang
3	dan	pasal
4	itu	dalam
5	aku	dimaksud
6	ia	ayat
7	tak	sebagaimana
8	dengan	pada
9	tidak	dengan
10	ke	daerah

5. PERSPEKTIF

5.1 Tata Bahasa Prekriptif dan Deskriptif

Akan dibahas dua sudut pandangan tata bahasa, yaitu preskriptif dan deksriptif. Sebagai pembuka, mari kita simak ilustrasi berikut. Manakah bentuk yang baku, *praktik*, atau *praktek*? Pertanyaan seperti ini bersifat preskriptif. Pendekatan preskriptif seperti resep obat (dalam bahasa Inggris *medical prescription* atau *prescription* saja) dari dokter yang harus diikuti oleh apoteker, baik dari jenis obat, maupun takarannya. Kembali ke pertanyaan tadi, jawaban preskriptifnya adalah *praktik* bukan *praktek*. Ini bisa kita lihat di KBBI versi online. Penjelasan untuk entri *praktik* adalah sebagai berikut.

prak.tik

bentuk tidak baku: **praktek**

→ Taurus

1. *n* pelaksanaan secara nyata apa yang disebut dalam teori: *teorinya mudah*
2. *n* pelaksanaan pekerjaan (tentang dokter, pengacara, dan sebagainya): --
3. *n* perbuatan menerapkan teori (keyakinan dan sebagainya); pelaksanaan:

Gambar 21 Praktik¹⁸ dalam KBBI online

Bagaimana dengan tata bahasa deksriptif? Tata bahasa deskriptif cenderung lebih cair. Misalnya, ketimbang memilih mana yang baku atau tidak, maka tata bahasa deskriptif bersifat menjelaskan. Misalnya, apakah keduanya digunakan? Jika iya, mana yang lebih sering digunakan dan dalam konteks apa?

Manakah yang sebaiknya digunakan? Pendekatan preskriptif atau deskriptif? Tentu saja ini bergantung dari tujuannya. Misalnya, untuk tujuan pedagogis pendidikan bahasa pada level sekolah dasar atau menengah, pendekatan preskriptif lebih mudah digunakan karena memiliki panduan yang sangat jelas. Namun untuk penelitian bahasa yang sifatnya saintifik dan tidak sekedar mengukur benar dan salah, misalnya skripsi, tesis, atau disertasi, penggunaan pendekatan deksriptif akan sangat berguna sesuai dengan level kompetensi pemikiran yang sudah dikuasai. Mereka akan diminta mengklasifikasikan, menganalisis, mensintesis dan menginterpretasikan

¹⁸ <https://kbbi.kemdikbud.go.id/entri/praktik> (diakses tanggal 6 April 2023)

temuan. Dengan level abstraksi seperti ini, cocok untuk peneliti dewasa dibandingkan anak sekolah dasar atau menengah.

5.2 Kombinasi Teknik Analisis

Pendekatan linguistik korpus banyak diasumsikan merupakan sebuah pendekatan yang cenderung kuantitatif. Inilah yang perlu diluruskan. Pendekatan linguistik korpus tidaklah bersifat eksklusif. Dia bersifat inklusif. Sehingga, baik penelitian yang sifatnya kuantitatif maupun kualitatif sama berharganya, dan dapat digunakan di linguistik korpus sesuai dengan fungsi dan porsinya. Buktinya dapat dilihat pada bagian-bagian sebelumnya. Misalnya, pada bagian 4.3, kita membahas tentang frekuensi yang cenderung menggunakan pendekatan kuantitatif. Namun pada bagian 3.1 tentang metadata, kita mendiskusikan sebuah ilustrasi yang menunjukkan bahwa data kuantitatif perlu diinterpretasikan menggunakan pendekatan kualitatif.

Selain sifat kuantitatif-kualitatif, banyak riset linguistik yang mengkombinasikan pendekatan linguistik korpus dengan pendekatan atau teknik analisis dari sub-disiplin linguistik lain, misalnya sosiolinguistik (misalnya [23] [24]), pragmatik (seperti [25] [26]), analisis wacana (misalnya [27], [28]). Oleh sebab itu, kita tidak perlu menganggap kehadiran linguistik korpus sebagai satu pendekatan yang kompetitif. Linguistik korpus bukanlah ancaman, karena keberadaan linguistik korpus tidak menegasikan pendekatan-pendekatan linguistik lain yang lebih dulu ada, maupun yang nantinya akan muncul.

Referensi

- [1] J N Sneddon, A Adelaar, D-N Djenar, and M-C Ewing, *Indonesian Reference Grammar: 2nd Edition*. New South Wales: Allen & Unwin, 2010.
- [2] H Alwi, S Dardjowidjojo, H Lapoliwa, and M Moeliono, *Tata Bahasa Baku Bahasa Indonesia (4th Edition)*. Jakarta: Badan Pengembangan dan Pembinaan Bahasa, 2017.
- [3] Prihantoro, *Buku referensi pengantar linguistik korpus: lensa digital data bahasa*. Semarang: Undip Press, 2022.
- [4] P Baker, "Corpus methods in linguistics," in *Research methods in linguistics*. London & New York: Continuum, 2010, pp. 93-113.

- [5] Tony Mc Enery and Andrew Hardie, *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press, 2012.
- [6] T McEnery, R Xiao, and Y Tono, *Corpus-based language studies: An advanced resource book*. Milton Park: Taylor & Francis, 2006.
- [7] H Schimdt, "Probabilistic Part-of-Speech Tagging Using Decision Trees," in *International Conference on New Methods in Language Processing*, Manchester, UK, 1994.
- [8] F Rashel, A Dinakaraman, and A Luthfy, *Building an Indonesian Rule-Based Part-of-Speech Tagger*, Minghui Dong et al., Eds. Sarawak: IALP, 2014, pp. 70-73.
- [9] S-D Larasati, Vladislav Kuboň, and Daniel Zeman, "Indonesian Morphology Tool (MorphInd): Towards an Indonesian Corpus," in *Systems and Frameworks for Computational Morphology*, Zurich, 2011, pp. 119-129.
- [10] Prihantoro, *SANTI-morf: A new morphological annotation system for Indonesian (PhD Thesis)*. Lancaster: Lancaster University, 2021.
- [11] F Pisceldo, R Mahendra, R Manurung, and I W Arka, "A Two Level Morphological Analyser for the Indonesian Language," in *Proceedings of Australasia Technology Association Workshop*, Nicola Stokes and David Powers, Eds. Hobart: ACL, 2008, pp. 142-150.
- [12] D Moeljadi, F Bond, and S Song, "Building an HPSG-based indonesian resource grammar (INDRA)," Singapore, 2015.
- [13] A Hardie, "CQPweb—combining power, flexibility and usability in a corpus analysis tool," *International journal of corpus linguistics*, 17(3), pp. 380-409, 2012.
- [14] A Kilgariff et al., "The Sketch Engine: ten years on," *Lexicography* (1) 1, pp. 7-36, 2014.
- [15] M Scott, *WordSmith manual v6*. Gloucestershire: Lexical Analysis Software ltd, 2015.
- [16] V Brezina, M Timperley, and T McEnery, *#LancsBox v. 4.x [software]*. Available at: <http://corpora.lancs.ac.uk/lancsbox>, 2018.
- [17] L. Anthony, "AntConc (Version 4.2.0)," Tokyo, Japan, 2022, [Computer Software]. Available: <https://www.laurenceanthony.net/software>.
- [18] L Burnard, *Reference Guide for the British National Corpus (XML Edition)*.: URL: <http://www.natcorp.ox.ac.uk/XMLedition/URG/>, 2007.
- [19] R Love, C Dembry, A Hardie, V Brezina, and T McEnery, "The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations," *International Journal of Corpus Linguistics*, 22(3), pp. 319-344, 2017.

- [20] A-F Wicaksono and A Purwarianti, "HMM Based Part-of-Speech Tagger for Bahasa Indonesia," in *Proceeding of the Fourth International MALINDO Workshop (MALINDO2010)*, Jakarta, 2010.
- [21] Mirna Adriani and Riza Hamam, "Research Report Phase 3.2: Final Report on Statistical Machine Translation for Bahasa Indonesia - English and English to Bahasa Indonesia," Jakarta, 2009.
- [22] G-P-W Rajeg, *Kajian konkordansi korpus terhadap perilaku konstruksional makna literal dan metaforis pasangan verba sinonim pandang dan tatap*. RINarxiv, Aug. 26, 2022. Accessed: Sep. 25, 2022. [Online]. Available: <https://rinarxiv.lipi.go.id/lipi/preprint/view/638, 2022>.
- [23] Ryan Howe, *The Use of Fuck: A Sociolinguistic Approach to the Usage of Fuck in the BNC and Blog Authorship Corpus*. Michigan: Michigan University Press, 2012.
- [24] V Brezina, R Love, and K Aijmer, *Corpus approaches to contemporary British speech: Sociolinguistic studies of the Spoken BNC2014*. Oxford and Philadelphia: Routledge, 2018.
- [25] K Aijmer and C Rühlemann, *Corpus pragmatics*. Cambridge: Cambridge University Press, 2017.
- [26] M Weisser and C Ruhlemann, "Speech act annotation," in *Corpus pragmatics: a handbook*. Cambridge: Cambridge university press, 2014, pp. 84-114.
- [27] P Baker, C Gabrielatos, and T McEnery, "Sketching Muslims: A corpus driven analysis of representations around the word 'Muslim' in the British press 1998–2009," *Applied linguistics*, 34(3), pp. 255-278, 2013.
- [28] S Jaworska and R Krishnamurthy, "On the F word: A corpus-based analysis of the media representation of feminism in British and German press discourse, 1990–2009," *Discourse & Society* 23(4), pp. 401-431, 2012.

AFIKSASI VERBA DALAM BAHASA INDONESIA

¹Gede Primahadi Wijaya Rajeg dan ²Karlina Denistia

¹Universitas Udayana; ²Universitas Sebelas Maret

¹primahadi_wijaya@unud.ac.id; ²karlinadenistia@staff.uns.ac.id

Abstrak

Afiksasi verba bahasa Indonesia (BI) telah banyak dibahas dalam setiap buku tata bahasa BI dan literatur-literatur teoretis. Akan tetapi, sejauh ini, pembahasan distribusi kuantitatif afiksasi verba BI masih minim. Makalah ini membawa kebaruan dengan menampilkan deskripsi bentuk, fungsi dan makna afiksasi verba berdasarkan bank data teks digital melimpah (korpus bahasa) dari berbagai ragam teks dalam sepuluh tahun terakhir. Data korpus digali untuk menampilkan contoh-contoh bentukan kata yang tidak dibahas sebelumnya, sehingga memperkaya deskripsi afiksasi tersebut. Selanjutnya, distribusi kuantitatif afiksasi verba yang diulas adalah tingkat produktifitasnya (i) secara menyeluruh, (ii) berdasarkan dimensi ragam teks, dan (iii) secara diakronis dalam rentang sepuluh tahun terakhir. Secara kuantitatif, produktivitas afiksasi verba menunjukkan variasi berdasarkan ragam teks. Misalnya, jumlah kata unik (frekuensi tipe) dan kata yang ditemukan sekali dalam korpus (hapax) secara relatif sangat rendah pada ragam formal normatif seperti Perundang-Undangan dan Surat Resmi, namun sangat tinggi pada Cerpen, Novel, dan Koran. Dalam rentang sepuluh tahun terakhir (2011-2020), afiksasi verba utama BI (ME-, DI-, BER-, dan TER-) mengalami tren penurunan dalam hal frekuensi tipe dan jumlah hapax-nya.

Kata kunci: morfologi, verba, produktivitas morfologis, linguistik korpus kuantitatif

1. PENDAHULUAN

Makalah ini menjelaskan tipe-tipe, struktur, fungsi dan makna imbuhan/afiksasi pembentukan kata kerja (atau verba) dalam bahasa Indonesia (BI) (§4.1). Deskripsi afiksasi verba telah banyak dijelaskan sebelumnya, baik dalam bentuk buku tata bahasa BI (mis. [1], [2]) ataupun yang bersifat teoritis (mis. [3], [4]). Salah satu kebaruan makalah kali ini daripada deskripsi sebelumnya adalah sumber data yang dilandasi atas bank data kebahasaan BI digital (korpus) masif kontemporer dengan ukuran ± 18 juta kemunculan

kata¹. Korpus tersebut dibangun dari dua belas (12) ragam teks tulis² BI dari tahun 2011-2020. Penggunaan korpus masif beragam tersebut memungkinkan pemerolehan contoh-contoh pemakaian imbuhan berdasarkan berbagai ragam pemakaian bahasa alamiah (bukan data rekaan/rekayasa dari peneliti). Keragaman data pemakaian bahasa alamiah dapat memberikan penyegaran terhadap topik klasik tata bahasa BI terkait pembentukan verba.

Kebaruan berikutnya, yang tidak banyak mendapat perhatian sebelumnya, merupakan implikasi dari pemanfaatan data masif, yaitu keberadaan data kuantitatif pemakaian bahasa. Dalam konteks kajian morfologi terkait pembentukan kata, data kuantitatif dapat diolah untuk membandingkan produktivitas tiap-tiap afiksasi (§4.2), yang sekaligus dapat memberikan pengukuran terhadap asumsi kualitatif terkait keutamaan/sentralitas afiksasi verba BI tertentu. Selain itu, produktivitas dan distribusi tiap-tiap afiks dapat dibandingkan berdasarkan ragam teks (*genre*) korpusnya (§4.3) dan tahun (§4.4). Misalnya, kita dapat membandingkan (i) ragam teks mana yang memiliki produktivitas lebih tinggi untuk tipe afiksasi tertentu, (ii) tipe afiksasi mana yang lebih produktif di satu ragam teks dibandingkan ragam teks lainnya, atau (iii) seberapa stabil produktivitas afiksasi verba dalam sepuluh tahun terakhir. Singkat kata, kedua variabel tersebut memperkaya deskripsi pemakaian afiksasi verba secara lebih terukur.

2. KAJIAN PUSTAKA

Penelitian morfologi BI berbasis data korpus dan analisis kuantitatif mulai muncul setakat ini. Salah satu topik pembahasannya adalah produktivitas afiksasi secara umum dan berdasarkan sifat-sifat akarnya [5]–[8]. Beberapa kajian lainnya menemukan bahwa kecenderungan makna yang diungkapkan suatu akar verba yang sama bisa berbeda (i) pada bentuk kalimat aktif *meN-* dan pasif *di-* [9]–[11] dan (ii) ketika muncul dengan akhiran yang berbeda

1 Ukuran korpus ini dihitung berdasarkan kata yang terdiri atas paling sedikit dua huruf (mis. *di*), mengesampingkan bilangan dan tanda baca, kecuali “-” untuk menjaga bentuk reduplikasi (mis. *anak-anak*).

2 Ragam teks tersebut adalah Biografi, Buku Teks, Cerpen, Disertasi/Tesis/Skripsi, Jurnal, Koran, Laman Resmi, Majalah, Novel, Perundang-undangan, Populer, dan Surat Resmi. Korpus yang dinamai Korpus Referensi Tata Bahasa Indonesia Kontemporer (TBIK) ini dikumpulkan oleh Pusat Pengembangan dan Pelindungan Bahasa dan Sastra, Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

(meskipun secara teoritis arti pasangan verba tersebut diasumsikan mirip) [12]–[14]. Terakhir, terdapat kajian yang melihat perbedaan serta gugusan semantis sebagian tipe afiksasi (verba dan nomina) berdasarkan model-model komputasional mutakhir [15]–[17]. Dibandingkan dengan kajian terdahulu, makalah kali ini meneliti distribusi dan produktivitas afiksasi verba secara lebih komprehensif dengan mengikutkan dimensi ragam teks dan tahun.

3. METODOLOGI

Makalah ini memadukan pendekatan komputasional, kualitatif, dan kuantitatif. Secara komputasional, peranti pemrograman R (v4.1.3) [18] (dan modul R *tidyverse* [19]) digunakan untuk menjalankan *MorphInd* [20] terhadap daftar frekuensi kata dalam korpus. Tujuannya untuk menemukan imbuhan, bentuk dasar/akar kata beserta kelas katanya (lihat Tabel 1).

Tabel 1 Nukilan pangkalan data verba dan luaran analisis oleh *MorphInd*

Tahun	Kata	Frekuensi Kata	Genre	Morphind	Akar	Kelas Kata Akar
2018	<i>berkelamin</i>	2	Biografi	ber+kelamin<n>_VSA	kelamin	n
2017	<i>diterbitkan</i>	19	Majalah	di+terbit<v>+kan_VSP	terbit	v
2015	<i>merobek</i>	3	Cerpen	meN+robek<a>_VSA	robek	a
2012	<i>termakan</i>	5	Populer	ter+makan<v>_VSP	makan	v

R juga digunakan untuk pengolahan data korpus, analisis statistik, dan visualisasi. Selain itu, data pada kamus morfologi *MALINDO Morph* [21] diintegrasikan untuk kata-kata yang tidak dikenali oleh *MorphInd*. Pendekatan kualitatif meliputi dua hal. Pertama, identifikasi serta perbaikan manual (berdasarkan KBBI daring) sejumlah hasil analisis komputasional sebelumnya, seperti penentuan kelas kata akar (yang tidak diberikan oleh *MALINDO Morph*), ataupun analisis morfologis untuk kata yang tidak dikenali oleh *MALINDO Morph* dan *MorphInd*. Kedua, analisis kualitatif digunakan untuk pemaknaan bentuk kata berdasarkan konteks pemakaian kata tersebut dalam korpus. Selanjutnya, analisis kuantitatif digunakan untuk mengukur tingkat produktivitas [5]–[7], [22] (subtipe) afiksasi verba (i) pada keseluruhan korpus (§4.2), dan (ii) berdasarkan ragam teks (§4.3) dan tahun

(§4.4). Data dan kode R untuk analisis statistik dan visualisasi data tersedia pada <https://osf.io/nuxd4/>.

4. HASIL DAN PEMBAHASAN

Bagian 4.1 mengulas afiksasi pembentuk dua jenis kata kerja utama, yaitu verba taktransitif dan transitif. Verba taktransitif adalah verba yang pada hakikatnya tidak memerlukan objek untuk bermakna dalam kalimat (§4.1.1.1, §4.1.2). Sedangkan verba transitif adalah verba yang mewajibkan peran objek untuk bermakna dalam kalimat (§4.1.1.2) dan biasanya memiliki varian dalam kalimat pasif. Ulasan kuantitatif terkait produktivitas afiksasi verba dan distribusinya berdasarkan ragam dan tahun dimulai dari §4.2 hingga §4.4.

4.1 Afiksasi Verba dalam Bahasa Indonesia

Bagian ini mengulas pembentukan verba dengan awalan (prefiks) (§4.1.1-§4.1.3), akhiran (sufiks) (§4.1.4-§4.1.5), dan penggabungan antara awalan dan akhiran (konfiks) (§4.1.6).

4.1.1 Verba dengan awalan *meN-*

Bagian ini menjelaskan afiksasi verba *meN-* yang dapat membentuk verba taktransitif (§4.1.1.1) dan verba transitif (§4.1.1.2).

4.1.1.1 Verba taktransitif dengan *meN-*

Awalan *meN-* dapat berfungsi untuk menghasilkan verba taktransitif ketika bergabung dengan akar verba (mis. *menancap*, *menyerah*, *meletus*, *menawan*, *meloncat*, *menurun*). Ketika melekat pada akar verba taktransitif, *meN-* tidak memiliki fungsi khusus selain untuk menghasilkan bentuk verba yang dapat digunakan secara gramatikal dalam kalimat [1, p. 69].

- (1) Tiang-tiang itu *menancap* di kedalaman 23 meter. (Majalah)
- (2) Terdengar suara gelembung air *meletus*. (Novel)

Selain berasal dari akar verba, verba taktransitif juga dapat dihasilkan dengan melekatkan *meN-* pada akar kata benda (nomina), dan memiliki tiga makna utama. Makna pertama adalah ‘menuju [akar]’ (mis. *membumi*, *mendarat*, *mengangkasa*, *mendunia*, *melaut*, *melangit*).

- (3) ...satelit buatan yang *mengangkasa* pada daerah orbit geostasioner...
(Laman_resmi)

Makna kedua dari verba taktransitif berakar nomina adalah ‘menghasilkan/mengeluarkan [akar]’ (mis. *merona*, *menggema*, *menderu*, *mengembik*, *mengeong*, *menggerung*, *mendesah*, *mendesis*, *mendengkur*). Perhatikan contoh berikut untuk *mengembik*.

- (4) Kambing pun *mengembik* dengan anehnya. (Cerpen)

Makna ketiga verba taktransitif berakar nomina adalah ‘menjadi/ menyerupai AKAR’ (mis. *melegenda*, *melengkung*, *menjanda*, *mengorbit*, *mengakar*, *menyemut*, *meroket*, *merakyat*).

- (5) Ada beberapa nama penting dan *melegenda* dalam ajang Asian Games... (Biografi)

Selanjutnya, kata sifat (adjektiva) juga dapat digunakan dengan *meN-* untuk membentuk verba taktransitif. Makna yang diungkapkan adalah ‘(berubah) menjadi [akar]’/‘bersifat seperti [akar]’. Beberapa contoh katanya adalah *mengglobal*, *memudar*, *menghijau*, *menyempit*, *merata*, *merindu*, *mengiba*, *menegang*, *mengencang*, *membengkok*, *menua*.

- (6) Malam ini dadaku menjadi sesak *merindu* (Cerpen)

Kelompok akar kata lain yang dapat menghasilkan verba taktransitif dengan *meN-* adalah kata bilangan (numeralia) (mis. *mendua*) (lihat (7)), kata keterangan (adverbia) (*menggeletak*) (8), kata ganti orang/persona (pronomina) (*mengaku*) (9), ataupun kata interjeksi (*mengaduh*) (10). Kalimat-kalimat berikut mencontohkan penggunaan verba turunan tersebut.

- (7) ...Cruyff masih *mendua* apakah menekuni sepak bola atau bisbol
(Majalah)
- (8) ...mengguncang bahu Sunu dan Alex yang *menggeletak* seperti sapi.
(Novel)

- (9) “Google Indonesia *mengaku* tidak memiliki dokumen yang kami minta.” (Majalah)
- (10) ...jika sakit gigi sedikit saja *mengaduh* siang malam... (Populer)

Terakhir, *meN-* juga dapat menandai verba taktransitif yang berasal dari akar kata yang terikat, dalam arti akar kata tersebut belum bisa berfungsi sebagai verba utama dalam kalimat jika tidak muncul dengan *meN-*. Perhatikan contoh-contoh berikut.

- (11) ...memencet Safari di layar sentuh, kemudian *meramban*. (Cerpen)
- (12) Tangan Lee *mengacung* keluar dari jendela mobil... (Novel)
- (13) ...air matanya menetes, *menggelincir* ke pipi. (Cerpen)

Bentuk akar *ramban* (11) dan *gelincir* (13) tanpa imbuhan tidak ditemukan dalam korpus, sedangkan *acung* (12) muncul tidak sebagai verba namun sebagai penjelas bagi nomina (*pe*)dagang (pada [*pe*]dagang *acung*) dan verba *berdagang* (*berdagang acung*). Jadi, awalan *meN-* membuat akar terikat tersebut berfungsi secara gramatikal sebagai verba dalam kalimat.

4.1.1.2 Verba transitif dengan *meN-*

Dalam konteks verba transitif, awalan *meN-* dapat ditambahkan pada akar verba yang sudah berupa verba transitif (yaitu verba yang pada hakikatnya memerlukan objek) (mis. *beli* → *membeli*; *lihat* → *melihat*; *cari* → *mencari*; *siram* → *menyiram*). Pada kasus ini, *meN-* tidak berfungsi untuk mengubah kelas kata akarnya (seperti pada §4.1.1.1) melainkan menjadi penanda penggunaan verba tersebut pada kalimat aktif. Bentuk pasif dari verba transitif tersebut ditandai dengan mengganti awalan *meN-* dengan awalan *di-* (atau *ter-*).

- (14) ...Qusai Abtini (...) sukses *merebut* hati pemirsa (Koran) (Aktif)
- (15) ...setelah kemerdekaan berhasil *direbut* dari tangan penjajah (Cerpen) (Pasif *di-*)

Selanjutnya, *meN-* dapat menurunkan verba transitif dari akar nomina serapan asing. Beberapa contohnya adalah *mendominasi*, *mensintesis*, *mereduksi*, *mengonsumsi*, *menganalisis*, *memotivasi*, *mengestimasi*, *mengisolasi*. Dalam hal ini *meN-* juga sekaligus menjadi penanda bahwa verba

tersebut digunakan dalam kalimat aktif (bdk. *direduksi*, *dianalisis*, *dikonsumsi* yang merupakan bentuk pasif verba aktifnya).

- (16) Penelitian ini bertujuan untuk *mensintesis* senyawa baru. (Jurnal)
- (17) Biotin di dalam usus besar dapat *disintesis* oleh bakteri... (Buku_Teks)

Pola di atas dimungkinkan mengingat dalam bahasa Inggris nomina serapan tersebut dilandasi atas verba transitif. Misalnya, *motivate*_{Vtrans} → *motivation*_N → *motivasi*_N; *synthesise*_{Vtrans} → *synthesis*_N → *sintesis*_N; *reduce*_{Vtrans} → *reduction*_N → *reduksi*_N. Pola serupa juga tercermin pada akar nomina bukan serapan lain (mis. *lecute*, *perban*, *borgol*) yang dapat digunakan sebagai verba transitif dengan *meN-* (aktif) ataupun *di-* (atau *ter-*) (pasif).

- (18) Hujan yang kerap membuatmu termangu, *melecut* keterjagaanmu. (Cerpen)
- (19) Jiwa kejantanannya sebagai lelaki seperti sedang *dilecut* oleh Ayah. (Biografi)
- (20) Kondisi itulah yang membuatnya *terlecut* untuk keluar dari profesinya... (Populer)

Dari segi makna, verba transitif *meN-* berakar nomina bermakna generik ‘melakukan sesuatu, berkaitan dengan [akar], terhadap objek’ (mis. *menganalisis X* ‘melakukan analisis terhadap X’). Untuk kelompok nomina tertentu yang merujuk pada alat, makna lebih khususnya dapat berupa ‘menggunakan [akar] terhadap objek’ (mis. *melecut X* ‘menggunakan lecut terhadap X’) (lihat §4.1.3-§4.1.5 untuk pembentukan verba transitif dengan imbuhan lainnya).

4.1.2 Verba dengan awalan ber-

Verba berawalan *ber-* bersifat taktransitif, yaitu tidak memerlukan objek (nomina) (lihat §4.1.6 untuk kombinasi *ber-* dan akhiran *-kan* guna menghasilkan verba transitif). Verba taktransitif *ber-* dapat berasal dari akar verba dan secara umum memiliki makna ‘sedang melakukan aktivitas, perbuatan, atau dalam suatu proses yang dinyatakan oleh [akar]’ (bdk. [2, p. 147]). Beberapa contohnya adalah *bergaul*, *bersembunyi*, *berlari*, *bertumbuh*,

berpaling, berlimpah, bergeser, terbuka, bertanya, berlatih, bertiup, bersantap, dan lainnya.

- (21) Si pipi gembil *bersantap* lahap. (Cerpen)

4.1.2.1 Verba taktransitif *ber-* dengan akar nomina

Mayoritas bentukan verba taktransitif *ber-* dilandasi atas akar nomina (Gambar 5), dan mengungkapkan sejumlah gugusan makna. Salah satunya yang bersifat generik adalah ‘melakukan sesuatu berkaitan dengan [akar]’ (mis. *berswafoto, berinstagram, beribadat, berhaji, berjihad, bergerilya, berdiplomasi, berdiaspora, berdemo, berdangdut, beraktifitas*).

- (22) ...mendampingi putra-putri dalam *berinstagram* ria. (Koran)
(23) ...*berdiaspora* menjadi anak-anak panah yang memajukan persyarikatan... (Biografi)

Kegiatan yang dilakukan di atas juga bisa bersifat spesifik, seperti merujuk pada kegiatan olahraga, misalnya *beraerobik, berakrobat, bertinju, bergulat, berolahraga*.

Berikutnya, makna yang sedikit lebih mengkhusus untuk verba taktransitif dengan *ber-* adalah ‘memiliki [akar]’. Misalnya *bertampang, bervalensi, berserver, ber-NPWP, ber-KTP, berijazah, berprasangka, berhasrat*.

- (24) ...keluarga miskin (...) *ber-KTP* wilayah Kota Yogyakarta. (Cerpen)

Makna selanjutnya merujuk pada medan makna yang lebih spesifik, yaitu ‘menggunakan [akar] sebagai transportasi’, yang dicerminkan oleh kata *berlayar, berkendara, berkereta, berperahu, berkuda, berunta, bersepeda, bermobil, berbahtera, berkomuter* dan lainnya.

- (25) Rumah Tom hanya 10 menit *bermobil* dari rumah saya. (Cerpen)
(26) ...dengan membawa 3000 pasukan *berunta*, 200 pasukan *berkuda*... (Buku_Teks)

Kemudian, *ber-* dengan akar nomina bertipe busana atau aksesoris memicu makna ‘mengenakan [akar]’. Terdapat banyak contoh untuk tipe

makna ini, beberapa di antaranya adalah *berserban*, *berseragam*, *berhijab*, *berdasi*, *berjaket*, *berbatik*, *berbusana*, *berpiama*.

(27) ...figur lelaki *berkalung* salib... (Koran)

Makna lainnya dari *ber-* dengan akar nomina adalah ‘menghasilkan/mengeluarkan [akar]’. Contohnya adalah *berbuah*, *berbau*, *berpijar*, *berdesir*, *bergema*, *berkelakar*, *berfatwa*.

(28) Lagu yang sama juga *bergema* di pasar valuta asing... (Majalah)

(29) ...layaknya ulama besar yang sedang *berfatwa*. (Biografi)

Terdapat sekelompok verba taktransitif *ber-* yang mengungkapkan makna ‘hubungan timbal balik’ [2, p. 151] (lihat juga §4.1.6). Contohnya adalah *bermitra*, *berkawan*, *berpartner*, *berteman*, *berkoalisi*, *berkomplot*, *berkolaborator*, *bertetangga*, *berkerabat*, *bermasyarakat*.

(30) ...maka peneliti yang *berkolaborator* dengan guru... (Disertasi_ Tesis_ Skripsi)

Makna lain untuk beberapa kelompok verba taktransitif dengan *ber-* adalah aktivitas terkait mata pencaharian atau pekerjaan, seperti *berbisnis*, *berakting*, *berwirausaha*, *berdagang*, *berladang*, *bercocok tanam*, *beternak*, *berfotografi*, *bergiat*, *berkuliah*.

(31) ...ia menyesal karena selama ini *bergiat* dalam politik... (Biografi)

(32) ...jalannya benar-benar mantap dari masa ke masa untuk *berfotografi*. (Biografi)

Terakhir, awalan *ber-* juga dapat menandai tindakan yang dilakukan oleh, dan diarahkan sendiri pada, subjek (diistilahkan dengan “refleksif”) [1, p. 66]. Contohnya adalah *berbedak*, *berpupur*, *bercermin*, *berkaca*, *bersuluh*, *bersisir*, dan *bercukur* (akar verba *cukur*).

(33) Dia *berkaca bersisir*. Semua dilakukannya dengan sangat tenang... (Novel)

4.1.2.2 Verba taktransitif *ber-* dengan akar adjektiva

Verba taktransitif *ber-* berakar adjektiva tidak seproduktif verba sejenis dengan akar verba ataupun nomina (Gambar 5). Contohnya di antaranya *bergegas*, *berpasrah*, *bersepakat*, *berkompeten*, *bergembira*, *berbahagia*, *bersabar*, *bersiap*, *berkeras*, *berkongkalikong*.

- (34) Keesokannya Sabari *berkongkalikong* dengan tukang parkir...
(Novel)
- (35) ...Kasimo dan teman-teman *berkeras* mempertahankan Pancasila.
(Biografi)

Kelompok verba taktransitif ini umumnya bermakna ‘dalam keadaan [akar]’ [2, p. 149].

4.1.3 Verba dengan awalan *per-*

Verba transitif juga dapat dibentuk dengan awalan *per-* yang mengungkapkan arti ‘kausatif’, yaitu ‘menyebabkan peningkatan sifat/kualitas [akar] yang dimiliki objek’ [1, p. 103], [2, p. 135], [5, p. 289]. Makna tersebut khususnya dipicu ketika *per-* melekat pada akar adjektiva (kata sifat). Sebagai contoh, *perketat X* (dari akar kata *ketat*) berarti ‘membuat X menjadi/memiliki sifat lebih ketat’. Contoh lain diantaranya *perlancar*, *perlebar*, *permurah*, *pertajam*, *perkokoh*, *pertinggi*, *perberat* (lihat [5, pp. 314–317] untuk daftar lengkapnya).

- (36) Teorinya, fenomena ini akan *mempermurah* ekspor dan *mempermahal* impor. (Koran)
- (37) Riset unggulan *dipertajam* dengan pendekatan (...) trans-disiplin.
(Laman_Resmi)

Verba transitif *per-* dengan akar adjektiva paling produktif ditemukan dalam korpus. Dari total 99 bentukan verba transitif *per-* (dalam bentuk aktif *memper-* dan pasif *diper-*), 70,7%-nya (n=70) berakar adjektiva. Kemudian, *per-* juga bisa melekat dengan akar nomina (15,2%; n=15), seperti *peralat*, *perdaya*, *pertuan*, *perhamba*, *perbudak*, *perkuda*, *peristri*, *persengketa*, *perantai*. Berikut ini contoh penggunaannya dalam korpus.

- (38) sinkronisasi program pemerintah dan pemerintah daerah yang *diperantai* oleh Peraturan Menteri Dalam Negeri... (Koran)
- (39) ...mereka telah membiarkan diri *diperkuda* oleh kaum politisi... (Biografi)

Secara umum, makna verba transitif *per-* berakar nomina di atas adalah ‘menjadikan objek sebagai [akar]’. Misalnya, *perkuda/perbudak X* berarti ‘menjadikan X sebagai kuda/budak’.

Ditemukan sedikit verba transitif *per-* berakar verba (6%; n=6 dari total 99 bentukan kata). Bentuk tersebut adalah *perbuat*, *perolok*, *perhambat*, *perkenal*, *perlihat*.

- (40) Gambar 5.11 *memperlihat* bentuk tangan kanan dan kiri... (Disertasi_Tesis_Skripsi)
- (41) mendorong berdirinya Assesment Centre, *memperkenal* sistem penilaian... (Populer)

Sulit untuk menyematkan makna khusus pada verba *per-* berakar verba, selain mengindikasikan secara umum bahwa subjek ‘melakukan sesuatu’ terhadap objek [2, p. 134].

4.1.4 Verba dengan akhiran -kan

Verba berakhiran *-kan* pada hakikatnya bersifat transitif dan mengungkapkan makna tertentu tergantung dari akar katanya. Pertama-tama akan dibahas pelekatan *-kan* pada akar non-verba (§4.1.4.1) yang dilanjutkan pada akar verba (§4.1.4.2).

4.1.4.1 Verba transitif dengan -kan berdasarkan akar non-verba

Akhiran *-kan* dapat melekat dengan dua akar kata non-verba utama, yaitu adjektiva dan nomina (lihat Gambar 6). Verba *-kan* berakar adjektiva mengungkapkan makna ‘kausatif’ (layaknya *per-* [§4.1.3]), yaitu ‘menyebabkan objek memiliki sifat yang dinyatakan oleh [akar]’. Misalnya, *longgarkan/intensifkan X* berarti ‘membuat X menjadi longgar/intensif’. Verba transitif *-kan* juga bisa digunakan dalam kalimat aktif *meN-* dan pasif *di-* (atau *ter-*).

- (42) Jajarannya bakal terus *mengintensifkan* pembersihan saluran air...
(Koran)
- (43) ...latihan makin *diintensifkan* satu minggu menjelang keberangkatan.
(Laman_Resmi)
- (44) Realitas tersebut, perlu *terintensifkan* dan terakselerasi...³

Sejumlah rujukan terdahulu [1], [2] menyatakan bahwa verba kausatif *-kan* berbeda dengan verba kausatif *per-*. Verba dengan *per-* diasumsikan meningkatkan kualitas/ciri yang sudah ada pada objek sedangkan *-kan* memberikan objek kualitas baru (yang diasumsikan belum ada). Penelitian lanjutan menunjukkan bahwa verba kausatif *per-* dan *-kan* (dengan akar adjektiva yang sama, misalnya *perbesar* vs. *besarkan*) juga (i) berbeda maknanya dari segi konteks pemakaiannya [12], dan (ii) berbeda dari segi kelompok adjektiva yang cenderung (secara statistik) digunakan dengan salah satu dari kedua afiks tersebut [23].

Makna kausatif *-kan* juga dapat diungkapkan dengan akar nomina. Terdapat variasi makna untuk jenis verba ini. Makna pertama yang lumrah adalah ‘menyebabkan objek berada pada [akar]’ dengan akar (yang dapat dipandang sebagai) lokasi. Contohnya *daftarkan*, *pasarkan*, *sarungkan*, *petakan*, *sekolahkan*, *kasetkan*, *penjarakan*, *depositokan*, *pojokkan*.

- (45) ...*mendepositokan* uangnya di Bank Indonesia... (Majalah)
- (46) ...wacana wayang golek yang telah *dikasetkan*. (Disertasi_Tesis_Skripsi)

Makna terkait lain yang dapat diasumsikan adalah ‘memandang objek sebagai [akar]’, seperti *syaratkan*, *tokohkan*, *manusiakan*, *prioritaskan*, *ibaratkan*, *gambarkan*, *rahasiakan*.

- (47) ...tantangan tersendiri baginya sebagai seorang yang *ditokohkan*.
(Cerpen)
- (48) ...peran sastra dalam *memanusiakan* lingkungan. (Buku_Teks)

3 Diperoleh dari: <https://www.harianbhirawa.co.id/antisipasi-potensi-buruk-keamanan-digital/> (akses terakhir: 19 September 2022). Bentuk *terintensifkan* tidak ditemukan dalam korpus yang digunakan untuk makalah ini. Namun, pencarian menggunakan Google dengan domain laman “.id” mengeluarkan sejumlah penggunaan bentuk *terintensifkan* (pola pencarian dalam Google: "terintensifkan"; site:.id).

Makna sebelumnya dapat terkait dengan makna kausatif yang lebih umum, yaitu ‘mengakibatkan objek menjadi [akar]’, seperti misalnya *korbankan*, *calonkan*, *cawapreskan*, *gulingkan*. Jadi, *gulingkan X* (secara literal) bermakna ‘mengakibatkan X menjadi guling/terguling’, yang kemudian lebih condong digunakan dalam konteks metaforis.

- (49) ...pengurus DPP PKB terus ngotot *mencawapreskan* Cak Imin.
(Koran)
- (50) Zulkifli Lubis mengadakan makar untuk *menggulingkan* pemerintahan... (Biografi)

Makna selanjutnya adalah ‘memberikan [akar] pada objek’. Beberapa contohnya adalah *tugaskan*, *promosikan*, *doakan*, *perintahkan*, *amanatkan*, *ilhamkan*, *tepungkan*.

- (51) 5. *menepungkan* parutan inti biji mangga. (Buku_Teks)
- (52) ...kita bisa ikut *mempromosikan* kursus-kursus online di Udemy...
(Populer)

Jadi, misalnya, *mempromosikan X* berarti ‘memberi X promosi’, dan *menepungkan X* berarti ‘memberi X tepung’. Pengelompokan makna lainnya untuk verba *-kan* berakar nomina mungkin dilakukan namun seringnya makna tersebut tidak dapat diprediksi [1, p. 82].

4.1.4.2 Verba transitif dengan *-kan* berdasarkan akar verba

Akhiran *-kan* dapat dilekatkan pada verba taktransitif (seperti *jatuh*, *tenggelam*, *datang*, *takluk*, *mundur*, *maju*, *putus*, *tampil*, *ledak*) dan mengubah verba tersebut menjadi verba transitif dengan makna kausatif. Jadi, *jatuhkan X* bermakna ‘menyebabkan X jatuh’ atau *taklukkan X* bermakna ‘menyebabkan X takluk’. Layaknya verba transitif lainnya, verba transitif dengan *-kan* dapat muncul pada kalimat aktif *meN-* maupun pasif *di-* (atau *ter-*).

- (53) Serena *memecahkan* rekor yang pernah dibuat rekan senegaranya...
(Koran)
- (54) ...peluang untuk *merekatkan* kembali hubungan kedua negara.
(Majalah)

Selain melekat pada verba taktransitif, *-kan* juga dapat melekat pada akar verba transitif dan menyampaikan beberapa makna yang tergantung atas jenis akar verbanya. Salah satu kelompok verba tersebut menyatakan bahwa seseorang diuntungkan dari perbuatan yang dinyatakan verbanya. Misalnya, *pilihkan, masakkan, belikan, buatkan, ambilkan*.

- (55) ...orang tua yang lebih suka *memilihkan* calon suami bagi putrinya... (Novel)
(56) ...suami hanya minta *diambilkan* air putih untuk minum... (Populer)

Pada kedua contoh di atas, *putrinya* dan *suami* adalah orang yang diuntungkan dari tindakan yang dinyatakan oleh verba *pilihkan* dan *ambilkan*. Makna selanjutnya dari *-kan* dengan akar verba transitif mengindikasikan bahwa objek verbanya menjadi alat/instrumen dari aksi yang dinyatakan oleh verba tersebut. Contohnya *pukulkan, tikamkan, sentuhkan, siramkan*.

- (57) Bintang laut raksasa itu *memukulkan* tangannya ke Ali. (Novel)
(58) Ia *menyiramkan* bensin itu ke tubuh Mira. (Majalah)

Terakhir, untuk sejumlah verba *-kan* lainnya, sangat sulit menyematkan makna khusus pada *-kan*, selain sebagai penanda bahwa objek verbanya berperan sebagai penderita (*patient*)/pengalaman (*experiencer*) dari aksi yang dinyatakan oleh verba tersebut [1, p. 74]. Contohnya *pikirkan, laksanakan, terjemahkan, siarkan, butuhkan, kerjakan, terapkan*.

4.1.5 Verba dengan akhiran -i

Layaknya verba berakhiran *-kan*, verba berakhiran *-i* juga merupakan verba transitif. Terdapat dua makna umum dari verba berakhiran *-i*: (i) menyatakan bahwa objek verba merupakan lokasi (baik lokasi secara harfiah ataupun metaforis) dimana aksi dilakukan atau diarahkan (mis., *jalani, tandai*); (ii) menyatakan aksi berulang/intensif (mis. *pukuli*) [1, p. 89], [2, p. 141]. Akhiran *-i* dapat melekat dengan akar non-verba (§4.1.5.1) dan verba (§4.1.5.2).

4.1.5.1 Verba transitif dengan -i berdasarkan akar non-verba

Akar non-verba dominan untuk verba *-i* adalah nomina (Gambar 7), dengan makna umum dominan yaitu ‘memberikan/menerapkan [akar] pada

objek' [1, p. 89], [2, p. 143]. Contohnya *sinari*, *namai*, *nodai*, *biayai*, *bekali*, *bentengi*, *obati*, *cintai*, *bumbui*, *jiwai*, *akhiri*.

- (59) ...kerja keras PLN dalam *melistriki* Papua, khususnya Wamena...
(Laman_Resmi)

Makna lokatif terkait lainnya yang bersifat minor (dalam arti direalisasikan oleh sedikit bentuk kata) adalah 'membuang [akar] dari objek' [2, p. 144], seperti *kuliti*, *rumputi*, dan *bului*. Jadi, *kuliti/rumputi/bului X* berarti 'membuang kulit/rumput/bulu dari X'.

Sekelompok verba *-i* berakar nomina juga dapat bermakna 'subjek berperan/bertindak sebagai [akar] terkait objek'. Contohnya adalah *pelopori*, *juarai*, *rajai*, *kepalai*, *bintang*, *dalangi*, *musuhi*, *temani*. Jadi, *pelopori X* bermakna 'berperan sebagai pelopor terkait X'.

- (60) Berdasarkan teori-teori yang *mendasari* kajian ini, ... (Disertasi_
Tesis_Skripsi)

Kemudian, akhiran *-i* juga dapat melekat pada akar adjektiva. Salah satu maknanya adalah 'menerapkan sifat [akar] pada objek'. Jadi, *terangi X* bermakna 'menerapkan sifat terang pada X'. Beberapa contoh lainnya adalah *bodohi*, *teduhi*, *jahati*, *basahi*, *lamuri*, *kotori*.

- (61) Darah mengucur dari mulut, menetes-netes *memerahi* air danau.
(Novel)
(62) Kau tau bibi sedang *merindui* ibunya. (Cerpen)

Berikutnya, *-i* dapat melekat pada akar (i) adjektiva yang merujuk pada jarak atau dimensi spasial/ruang, seperti *jauh*, *dekat*, *dalam*, dan (ii) nomina ruang, seperti *tengah*, *atas*, *bawah*, *belakang*. Kelompok verba dengan akar kata spasial/lokatif ini juga memicu makna lokatif dari *-i*, khususnya 'subjek berada pada jarak atau lokasi yang dinyatakan oleh [akar] berkaitan dengan objek'. Perhatikan contoh kalimat berikut.

- (63) ...mereka berdua diminta *mendalami* ilmu tafsir... (Biografi)
(64) ...Anda akan dapat *mengatasi* keterpurukan Anda. (Populer)

Jadi, *mengatasi X* berarti ‘subjek (secara metaforis) berada di atas X’ (yaitu, ‘menaklukkan X’), dan *mendalami X* berarti ‘subjek (secara metaforis) berada di dalam X’ (dalam hal ini, ‘memahami X’). Akan tetapi, menurut Moeliono dkk. [2, p. 144], untuk kasus *bawahi* dan *belakangi*, objeklah (bukannya subjek) yang berada di posisi bawah dan belakang.

4.1.5.2 Verba transitif dengan -i berdasarkan akar verba

Akar kata yang juga dominan muncul dengan verba transitif -i adalah akar verba (Gambar 7). Sebagian verba -i tersebut dapat dipandang mengungkapkan makna yang secara skematis mirip dengan kemunculan verba akar yang sama dalam konstruksi taktransitif dengan frase lokatif pada konteks tertentu. Perhatikan pasangan contoh berikut untuk *naik* dan *naiki*.

- (65) ...sepeda Wawan mengarah *naik ke* bukit. (Majalah)
(66) ...memacu kudanya *menaiki* bukit. (Disertasi_Tesis_Skripsi)

Tidak semua verba berakhiran -i memiliki pasangan dalam versi intransitif dan meskipun terdapat keberpasangan, makna di antara verba taktransitif dan ketika muncul dengan -i bisa berbeda. Perhatikan kembali contoh lain berikut ini antara *naik* dan *naiki*.

- (67) ...membantu pengguna kursi roda untuk *naik ke* bus. (Majalah)
(68) ...Lail kembali *menaiki* bus rute 12... (Novel)

Dalam konteks kedua contoh spesifik di atas, *naiki bus* (68) telah bermakna khusus ‘menggunakan objek (bus) sebagai moda transportasi’ sedangkan *naik ke bus* (67) menonjolkan makna ‘pergerakan’, yaitu ‘naik menuju bagian dalam bus’ (meskipun secara implisit makna transportasi bisa muncul). Contoh lain keberpasangan verba taktransitif dan versi transitif -i adalah *tidur-tiduri, jatuh-jatuhi, hadir-hadiri, lewat-lewati* [1, pp. 92–93].

Selanjutnya, akhiran -i dapat melekat pada akar verba transitif (mis. *jilat, cabut, pukul, pandang, cium*) dan mengungkapkan implikasi makna ‘aksi [akar] yang berulang atau intensif’ [1, p. 98], [2, p. 145]. Sehingga, *cabuti X* mengikutkan implikasi makna ‘mencabut X secara berulang/intensif’, implikasi yang tidak dipicu oleh bentuk dasar *cabut X*.

- (69) Ia pun sempat akan *mencium* pipi perempuan itu. (Cermen)
(70) Ia mulai *menciumi* perempuan itu dengan ganas... (Novel)

Bentuk *ciumi pipi* mengindikasikan aksi ‘mencium pipi secara berulang’ namun *cium pipi* tidak membawa implikasi repetitif. Contoh lain adalah *tembaki, tebangi, tendangi*.

4.1.6 Verba dengan awalan + akhiran (konfiks)

Afiksasi verba BI dapat berupa gabungan antara awalan dan akhiran (konfiks), dan bagian ini membahas secara ringkas beberapa konfiks verba yang dominan ditemukan di korpus. Konfiks verba BI (mis. *per-/kan* [*permainkan*], atau *ber-/an* [*berpegangan*]) tidak sedominan bentuk yang lebih sederhana (*meN-*, *per-*, *ber-*, *-kan*, *-i*) (Gambar 2-Gambar 4).

Konfiks pertama adalah *ber-/an*, yang secara relatif paling dominan di antara konfiks lainnya (lihat Gambar 2-Gambar 4). Dari 305 bentukan kata dengan *ber-/an*, 58,7%-nya (n=179) dilandasi akar verba, diikuti oleh akar nomina (25,6%; n=78), adjektiva (11,5%; n=35), dan akar lainnya (4,26%; n=13). Salah satu makna dari *ber-/an* adalah ‘kejadian/aksi “timbal balik atau berbalasan”’ [2, p. 155], atau diistilahkan dengan makna ‘resiprokal’.

- (71) Bahagia rasanya melihat Ingrid dan Isabela *berpelukan* seperti itu. (Novel)
(72) ...asalkan ada waktu dan tidak *berbentrokan* dengan kegiatan kenegaran... (Koran)

Makna resiprokal ini dapat merujuk pada (i) dua orang saling melakukan aksi yang sama (mis. *bertatapan, berpandangan, bergandengan, berpapasan, bersahutan, bermesraan*), ataupun (ii) dua orang yang berada pada hubungan yang sama, baik hubungan spasial/lokatif (mis. *bertempelan, beririsan, berbatasan, berseberangan*) maupun hubungan personal (mis. *berpasangan, bertunangan, berbesan, bermusuhan, berkawanan*). Makna lain dari konfiks *ber-/an* adalah ‘kejadian yang tidak beraturan, tidak tentu, atau merambang’ [1, p. 113], [2, p. 155]. Contohnya adalah *berpatahan, berlarian, berpantulan, bercipratan, berkobaran, beterbangan, bercucuran, bertumbangan, berhamburan, berseliweran, berdebaran*.

Konfiks verba selanjutnya adalah *ber-/kan* (mis. *bersiluetkan, bernafaskan, berlantaikan*), yang menyatakan keadaan, bukannya aksi.

Meskipun diakhiri dengan *-kan* dan tampak seperti verba transitif karena diikuti oleh nomina (mis. *semak* dan *pohon* pada contoh (73)), verba *ber-/kan* tidak memiliki bentuk pasif *di-* [1, pp. 114–115]. Atas dasar tersebut, nomina yang mengikuti verba *ber-/kan* disebut pelengkap (bukan objek langsung).

- (73) ...jalan aspal hitam yang (...) *bersiluetkan* semak dan pohon. (Novel)
(74) ...Musik dan lagu yang *bernafaskan* agama... (Buku_Teks)

Makna dari konfiks *ber-/kan* adalah ‘subjek menjadikan pelengkap sebagai [akar]’, sehingga *bersiluetkan semak* berarti ‘menjadikan semak sebagai siluet’, *bernafaskan agama* berarti (secara harfiah) ‘menjadikan agama sebagai nafas’. Dari total 98 bentukan kata dengan konfiks *ber-/kan*, 69,4% (n=68 tipe) didominasi oleh akar nomina [1, p. 114], [2, p. 153].

Konfiks dominan terakhir yang dibahas adalah *per-/kan*. Konfiks ini menghasilkan verba transitif yang diikuti oleh objek langsung sehingga verba *per-/kan* dapat digunakan dalam konstruksi aktif *meN-* (*memper-/kan*) dan pasif (*diper-/kan*). Dari 104 bentukan kata *per-/kan* (baik dalam bentuk aktif dan pasif), 46,2%-nya (n=48) berasal dari akar verba, diikuti oleh akar nomina (36,5%; n=38), adjektiva (12,5%; n=13), dan lainnya (4,81%; n=5).

- (75) ...orang tua *mempersalahkan* lingkungan para remaja yang tidak baik... (Populer)
(76) ...ia juga tengah *mempertukarkan* nasib keduanya. (Cerpen)
(77) ...ia *dipergunjingkan* sebagai anak pelacur. (Majalah)
(78) Dokumen lain yang *dipersamakan* sebagai Faktur pajak... (Surat_Resmi)

Salah satu makna utama dari konfiks *per-/kan* adalah kausatif (‘membuat/menjadikan objek [akar]’) [1, p. 105]. Misalnya, *mempersalahkan X* pada contoh (75) berarti ‘menjadikan X salah’, *mempertukarkan X* pada (76) diartikan ‘menjadikan X tertukar’, atau *X dipersamakan (dengan/sebagai Y)* pada (78) berarti ‘X dibuat/dijadikan sama (dengan/sebagai Y)’. Untuk verba *per-/kan* lainnya yang dilandasi akar verba, konfiks ini tidak menyampaikan makna khusus selain (i) menyatakan bahwa objek verbanya adalah penderita dari aksi yang dinyatakan verba tersebut, dan (ii) membutuhkan kehadiran objek langsung untuk keberterimaan verbanya dalam kalimat [1, p. 106]. Beberapa contohnya adalah *pertimbangkan*, *perkiraan*, *perhitungan*,

permainkan, perbandingan. Selain *per-/kan*, terdapat juga konfiks verba transitif *per-/i* (mis. *perlengkapi, peringati, perdayai, persenjatai*), namun produktivitasnya rendah (perhatikan Gambar 2-Gambar 4).

4.2 Produktivitas Afiksasi Verba Bahasa Indonesia

Bagian ini dan dua bagian selanjutnya (§4.3 dan §4.4) melengkapi paparan kualitatif sebelumnya guna menampilkan profil kuantitatif yang mencirikan produktivitas imbuhan [22], [5], [6]. Tingkat produktivitas yang dibahas pada bagian ini (i) merupakan gabungan dari keseluruhan ragam teks dan tahun (yaitu, produktivitas pada keseluruhan korpus), dan (ii) dikhususkan untuk empat awalan verba utama, yaitu ME-⁴, DI-, BER-, dan TER-. Subtipe dari keempat awalan tersebut, yang mengandung akhiran, dibahas pada §4.2.1 dan §4.2.2.

Profil kuantitatif pertama yang dapat mencirikan produktivitas suatu afiks adalah **kekerapan/frekuensi token** (*token frequency*). Frekuensi token/kekerapan suatu afiks adalah total kemunculan/penggunaan tiap-tiap bentukan kata dengan afiks tersebut di dalam korpus. Frekuensi token yang dilaporkan adalah yang direlatifkan per satu juta kemunculan kata (panel kiri pada Gambar 1). Profil kuantitatif kedua adalah **frekuensi tipe** (*type frequency*) yang merujuk pada jumlah bentukan kata unik untuk suatu afiks; frekuensi tipe juga direlatifkan berdasarkan nilai per satu juta kata (panel tengah pada Gambar 1). Profil kuantitatif ketiga adalah jumlah **hapax legomena**, yaitu bentukan kata dari suatu afiks yang memiliki frekuensi token satu; nilai hapax juga direlatifkan per satu juta kemunculan kata dalam korpus (panel kanan pada Gambar 1). Pangkalan data verba makalah ini mengandung 1.981.462 token, 13.646 tipe, dan 3.444 hapax legomena.

4 Afiksasi verba utama yang dicetak dengan huruf kapital mewakili dan mencakup subtipe afiksasinya. Misalnya, distribusi untuk ME- pada Gambar 1 meliputi distribusi untuk *meN-*, *meN-/kan*, *meN-/i*, *memper-*, *memper-/kan*, dll. Hal serupa berlaku untuk awalan DI-, BER-, dan TER-. Lihat §4.2.1 untuk ulasan produktivitas masing-masing subtipe dari ME-, DI-, BER-, dan TER-.



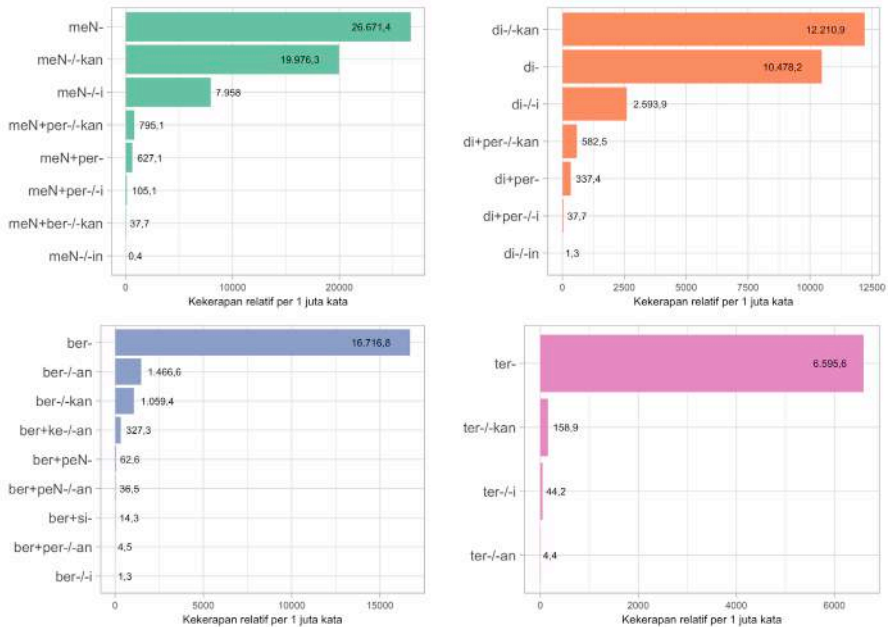
Gambar 1 Produktivitas imbuhan verba gabungan di keseluruhan korpus

Gambar 1 menunjukkan bahwa verba berawalan ME- (termasuk yang memiliki akhiran) paling produktif di ketiga jenis profil kuantitatifnya, yaitu dari segi (i) kekerapan (atau frekuensi token relatif) verba tersebut, (ii) jumlah bentukan kata (frekuensi tipe relatif), dan (iii) jumlah bentukan kata berfrekuensi token satu (yaitu jumlah **hapax**)⁵. Hapax adalah bentukan kata yang secara potensial bisa berupa (meskipun tidak selalu) bentukan baru [22, pp. 905–906]. Sebaliknya, verba berawalan TER-, yang mencirikan verba pasif statis [24], adalah yang terendah di ketiga profil kuantitatifnya dibandingkan ketiga awalan yang lain.

4.2.1 Produktivitas sub tipe afiksasi verba

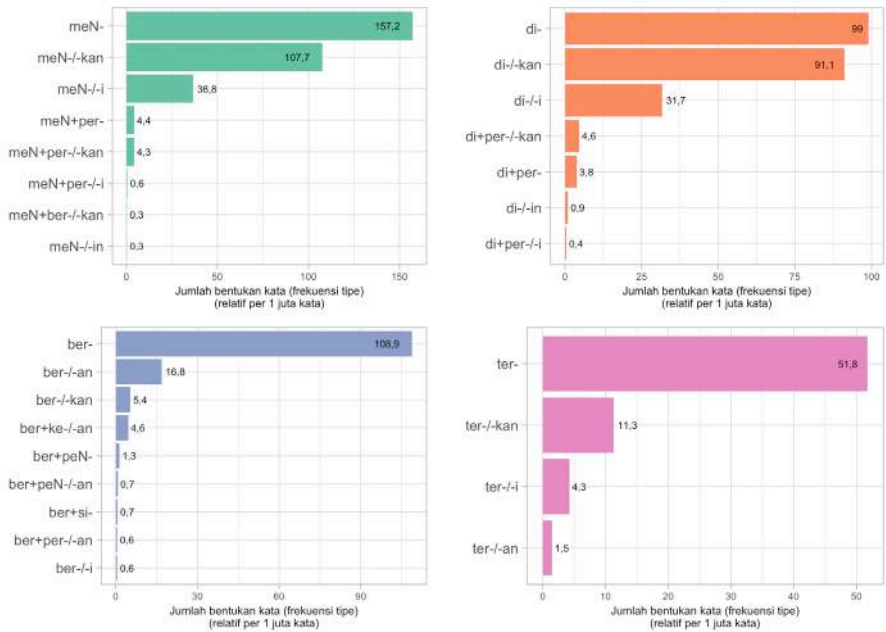
Subbagian ini berfokus pada produktivitas sub tipe afiksasi verba untuk ME-, DI-, BER-, dan TER- berdasarkan tiga profil kuantitatif yang telah dijabarkan sebelumnya (§4.2). Sub tipe afiksasi yang diulas adalah yang memiliki frekuensi tipe riil lebih dari lima kata; nilai “riil” maksudnya adalah nilai kemunculan dalam korpus yang belum diubah menjadi nilai relatif per satu juta kemunculan kata. Gambar 2 berikut menyajikan frekuensi token dari sub tipe afiksasi verba ME-, DI-, BER-, dan TER-.

5 Berkas visualisasi berwarna dengan kualitas tinggi tersedia pada direktori *figures* pada <https://osf.io/nuxd4/>



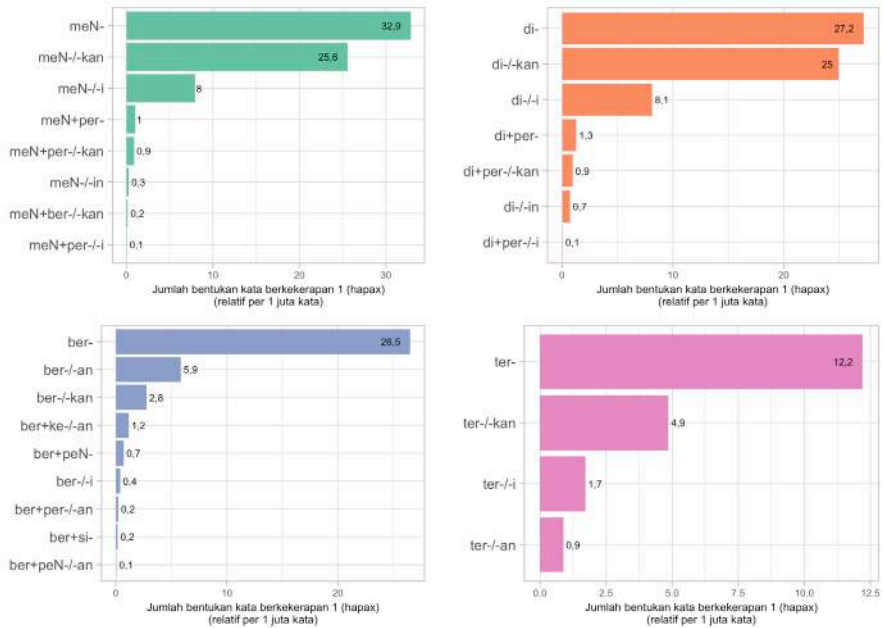
Gambar 2 Frekuensi token relatif sub tipe afiksasi
(dengan frekuensi tipe riil lebih besar dari 5)

Untuk sub tipe afiksasi yang frekuensi tipe-nya lebih tinggi dari lima, BER- memiliki sub tipe yang paling beragam, yaitu sembilan sub tipe. Tiga sub tipe BER- yang kerap digunakan berdasarkan Gambar 2 adalah *ber-*, *ber-/an*, dan *ber-/kan*; frekuensi token *ber-* yang paling dominan. Sebaliknya, TER- hanya memiliki empat sub tipe dengan frekuensi tipe riil di atas lima, dan hanya *ter-* yang paling kerap digunakan. Kemudian, untuk ME-, hanya tiga dari delapan sub tipe yang frekuensi tokennya relatif tinggi, yaitu *meN-*, *meN-/kan*, dan *meN-/i*. Di sisi lain, DI- juga memiliki tiga sub tipe dominan yang mencerminkan bentuk pasif dari ketiga sub tipe dominan ME-: *di-* dan *di-/kan* tampak mendominasi, sedangkan bentuk *-i*, yaitu *di-/i* jauh di bawah dan lebih rendah secara relatif dibandingkan bentuk aktif *me-/i*. Gambar 3 berikut menyajikan data frekuensi tipe relatif untuk tiap-tiap sub tipe afiksasi.



Gambar 3 Frekuensi tipe relatif subtype afiksasi (dengan frekuensi tipe riil lebih besar dari 5)

Dari segi frekuensi tipe, ketiga subtype afiksasi ME- (*me-*, *me-/kan*, dan *me-/i*) sama-sama dominan dibandingkan dengan subtype ME- lainnya. Sebaliknya untuk DI-, meskipun ketiga subtype dominan berdasarkan frekuensi token juga dominan dari segi frekuensi tipenya, namun terdapat sedikit perbedaan, yaitu *di-* memiliki tipe terbanyak meskipun tidak sekerap *di-/kan* untuk frekuensi tokennya. Kebertahanan dominasi *ber-* dan *ter-* pada frekuensi token juga tercermin pada frekuensi tipenya dibandingkan dengan subtype BER- dan TER- lainnya. Gambar 4 berikut menyajikan distribusi relatif jumlah hapax subtype afiksasi verba.

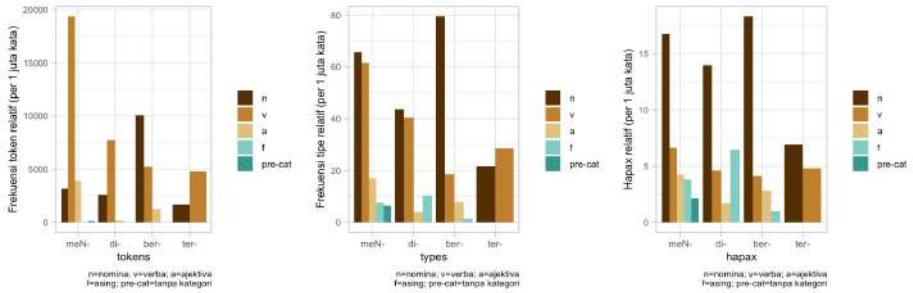


Gambar 4 Jumlah hapax relatif sub tipe afiksasi (dengan frekuensi tipe riil lebih besar dari 5)

Tren serupa untuk distribusi hapax juga ditunjukkan oleh sub tipe afiksasi yang dominan pada dua profil kuantitatif sebelumnya. Secara statistik, terdapat korelasi positif di antara ketiga profil kuantitatif untuk tiap-tiap sub tipe afiksasi. Sebagai contoh, semakin tinggi frekuensi token suatu sub tipe afiksasi (mis. *me-*), semakin tinggi pula frekuensi tipe dan hapax dari sub tipe afiksasi tersebut. Lebih lanjut, Gambar 2-Gambar 4 menunjukkan minornya distribusi konfiks tertentu (§4.1.6) dan sekaligus memberikan bukti kuantitatif terhadap pencirian Sneddon dkk. [1, p. 108] bahwa konfiks merupakan salah satu jenis afiksasi verba sekunder.

4.2.2 Produktivitas jenis kelas kata pada sub tipe afiksasi verba

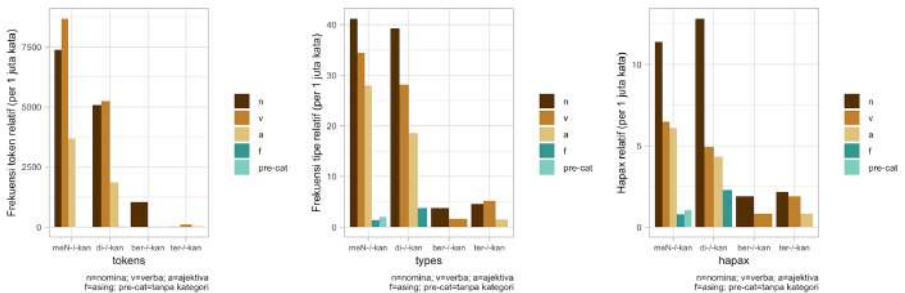
Sub-bagian ini menjabarkan produktivitas kelas kata dari bentuk akar untuk tiga sub tipe afiksasi dominan. Kelas kata yang difokuskan adalah yang berfrekuensi tipe relatif lebih besar dari satu. Gambar 5 menyarikan distribusi kelas kata untuk *meN-/di-/ber-/ter-+[akar]*.



Gambar 5 Produktivitas kelas kata akar pada sub tipe *meN/di/ber/ter*+*[akar]*

Kekerapan (panel kiri) sub tipe *meN-*, *di-*, dan *ter-* dengan akar verba paling tinggi dibandingkan dengan kelas kata lain sedangkan kekerapan *ber-* dengan akar nomina paling dominan dibandingkan dengan kelas kata lain. Tendensi sebaliknya ditunjukkan oleh pengukuran frekuensi tipe (panel tengah). Frekuensi tipe *meN-* dan *di-* dengan akar nomina dan verba lebih dominan daripada akar lainnya. Kemudian, frekuensi tipe *ber-* dengan akar nomina juga paling dominan dibandingkan dengan kelas kata akar lain; hal ini sejalan dengan [1, p. 65], dan [2, p. 149]. Untuk nilai hapax, keempat sub tipe secara signifikan memiliki jumlah hapax nomina tertinggi dibandingkan jumlah hapax dengan kelas kata yang lain. Pengamatan menarik lainnya dari Gambar 5 adalah akar asing (*f*) mulai terlihat pada pengukuran frekuensi tipe dan hapax dibandingkan pada frekuensi token. Hal ini mungkin mencerminkan bahwa pengayaan bentukan tipe-tipe kosakata baru dengan afiks tertentu juga dapat dipicu melalui penggabungan afiks tersebut dengan akar asing.

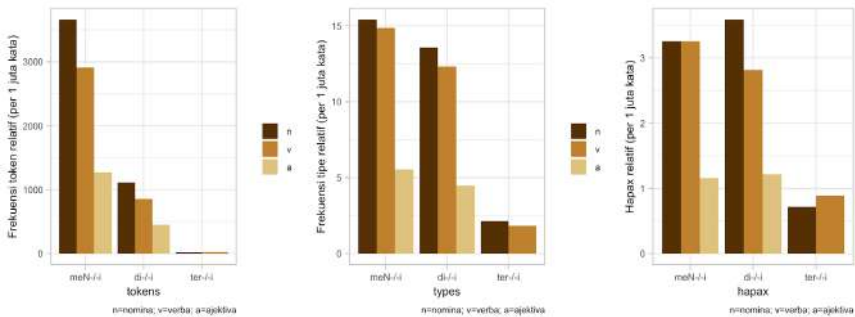
Gambar 6 menyajikan produktivitas kelas kata untuk sub tipe afiksasi berakhiran *-kan*, yaitu *meN/-kan*, *di/-kan*, *ber/-kan*, dan *ter/-kan*.



Gambar 6 Produktivitas kelas kata akar pada sub tipe *meN/di/ber/ter*+*[akar]*+*kan*

Dapat dilihat bahwa frekuensi token (panel kiri) untuk *ber-/kan* dan *ter-/kan* sangat rendah, dibandingkan *meN-/kan* dan *di-/kan*. Kemudian, kekerapan *meN-/kan* dan *di-/kan* dengan akar adjektiva adalah yang terendah dibandingkan dengan akar lain, utamanya verba, yang lebih tinggi dibandingkan nomina. Akan tetapi, frekuensi tipe (panel tengah) dan juga hapax (panel kanan) untuk akar kata nomina yang paling dominan untuk keempat sub tipe afiksasi. Tren dominasi nomina pada frekuensi tipe dan hapax untuk sub tipe dengan *-kan* ini sejalan dengan yang ditemukan pada verba berawalan *meN-*, *di-*, dan *ber-* (Gambar 5).

Sub tipe terakhir yang diulas adalah produktivitas kelas akar kata untuk afiksasi verba dengan *-i* pada *meN-/i*, *di-/i*, dan *ter-/i*. Gambar 7 menyarikan ketiga profil kuantitatifnya.



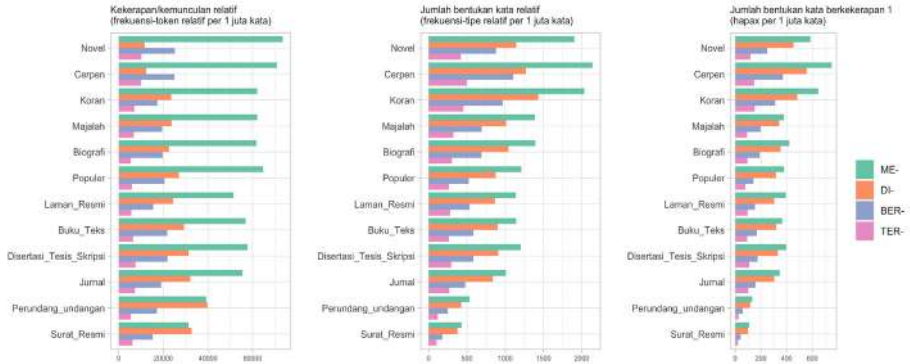
Gambar 7 Produktivitas kelas kata akar pada sub tipe *meN/di/ber/ter+[akar]+i*

Ketiga sub tipe *-i* dengan akar nomina memiliki frekuensi token relatif paling tinggi daripada kelas kata lainnya⁶. Untuk frekuensi tipe, jumlah bentukan kata sub tipe *-i* dengan nomina dan verba lebih mendominasi dibandingkan adjektiva, utamanya untuk *meN-/i* dan *di-/i*. Tendensi serupa ditunjukkan kedua sub tipe tersebut pada jumlah hapaxnya, yang masih didominasi oleh nomina dan verba, meskipun hapax nomina dan verba untuk *meN-/i* dan *di-/i* setara. Keterbatasan akar adjektiva dan dominasi akar nomina yang muncul dengan *-i* sejalan dengan pengamatan yang diajukan oleh Sneddon dkk. [1, pp. 89–91].

⁶ Untuk *ter-/i*, yang secara grafik hampir tidak terlihat, frekuensi token relatifnya dengan akar nomina adalah 2,15 per sejuta kata ($n=361$) dan 1,82 per sejuta kata dengan akar verba ($n=377$).

4.3 Produktivitas Afiksasi Verba Bahasa Indonesia Berdasarkan Ragam Teks

Gambar 8 berikut menampilkan distribusi ME-, DI-, BER-, dan TER- (layaknya pada Gambar 1) terhadap ragam teks dalam korpus. Produktivitas subtype keempat afiksasi tersebut disajikan pada Gambar 9 hingga Gambar 11.



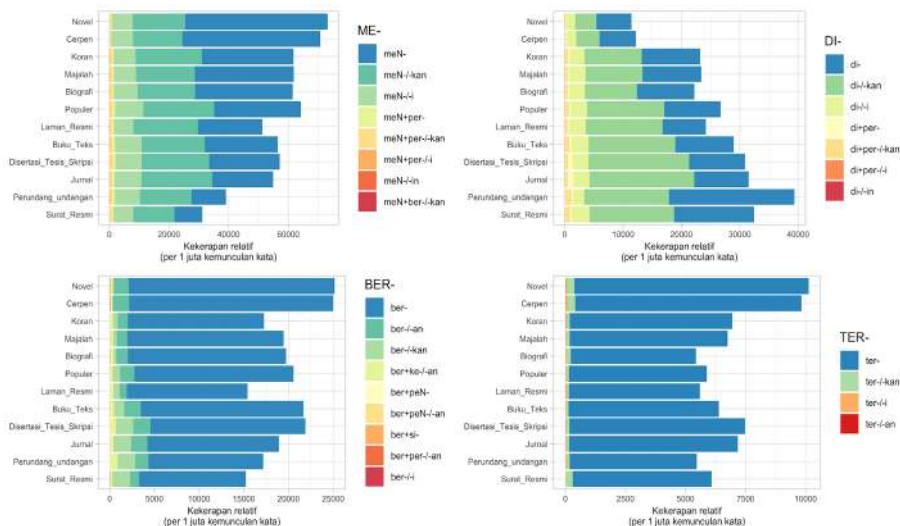
Gambar 8 Produktivitas awalan verba berdasarkan ragam teks dalam korpus

Untuk frekuensi token (panel kiri), salah satu tren utamanya adalah begitu dominannya afiks ME- pada sepuluh ragam teks, kecuali pada Perundang_undangan dan Surat_Resmi, yang frekuensi tokennya lebih tinggi untuk DI- dibandingkan ME-. Kemudian, frekuensi token DI- yang paling rendah ditemukan pada ragam Novel dan Cerpen, dua ragam yang memiliki kemunculan tertinggi untuk BER- dibandingkan ragam lainnya. Terakhir, TER- secara konsisten berkekerapan paling rendah (termasuk juga untuk frekuensi tipe dan hapax TER-), sejalan dengan tren gabungan TER-di keseluruhan korpus (Gambar 1). Distribusi keempat afiks terkait frekuensi tipe (panel tengah) dan hapax-nya (panel kanan) mencerminkan tren umum di keseluruhan korpus yang ditampilkan pada Gambar 1. Yaitu, pada tiap-tiap ragam teks, ME- selalu memiliki bentukan kata unik dan hapax terbanyak, diikuti oleh DI-, BER- dan TER- yang terendah.

Dari sudut pandang ragam teks, ragam cerita Novel dan Cerpen, serta Koran, menjadi tiga ragam dengan jumlah bentukan kata (frekuensi tipe) dan kemunculan hapax (untuk keempat awalan secara total) yang paling tinggi. Terkait dominasi ketiga ragam tersebut pada frekuensi tipe dan hapax, kami

berasumsi bahwa ketiga ragam tersebut (i) mawadahi/memicu pembentukan kata baru pada ragam tulis, dan (ii) memiliki tingkat keragaman leksikal yang lebih tinggi dibandingkan ragam teks yang lain. Sebaliknya, ragam formal normatif Perundang_undangan dan Surat_Resmi tidak hanya memiliki frekuensi token terendah tapi juga memiliki frekuensi tipe dan hapax terendah untuk keempat awalan. Kami berasumsi bahwa rendahnya produktivitas keempat afiks pada kedua ragam ini, utamanya dari jumlah bentukan kata dan hapax-nya, mungkin mencerminkan (i) minimnya variasi/pilihan leksikal afiksasi verbanya dan (ii) konvensionalitas kandungan teks pada kedua ragam tersebut. Bukti bahwa afiksasi verba menunjukkan variasi terkait ragam teks menjadi wawasan tambahan dalam deskripsi tatabahasa dan hanya akan diperoleh dengan mengikutkan variabel ragam teks dalam melihat distribusi afiks dalam korpus (lihat misalnya [25] yang mengikutkan dimensi ragam teks dalam deskripsi tatabahasa bahasa Inggris berbasis korpus).

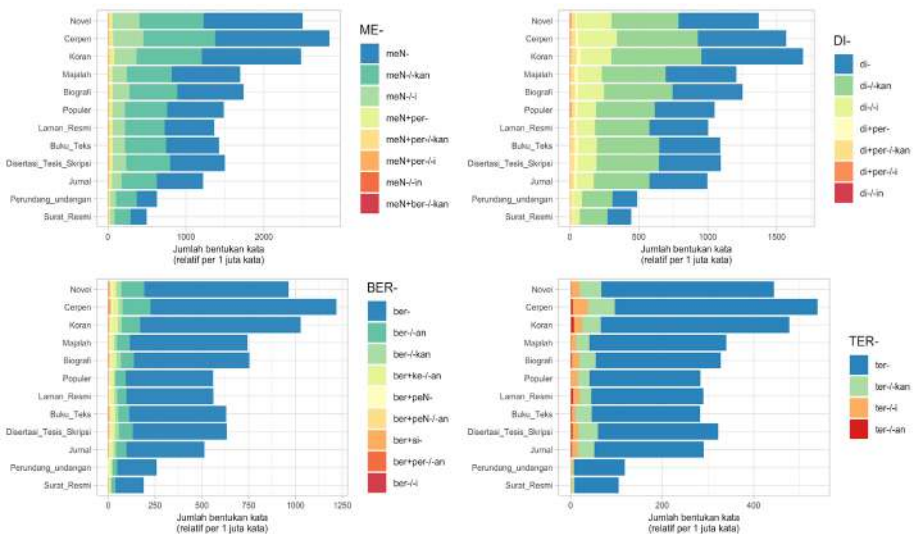
Gambar 9 berikut menyajikan frekuensi token sub tipe afiksasi berdasarkan ragam teks.



Gambar 9 Frekuensi token relatif berdasarkan ragam teks

Terdapat sejumlah variasi menarik pada Gambar 9. Kekerapan sub tipe untuk DI- (utamanya *di-*, *di/-kan*, dan *di/-i*) paling rendah pada ragam cerita seperti Novel dan Cerpen (sejalan dengan tren keseluruhan; lihat kembali

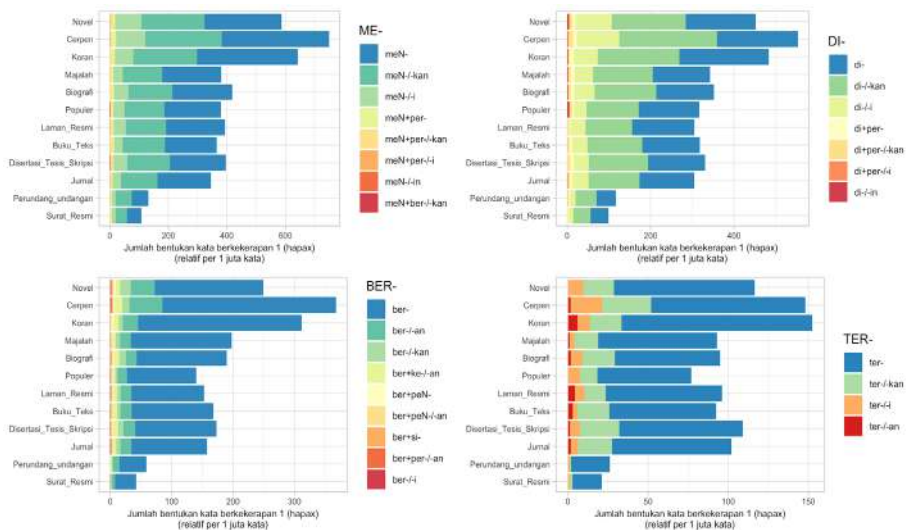
Gambar 8). Subtipe *di-/-kan* dan *di-* dominan pada ragam formal (Perundang_undangan, Surat_Resmi, Laman_Resmi) dan ilmiah (Jurnal, Disertasi/Tesis/Skripsi, dan Buku_Teks). Tendensi ini mencerminkan ciri formal dan ilmiah pemakaian pasif DI-. Tendensi yang berbeda ditunjukkan oleh subtipe pasif statis TER-, utamanya *ter-*, yang paling kerap muncul pada ragam Novel dan Cerpen, meskipun *ter-* juga cukup berimbang pada ragam teks lainnya. Selanjutnya, berseberangan dengan DI-, bentuk aktif ME- dan subtipenya memiliki kekerapan terendah pada ragam formal Perundang_undangan dan Surat_Resmi; yang juga menarik adalah subtipe *me-/-kan* dibandingkan subtipe ME- lainnya paling dominan pada Perundang_undangan, Surat_Resmi dan Jurnal. Tendensi akhiran *-kan* yang dominan pada ragam formal dan ilmiah juga ditunjukkan oleh konfiks *ber-/-kan* (lihat §4.1.6), yang kekerapannya lebih tinggi dan/atau mulai berimbang dengan *ber-/-an* hanya pada ragam Surat_Resmi, Perundang_undangan, Jurnal, dan Disertasi/Tesis/Skripsi. Gambar 10 berikut menyajikan distribusi jumlah bentukan kata (frekuensi tipe) relatif untuk subtipe ME, DI-, BER-, dan TER- pada masing-masing ragam teks.



Gambar 10 Frekuensi tipe relatif berdasarkan ragam teks

Satu tren menarik yang sejalan dengan tren pada Gambar 8 adalah bahwa ragam formal Perundang_undangan dan Surat_Resmi memiliki

jumlah bentukan kata terendah untuk keempat afiks beserta subtipenya. Hal ini mengisyaratkan rendahnya variasi leksikal untuk kedua ragam formal tersebut. Cerpen, Novel, dan Koran sebaliknya menunjukkan kekayaan tipe kata untuk sub tipe dominan dari masing-masing afiks. Lebih lanjut, perbedaan produktivitas sub tipe dominan untuk ME- dan DI- (*meN-*, *di-*, *meN/di-/-kan*, *meN/di-/-i*) pada keseluruhan korpus (Gambar 3) juga tercermin pada tiap-tiap ragam teks. Pola distribusional serupa ditunjukkan oleh sub tipe dari BER- dan TER-. Gambar 11 berikut menampilkan distribusi relatif hapax masing-masing sub tipe pada masing-masing ragam teks.

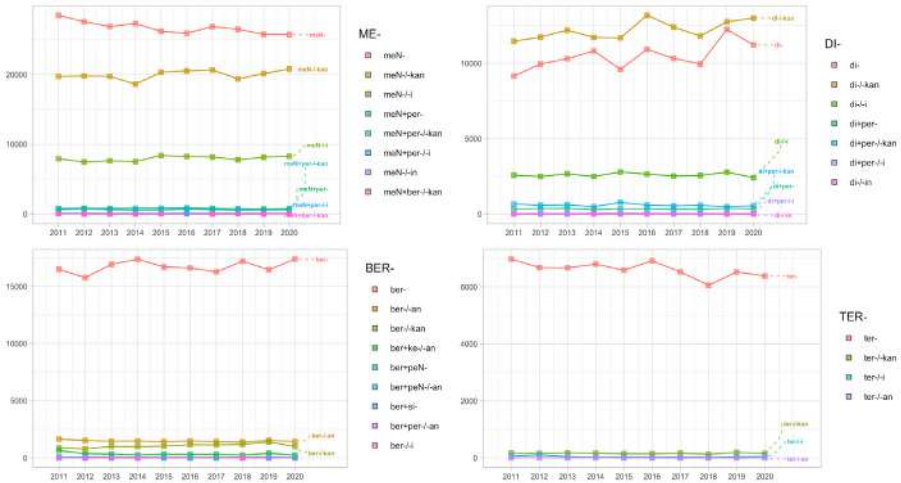


Gambar 11 Jumlah hapax relatif berdasarkan ragam teks

Pengukuran nilai hapax juga menunjukkan pola yang sejalan dengan frekuensi tipe. Sub tipe afiksasi yang dominan terkait frekuensi tipenya (mis. *meN-*, *di-*, *meN/di-/-kan*, *ber-*, *ber/-an*, dan *ter-*) juga dominan untuk nilai hapax-nya pada tiap-tiap ragam teks. Sejalan dengan pengamatan frekuensi tipe, ragam Perundang_undangan dan Surat_Resmi juga memiliki jumlah hapax terendah dibandingkan ragam teks lainnya untuk keempat afiksasi. Hal ini mengindikasikan bahwa kedua ragam teks tersebut tidak menjadi ranah dalam menghasilkan bentukan-bentukan baru, yang menjadi ranah dari Cerpen, Koran, dan Novel.

4.4 Produktivitas afiksasi verba bahasa Indonesia dari tahun ke tahun (2011-2020)

Informasi tahun dalam korpus yang digunakan dapat diolah untuk mengamati distribusi sub tipe afiksasi verba dalam sepuluh tahun terakhir (2011-2020). Gambar 12 berikut menyajikan frekuensi token relatif sub tipe ME-, DI-, BER-, dan TER- sepuluh tahun terakhir.



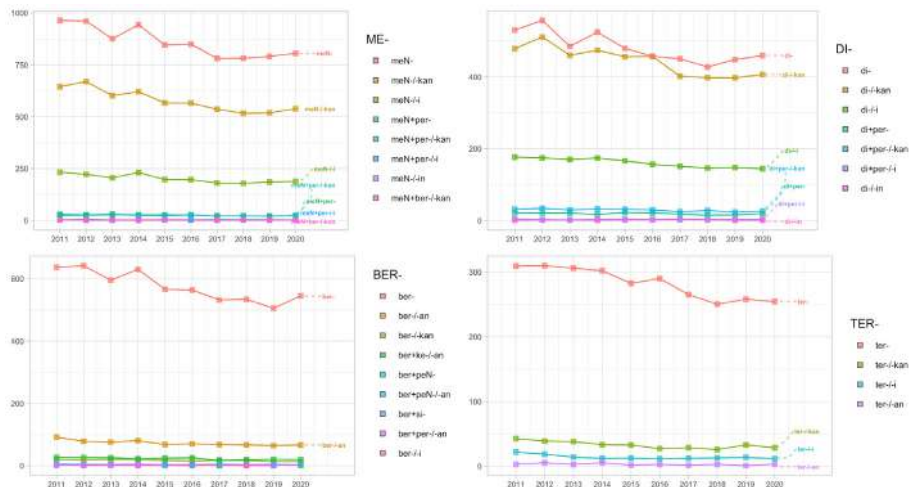
Gambar 12 Frekuensi token relatif berdasarkan tahun

Terdapat dominasi konsisten dari *meN-*, *meN/-kan*, *meN/-i*, *di-*, *di/-kan*, *di/-i*, *ber-*, dan *ter-* pada satu dekade terakhir, sejalan dengan data keseluruhan korpus (§4.2.1). Bentuk *ber-* dan *ter-* adalah yang paling dominan untuk BER- dan TER-, meskipun tren untuk *ter-* menurun signifikan ($\tau=-0,64$, $p < 0.001$)⁷. Yang menarik adalah kekerapan *meN-* dan *meN/-kan* tidak sedekat ketika dalam bentuk pasif *di-* dan *di/-kan*, yang kedekatannya konsisten. Sub tipe *meN/-i* dan *di/-i* secara konsisten token relatifnya jauh lebih rendah daripada dua sub tipe sebelumnya. Kekerapan relatif *di-* dan *di/-kan* menunjukkan tren meningkat yang tidak signifikan ($\tau=0,51$, $p=0.05$ untuk *di-*, dan $\tau=0,51$, $p=0.05$ untuk *di/-kan*); sebaliknya kekerapan relatif *meN-* menurun sangat signifikan ($\tau=-0,78$, $p < 0.001$).

Persaingan distribusi *di-* dan *di/-kan* tampak lebih ketat (yaitu, tidak

7 Uji statistik yang digunakan adalah uji korelasi *Kendall* dengan fungsi *cor.test()* di R [26, p. 378].

berbeda secara signifikan) pada frekuensi tipe relatif keduanya ($t_{\text{Welch}}=2,0499$; $df=17,964$; $p_{\text{two-tailed}}=0.055$; Cohen's $d=0,916$)⁸ (perhatikan Gambar 13 berikut).

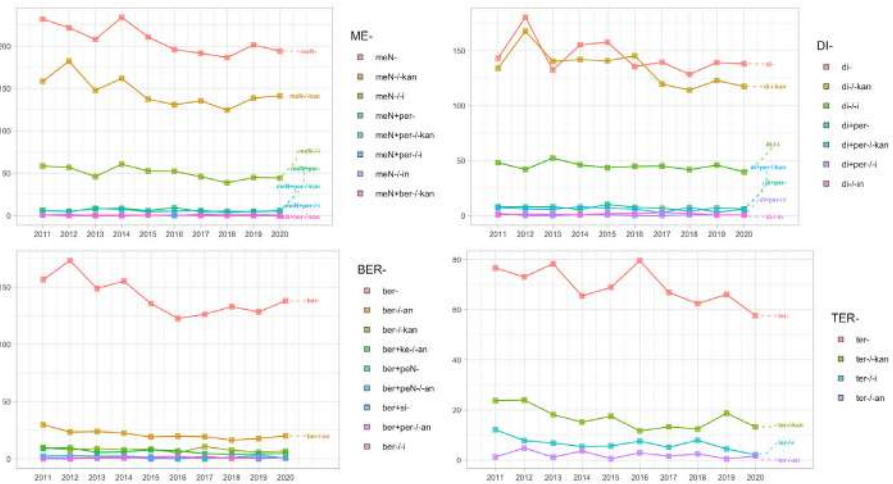


Gambar 13 Frekuensi tipe relatif berdasarkan tahun

Secara umum, frekuensi tipe relatif untuk sub tipe afiksasi dominan (yaitu *meN-*, *meN-/kan*, *meN-/i*, *di-*, *di-/kan*, *di-/i*, *ber-* dan *ter-*) juga mengalami penurunan kuat yang signifikan⁹. Asumsi yang kami tawarkan terkait tren terbalik antara peningkatan frekuensi token khususnya *di-* dan *di-/kan* dan penurunan signifikan frekuensi tipe *di-* dan *di-/kan* adalah stagnasi pembentukan kata untuk afiks yang tokennya tinggi. Kini mari kita lihat distribusi hapax yang ditampilkan pada Gambar 14 berikut ini.

8 Uji statistik yang digunakan adalah *T-Test for Independent Sample* dengan nilai efek *Cohen's d*.

9 Klaim penurunan kuat dan signifikan ini dilandasi atas hasil analisis statistik. Analisis tersebut dapat pembaca hasilkan dengan mengunduh data pendukung makalah ini. Kemudian, jalankan kode pemrograman dalam berkas R dalam data pendukung tersebut. Kode R tersebut berisi langkah-langkah analisis kuantitatif makalah ini, utamanya jalankan kode pada baris dengan label “4.2.1 correlation test for type frequency of some subtypes”.



Gambar 14 Jumlah hapax relatif berdasarkan tahun

Nilai hapax untuk subtipen afiksasi dominan juga menunjukkan tren penurunan (nilai negatif uji korelasi *Kendall*), meskipun hanya subtipen *meN-*, *meN/-i* dan *ter-* yang memiliki derajat penurunan yang kuat dan signifikan secara statistik¹⁰. Hal yang patut juga diperhatikan pada Gambar 14 adalah tensi ketat untuk variasi nilai hapax relatif di antara *di-* dan *di/-kan*.

5. SIMPULAN DAN SARAN

Makalah ini telah membahas afiksasi verba BI (§4.1) terkait (i) bentuknya (awalan, akhiran, dan gabungan keduanya, atau konfiks), (ii) fungsinya dalam menghasilkan verba transitif dan taktransitif, (iii) makna yang diungkapkan oleh verba berafiks tertentu sehubungan dengan jenis akar katanya, serta (iv) contoh-contoh pemakaiannya yang diperoleh dari korpus BI digital masif dari berbagai ragam teks dalam rentang 2011-2020. Keterbatasan ruang menyebabkan beberapa jenis afiksasi tidak dapat diulas secara rinci, seperti verba pasif *ke/-an* (mis. *kecipratan*, *kedahuluhan*, *kecampuran*, *kecurian*)

10 Klaim penurunan kuat dan signifikan ini dilandasi atas hasil analisis statistik. Analisis tersebut dapat pembaca hasilkan dengan mengunduh data pendukung makalah ini. Kemudian, jalankan kode pemrograman dalam berkas R dalam data pendukung tersebut. Kode R tersebut berisi langkah-langkah analisis kuantitatif makalah ini, utamanya (jalankan kode pada baris dengan label “4.3.1 correlation test for relative hapax of some subtypes”).

maupun verba pasif statis *ter-* yang bisa bermakna ‘kemampuan/abilitatif’ (mis. *terbayar*, *terkendalikan*), ‘kejadian tak disengaja’ (mis. *terseret*, *tertima*), ataupun ‘keadaan’ (mis. *terkenal*, *tergabung*).

Selanjutnya, makalah ini juga menyajikan hasil analisis kuantitatif atas beberapa aspek pemakaian afiksasi verba, yang sebelumnya tidak diulas secara mendalam dan sistematis, yaitu (i) produktivitas afiks keseluruhan (§4.2; §4.2.1), (ii) produktivitas kelas kata akar dari suatu jenis afiks (§4.2.2), dan (iii) produktivitas afiks terkait ragam teks (§4.3) dan tahun (§4.4). Secara umum, verba berawalan ME- paling produktif, diikuti oleh DI-, BER-, dan TER- (Gambar 1). Subtipe dominan dan produktif dari keempat awalan verba ini adalah *meN-*, *meN-/kan*, *meN-/i* (untuk ME-), *di-*, *di-/kan*, *di-/i* (DI-), *ber-* dan *ber-/an* (BER-), dan *ter-* (TER-) (§4.2.1). Kemudian analisis distribusi kelas kata akar mengkonfirmasi pernyataan introspektif terdahulu terkait dominasi jenis akar tertentu untuk afiksasi tertentu (mis. tingginya frekuensi tipe dan hapax akar nomina yang melekat dengan *ber-*; Gambar 5).

Dengan mengikutkan dimensi ragam teks (§4.3), kami menemukan variasi distribusi dan produktivitas subtipe afiksasi verba. Contohnya, teks Perundang_undangan dan Surat_Resmi memiliki frekuensi tipe dan jumlah hapax terendah untuk keempat afiksasi verba beserta subtipenya; distribusi ini mengindikasikan rendahnya variasi leksikal kedua ragam tersebut. Sebaliknya, cerita naratif, seperti Cerpen dan Novel, serta ragam informatif seperti Koran memiliki frekuensi tipe dan hapax paling dominan; distribusi ini mengindikasikan tingginya keragaman leksikal dan potensi ragam tersebut sebagai wadah bentukan kata baru. Terkait dimensi tahun (§4.4), sejumlah subtipe afiksasi verba dominan mengalami penurunan kuat dan signifikan pada tataran frekuensi tipe dan hapax. Hal ini mengindikasikan stagnasi bertahap subtipe afiksasi dominan tersebut terkait kekayaan leksikalnya (i.e., bentukan kata yang dibingkai oleh subtipe tersebut).

Sebagai penutup, kajian morfologi BI secara umum akan lebih komprehensif dengan memanfaatkan (i) keberlimpahan data korpus, guna memperoleh data alamiah serta informasi kuantitatif, dan (ii) dimensi non-linguistik (seperti ragam teks dan tahun) untuk mengamati distribusi afiksasi terhadap ragam teks dan rentang periode tertentu.

Referensi

- [1] J. N. Sneddon, A. Adelaar, D. N. Djenar, and M. C. Ewing, *Indonesian reference grammar*, 2nd ed. Crows Nest, New South Wales, Australia: Allen & Unwin, 2010.
- [2] A. M. Moeliono, H. Lapoliwa, H. Alwi, S. S. Tjatur, W. Sasangka, and S. Sugiyono, *Tata bahasa baku bahasa Indonesia*, Edisi Keempat. Jakarta: Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan, 2017. Accessed: Mar. 21, 2021. [Online]. Available: <http://repositori.kemdikbud.go.id/16351/>
- [3] I. W. Arka and N. Yannuar, ‘On the morphosyntax and pragmatics of *-in* in Colloquial Jakartan Indonesian’, *Indones. Malay World*, pp. 1–23, Sep. 2016, doi: 10.1080/13639811.2016.1215129.
- [4] J. Beavers and I. N. Udayana, ‘Middle voice as generalized argument suppression’, *Nat. Lang. Linguist. Theory*, Jun. 2022, doi: 10.1007/s11049-022-09542-5.
- [5] G. P. W. Rajeg and I. M. Rajeg, ‘Mempertemukan morfologi dan linguistik korpus: Kajian konstruksi pembentukan kata kerja [*per-*+Ajektiva] dalam Bahasa Indonesia’, in *Rona Bahasa: Buku persembahan kepada Prof. Dr. Aron Meko Mbete memasuki masa purnatugas*, I. N. Sudipa and M. S. Satyawati, Eds. Denpasar, Bali, Indonesia: Swasta Nulus, 2017, pp. 288–327. [Online]. Available: <https://doi.org/10.4225/03/5a0627de02453>
- [6] K. Denistia and R. H. Baayen, ‘The Indonesian prefixes PE- and PEN-: A study in productivity and allomorphy’, *Morphology*, Feb. 2019, doi: 10.1007/s11525-019-09340-7.
- [7] G. P. W. Rajeg and K. Denistia, ‘A study in productivity of Indonesian causative *per-* and *-kan*’, presented at the The 24th International Symposium on Malay/Indonesian Linguistics (ISMIL 24), Online, May 21, 2021. [Online]. Available: <https://doi.org/10.6084/m9.figshare.14633133>
- [8] A. C. Fadillah, I. Nurhayani, and S. E. Tabiati, ‘The Addition of Indonesian Prefixes *meN-* and *di-* to English Bases: A Corpus-based Study’, *J. Lang. Lit.*, vol. 21, no. 2, Art. no. 2, Sep. 2021, doi: 10.24071/joll.v21i2.3252.
- [9] G. P. W. Rajeg and I. M. Rajeg, ‘Pemahaman kuantitatif dasar dan penerapannya dalam mengkaji keterkaitan antara bentuk dan makna’, *Linguist. Indones.*, vol. 37, no. 1, pp. 13–31, 2019, doi: 10.26499/li.v37i1.87.

- [10] G. P. W. Rajeg, I. M. Rajeg, and I. W. Arka, 'Corpus-based approach meets LFG: the puzzling case of voice alternations of *kena*-verbs in Indonesian', in *Proceedings of the LFG'20 conference, on-line*, Stanford, 2020, pp. 307–327. doi: 10.6084/m9.figshare.12423788.
- [11] I. M. Rajeg, G. P. W. Rajeg, and I. W. Arka, 'Corpus linguistic and experimental studies on the meaning-preserving hypothesis in Indonesian voice alternations', *Linguist. Vanguard*, vol. 8, no. 1, pp. 1–16, 2022, doi: 10.1515/lingvan-2020-0104.
- [12] G. P. W. Rajeg and I. M. Rajeg, 'Analisis Koleksem Khas dan potensinya untuk kajian kemiripan makna konstruksional dalam Bahasa Indonesia', in *ETIKA BAHASA Buku persembahan menapaki usia pensiun: I Ketut Tika*, vol. 1, I. N. Sudipa, Ed. Denpasar, Bali, Indonesia: Swasta Nulus, 2019, pp. 65–83. Accessed: Jan. 30, 2019. [Online]. Available: <https://doi.org/10.31227/osf.io/uwzts>
- [13] G. P. W. Rajeg, I. M. Rajeg, and I. W. Arka, 'Contrasting the semantics of Indonesian *-kan* and *-i* verb pairs: A usage-based, constructional approach', in *Prosiding Seminar Nasional Bahasa Ibu XII*, Denpasar, Bali, Indonesia, 2020, pp. 328–344. [Online]. Available: <https://doi.org/10.6084/m9.figshare.12311192>
- [14] H. Y. J. Choi, 'A corpus based analysis of *-kan* and *-i* in Indonesian', Master's thesis, Nanyang Technological University, Singapore, 2019. Accessed: Apr. 01, 2020. [Online]. Available: <https://hdl.handle.net/10356/136955>
- [15] G. P. W. Rajeg, K. Denistia, and S. Musgrave, 'Vector Space Models and the usage patterns of Indonesian denominal verbs: A case study of verbs with *meN-*, *meN-/-kan*, and *meN-/-i* affixes', *NUSA*, vol. 67, pp. 35–75, 2019, doi: 10.15026/94452.
- [16] K. Denistia, E. Shafaei-Bajestan, and R. H. Baayen, 'Exploring semantic differences between the Indonesian prefixes *PE-* and *PEN-* using a vector space model', *Corpus Linguist. Linguist. Theory*, Apr. 2021, doi: 10.1515/cllt-2020-0023.
- [17] K. Denistia and R. H. Baayen, 'The Morphology of Indonesian: Data and quantitative modeling', in *The Routledge handbook of Asian linguistics*, C. Shei and S. Li Rasmussen, Eds. Abingdon, Oxon ; New York, NY: Routledge, 2022. [Online]. Available: <http://www.sfs.uni-tuebingen.de/~hbaayen/publications/DenistiaBaayen2021.pdf>
- [18] R Core Team, 'R: A language and environment for statistical computing', Vienna, Austria, manual, 2020. [Online]. Available: <https://www.R-project.org/>

- [19] H. Wickham *et al.*, ‘Welcome to the Tidyverse’, *J. Open Source Softw.*, vol. 4, no. 43, p. 1686, Nov. 2019, doi: 10.21105/joss.01686.
- [20] S. D. Larasati, V. Kuboň, and D. Zeman, ‘Indonesian Morphology Tool (MorphInd): Towards an Indonesian Corpus’, in *Systems and Frameworks for Computational Morphology*, Aug. 2011, pp. 119–129. doi: 10.1007/978-3-642-23138-4_8.
- [21] H. Nomoto, H. Choi, D. Moeljadi, and F. Bond, ‘MALINDO Morph: Morphological dictionary and analyser for Malay/Indonesian’, in *Proceedings of the LREC 2018 Workshop ‘The 13th Workshop on Asian Language Resources’*, 2018, pp. 36–43. [Online]. Available: http://lrec-conf.org/workshops/lrec2018/W29/pdf/8_W29.pdf
- [22] R. H. Baayen, ‘Corpus linguistics in morphology: Morphological productivity’, in *Corpus linguistics: An international handbook*, vol. 2, Anke Lüdeling and Merja Kytö, Eds. Berlin: Mouton de Gruyter, 2009, pp. 899–919.
- [23] G. P. W. Rajeg and K. Denistia, ‘Distinctive Collexeme Analysis of Indonesian Causative Rival Affixes *per-* and *-kan*’, Jul. 15, 2021. Accessed: Aug. 28, 2021. [Online]. Available: <https://doi.org/10.6084/m9.figshare.13602155>
- [24] I. W. Arka, ‘Dynamic and stative passives in Indonesian & their computational implementation’, presented at the MALINDO Workshop, Jakarta, Aug. 02, 2010.
- [25] D. Biber, S. Conrad, and G. Leech, *Longman Student Grammar of Spoken and Written English*, 1st edition. Harlow: Pearson Education ESL, 2002.
- [26] S. Th. Gries, ‘Elementary statistical testing with R’, in *Research methods in language variation and change*, M. Krug and J. Schlüter, Eds. Cambridge: Cambridge University Press, 2013, pp. 361–381. doi: 10.1017/CBO9780511792519.024.

AFIKSASI NOMINA DALAM BAHASA INDONESIA

¹Karlina Denistia dan ²Gede Primahadi Wijaya Rajeg

¹Sekolah Vokasi, Universitas Sebelas Maret

²Universitas Udayana

¹karlinadenistia@staff.uns.ac.id; ²primahadi_wijaya@unud.ac.id

Abstrak

Makalah ini mendeskripsikan pembentukan nomina dalam bahasa Indonesia melalui proses afiksasi berbasis korpus Tata Bahasa Indonesia Kontemporer yang berjumlah sekitar 18 juta kata. Pendekatan yang digunakan dalam makalah ini mencakup pendekatan kualitatif dan kuantitatif. Secara kualitatif, kami mendeskripsikan pembentukan kata benda dari segi bentuk, makna, dan kelas-kata kata-dasar. Secara kuantitatif, kami melakukan analisis produktivitas setiap imbuhan pembentuk kata kerja, analisis perkembangan produktivitas imbuhan per tahun, dan distribusi kelas-kata kata-dasar untuk lima imbuhan pembentuk nomina yang paling produktif: *-an*, *ke-/an*, *peN-*, *peN-/an*, dan *per-/an*. Makalah ini juga mengangkat topik dua imbuhan pembentuk kata benda yang mirip secara bentuk, namun berbeda secara kuantitatif, yaitu imbuhan *pe-* dan *peN-*. Dengan menggunakan pendekatan korpus, kami menemukan beberapa makna baru yang tercipta dari pembentukan nomina melalui proses afiksasi dalam bahasa Indonesia. Selain itu, data menunjukkan bahwa afiks *ke-/an*, *peN-/an*, dan *-an* merupakan afiks pembentuk nomina yang paling produktif dibandingkan afiks lainnya. Pendekatan kuantitatif juga memungkinkan kami untuk mendeskripsikan preferensi kelas-kata kata-dasar pada afiks pembentuk nomina. Temuan kami menunjukkan bahwa imbuhan *ke-/an* memiliki kecenderungan berkata-dasar kata sifat, imbuhan *per-/an* memiliki representasi berlebih untuk kata-dasar kata benda, dan kata kerja memiliki representasi berlebih sebagai kata dasar imbuhan *-an*, *peN-*, dan *peN-/an*. Penelitian ini mengimplikasikan sinergi penerapan metode linguistik korpus, linguistik kuantitatif, dan linguistik kualitatif

Kata kunci: linguistik korpus, analisis kuantitatif, produktivitas, distribusi kata dasar

1. PENDAHULUAN

Nomina, atau bisa juga disebut kata benda, adalah kelas kata yang secara tidak dapat dinegasikan dengan kata *tidak*. Negasi nomina menggunakan kata *bukan*. Kridalaksana [1] menyebutkan bahwa nomina terbagi menjadi dua kategori, yaitu nomina dasar (nomina yang berupa morfem bebas), contohnya: *batu, kertas, radio, udara, kereta*, dan nomina turunan (nomina yang terbentuk dari proses morfologi), contohnya: *keuangan* yang merupakan bentukan dari kata dasar *uang* dan imbuhan ke-/an, *penyala* yang merupakan bentukan dari kata dasar *nyala* dan imbuhan peN-, *pelajaran* yang merupakan bentukan dari kata dasar *ajar* dan imbuhan peN-/an, serta *perluasan* yang merupakan bentukan dari kata dasar *luas* dan imbuhan per-/an. Kridalaksana [1] juga menjelaskan mengenai subkategorisasi kata benda seperti yang ada pada Tabel 1:

Tabel 1 Subkategorisasi kata benda

Subkategorisasi nomina		Contoh kata
Nomina bernyawa dan tak bernyawa	bernyawa	<ul style="list-style-type: none"> • nama diri seperti <i>Joko, Karlina</i> • kekerabatan seperti <i>nenek, kakek</i> • orang atau yang diperlakukan seperti orang, misalnya <i>tuan, nyonya, hantu, raksasa</i> • nama kelompok manusia seperti <i>Jepang, Indonesia</i>
	nomina tak bernyawa	<ul style="list-style-type: none"> • nama geografis seperti <i>Inggris, Bali, Jawa, utara</i> • nama lembaga seperti <i>DPR, MPR</i> • waktu seperti <i>senin, Selasa, Januari, pukul 8</i> • ukuran dan takaran seperti <i>kali, pikul, goni, lusin</i> • tiruan bunyi seperti <i>aum, dengung, kokok</i>
Nomina terbilang dan tak terbilang	nomina terbilang	dapat dihitung atau diikuti oleh kata bilangan seperti <i>buah, orang, helai, dan lembar</i> .
	nomina tak terbilang	tidak dapat diikuti oleh kata bilangan. Contoh nomina tak terbilang: <i>udara, kebersihan, kemanusiaan</i> , termasuk pula <i>nama diri</i> dan <i>nama geografis</i>

Nomina kolektif dan bukan kolektif	nomina kolektif (dapat diperinci bagian-bagiannya)	<ul style="list-style-type: none"> ● nomina dasar seperti <i>tentara, keluarga</i> ● nomina turunan seperti <i>wangi-wangian, tepung-tepungan, minuman</i>
	nomina bukan kolektif (nomina yang tidak diperinci bagiannya)	<i>asinan, cairan, hadirin, keluarga, kawan, kelompok</i>

Selain tidak bisa dinegasikan dengan kata *tidak*, nomina memiliki beberapa ciri tambahan, yaitu (1) berfungsi sebagai subjek maupun objek dalam sebuah kalimat, (2) dapat dimodifikasi dengan menambahkan adjektiva di belakang kata benda, (3) dapat dimodifikasi dengan menambahkan kata tunjuk *ini* dan *itu* di awal kata benda, (4) dapat dimodifikasi dengan menambahkan kata bilangan di awal kata benda terbilang, (5) dapat dimodifikasi dengan menambahkan *quantifier* (misalnya *banyak, beberapa*) di awal kata benda.

Artikel ini membahas pembentukan nomina melalui proses imbuhan (afiksasi). Dalam artikel ini, kami akan menjelaskan (1) jenis kata dasar yang melekat pada afiksasi pembentuk nomina, (2) makna kata bentukan yang ditemukan di dalam korpus, (3) frekuensi kemunculan dari setiap afiks, dan (4) preferensi kata dasar pada setiap afiks.

Kata yang melekat pada imbuhan merupakan kata dasar, dan hasil dari pembentukan nomina dengan proses afiksasi dinamakan nomina turunan [2]. Misalnya, nomina turunan *pembangunan* memiliki awalan *peN-* yang melekat pada kata dasar *bangun*. Imbuhan yang dibahas dalam tulisan ini adalah awalan *pe-* dan *peN-*, akhiran *-an*, *-wan*, dan *-wati*, serta gabungan awalan dan akhiran *peN-/an*, *per-/an*, dan *ke-/an*.

Pembentukan nomina turunan telah banyak diteliti oleh linguist Indonesia. Kridalaksana [1] dan Moeliono et al. [2] menjelaskan mengenai gambaran umum terkait jenis, ciri, dan klasifikasi kata benda. Pembentukan kata benda melalui proses afiksasi dibahas secara umum oleh Sneddon et al. [3] dan Moeliono et al. [2]. Pada dasarnya, mereka menyebutkan bahwa afiksasi pembentuk nomina di antaranya adalah *ke-* seperti pada *kekasih*, *-an* seperti pada *anjuran*, *peN-* seperti pada *peninju*, *pe-* seperti pada *petinju*, *per-* seperti pada *pertanda*, *ke-/an* seperti pada *ketinggian*, *per-/an* seperti pada *persatuan*, *peN-/an* seperti pada *penyatuan*, *-wan* seperti pada *karyawan*, dan *-wati* seperti pada *karyawati*. Setiap afiks punya makna tersendiri, bahkan tidak menutup kemungkinan bahwa satu afiks memiliki lebih dari satu makna.

Beberapa artikel lebih spesifik membahas imbuhan-imbuhan pembentuk nomina dalam bahasa Indonesia [4]–[7]. Tabel 2 merangkum beberapa imbuhan pembentuk nomina dari segi bentuk dan maknanya. Klasifikasi imbuhan *pe-*, *per-*, dan *peN-* pada Tabel 2 masih mengikuti paparan deskriptif kualitatif berdasarkan sumber artikel yang kami jadikan acuan. Akan ada pembahasan lebih lanjut mengenai status morfologi imbuhan *pe-* dan *peN-* pada subbab “Imbuhan *pe-* dan *peN-*” di artikel ini.

Tabel 2 Afiksasi pembentuk nomina dalam bahasa Indonesia

Imbuhan	Kata Dasar	Kata Turunan	Makna Imbuhan
<i>pe-</i>	kebun	pekebun	agen
	tinju	petinju	atlet
	suruh	pesuruh	pasien
<i>peN-</i>	tinju	peninju	agen
	buka	pembuka	instrumen
	dengar	(alat) pendengar	agen dan instrument
	malu	pemalu	agen dengan sifat X
<i>per-</i>	tapa	pertapa	agen
<i>-an</i>	lukis	lukisan	hasil meN-X
	ayun	ayunan	instrumen untuk meN-X
	belok	belokan	lokasi meN-X
	sayur	sayuran	kumpulan X
<i>-wan</i>	karya	karyawan	agen laki-laki
	budaya	budayawan	laki-laki yang ahli dalam bidang X
<i>-wati</i>	karya	karyawati	agen wanita
<i>ke-/ -an</i>	seimbang	keseimbangan	keadaan X
	pemimpin	kepemimpinan	berkaitan dengan X
	kaisar	kekaisaran	wilayah X
	pulau	kepulauan	kumpulan X
	pergi	kepergian	hasil dari X
<i>per-/ -an</i>	tolong	pertolongan	kegiatan meN-X
	gerak	pergerakan	kegiatan ber-X
	minta	permintaan	hasil meN-X
	cetak	percetakan	tempat meN-X
	budak	perbudakan	terkait dengan X
	istilah	peristilahan	sekumpulan X
<i>peN-/ -an</i>	tulis	penulisan	proses meN-X
	cuci	pencucian	tempat meN-X
	temu	penemuan	hasil meN-X

Selain imbuhan yang ada pada Tabel 2, ada akhiran asing *-isme* (*sukuisme*), *-isasi* (*kaderisasi*), *-logi* (*biologi*), dan *-tas* (*spontanitas*) yang

juga berfungsi sebagai pembentuk nomina dalam bahasa Indonesia.

Dalam makalah ini, kami ingin mendeskripsikan proses pembentukan kata benda dari segi bentuk dan makna untuk afiks *pe-*, *peN-*, *-wan*, *-wati*, *per-/an*, *peN-/an*, dan *ke-/an*. Selain itu, kami menggunakan metode kuantitatif untuk dapat menjawab pertanyaan: (1) bagaimana produktivitas afiks pembentuk nomina dalam bahasa Indonesia?, dan (2) bagaimana distribusi kata-dasar kelas-kata untuk afiks pembentuk nomina dalam bahasa Indonesia?.

Salah satu kebaruan yang ada pada penelitian ini adalah bahwa contoh dan analisis data yang disajikan berasal dari korpus bahasa Indonesia yang berukuran sekitar 18 juta kata. Sepengamatan kami, penelitian imbuhan bahasa Indonesia dengan menggunakan korpus mulai banyak dilakukan sejak tahun 2018. Misalnya, penelitian yang membahas perbedaan imbuhan *pe-* dan *peN-*, seperti pada kata *pelajar* dan *pengajar*; dari segi produktivitas, linearitas dengan imbuhan *meN-*, distribusi semantik, dan proses pembelajaran mesin [8]–[11]. Selain itu, penelitian kebahasaan menggunakan korpus mampu menjelaskan perbedaan verba berimbuhan *meN-*, *meN-/kan*, dan *meN-/i* untuk kata dasar yang melekat pada ketiga imbuhan tersebut. Data korpus yang dianalisis menggunakan metode kuantitatif menunjukkan bahwa penggunaan ketiga imbuhan ini sama-sama berada di domain (1) pergerakan kaki, misalnya *melangkah-melangkahkan-melangkahi*, *menjejak-menjeakkan-menjejaki*, dan *menapak-menapakkan-menapaki*, (2) perasaan, misalnya *menyesal-menyesali-menyesalkan* [12].

Penelitian ini fokus pada afiksasi turunan untuk imbuhan-imbuhan yang produktif, yaitu *pe-*, *peN-*, *-wan*, *-wati*, *per-/an*, *peN-/an*, dan *ke-/an*. Penggunaan korpus dalam penelitian ini memperbesar kemungkinan munculnya contoh pemakaian imbuhan yang baru dan belum pernah dibahas sebelumnya. Contoh yang disajikan dalam tulisan ini pun merupakan contoh nyata, tidak dibuat-buat, dan memang didasarkan pada pemakaian bahasa tulis yang alami. Selain itu, data korpus juga dapat digunakan untuk mendeskripsikan tipe-tipe, struktur, fungsi dan makna imbuhan dalam menghasilkan kata nomina dalam bahasa Indonesia. Kami berharap bahwa analisis yang ada pada artikel ini dapat memberikan penyegaran dalam tata bahasa bahasa Indonesia terkait proses pembentukan kata benda.

Dengan adanya jumlah data yang besar, kami dapat melakukan analisis kuantitatif untuk memperkuat analisis kualitatif pembentukan nomina dalam bahasa Indonesia. Contohnya, data kuantitatif yang berupa angka dapat diolah untuk membandingkan distribusi dan produktivitas setiap afiks. Kemudian,

distribusi dan produktivitas setiap afiks dapat dibandingkan berdasarkan tahunnya. Dengan demikian, kita dapat membandingkan afiks mana yang penggunaannya cenderung meningkat, stagnan, atau bahkan menurun. Hal lain yang dapat dianalisis menggunakan pendekatan kuantitatif adalah distribusi kata dasar. Tidak menutup kemungkinan bahwa sebuah afiks akan memiliki kecenderungan melekat pada kata dasar tertentu. Misalnya, Karlina et al. [10] dan Ramlan [13] menyebutkan bahwa imbuhan *pe-* cenderung melekat pada kata benda yang yang terkait dengan olah raga, seperti pada kata *pegolf*, *petenis*, dan *perenang*.

Pembahasan selanjutnya akan mencakup metode pengumpulan dan pengolahan data, deskripsi bentuk dan makna imbuhan, dan distribusi imbuhan berbasis korpus. Tulisan ini akan ditutup dengan kesimpulan dan saran.

2. METODE PENGUMPULAN DAN PENGOLAHAN DATA

Data berupa kalimat yang dianalisis dan dijadikan contoh dalam penelitian ini diambil dari korpus Tata Bahasa Indonesia Kontemporer yang dikumpulkan oleh Pusat Pengembangan dan Pelindungan Bahasa dan Sastra, Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi dari tahun 2011 sampai 2020. Sumber data dari korpus ini berupa kalimat yang diambil dari koran, majalah, cerpen, novel, buku teks, jurnal, skripsi, tesis, disertasi, biografi, tulisan populer, undang-undang, laman formal, dan surat formal. Korpus yang digunakan dalam penelitian ini berjumlah sekitar 14.000 kalimat yang terdiri dari sekitar 18.000.000 token kata, 436.000 tipe kata, dan 98.000 *hapax legomena* (kata yang muncul sekali di dalam korpus; [14]).

Produktivitas imbuhan dapat ditentukan dengan melihat jumlah frekuensi token (total penggunaan atau kemunculan imbuhan di dalam korpus), frekuensi tipe (jumlah kata unik yang ada di dalam korpus), dan *hapax legomena* (jumlah kata yang muncul satu kali di dalam korpus). Perhitungan token, tipe, dan *hapax legomena* yang kami lakukan tidak memasukkan angka dan tanda baca. Seluruh kapitalisasi yang ada di dalam korpus juga kami ubah menjadi huruf kecil. Sebagai gambaran sederhana, dalam kalimat „Aku cinta kamu karena kamu cinta aku“ terdapat 7 token yaitu `aku`, `cinta`, `kamu`, `karena`, `kamu`, `cinta`, `aku`, 4 tipe yaitu `aku`, `cinta`, `kamu`, `karena`, dan 1 *hapax legomena* yaitu `karena`.

Karena penelitian ini fokus pada pembentukan nomina dalam bahasa Indonesia, kami mengambil data kata benda yang ada di dalam korpus Tata Bahasa Indonesia Kontemporer. Kami menggunakan piranti *MorphInd* [15] untuk menyaring data berupa kata benda. Secara keseluruhan, data nomina yang terkumpul berjumlah 6.362.290 token kata yang terdiri dari 15.352 tipe kata dan 3.108 *hapax legomena*. Setelah data terkumpul, kami melakukan pemisahan antara kata dasar dengan imbuhan untuk kata non-majemuk secara komputasi menggunakan *MorphInd* [15]. Dari luaran *MorphInd*, kami juga memperoleh informasi mengenai kata dasar, kelas-kata kata-dasar, imbuhan, dan klitik yang ada pada sebuah kata turunan. Untuk kelas-kata kata-dasar, kami menggunakan kata dasar yang berupa morfem tidak terikat. Dengan demikian, kata dasar dari ‘pembelajaran’ adalah ‘ajar’ dan bukan ‘belajar’. Khusus untuk imbuhan *pe-* dan *peN-*, kami menambahkan informasi alomorfi dan makna kata secara manual. Hasil luaran *MorphInd* yang kemudian diolah menjadi data kata benda dapat dilihat pada Gambar 1.

Kami juga menghitung frekuensi penggunaan kata di setiap genre teks (lihat Gambar 2), frekuensi penggunaan kata di setiap tahun (lihat Gambar 3), serta total frekuensi penggunaan setiap kata yang ada di korpus (lihat Gambar 3, kolom ‘Freq’). Seluruh proses pengumpulan dan pengolahan data dilakukan dengan menggunakan bahasa pemrograman R (versi 4.2.1) di R Studio [16]. Data yang digunakan dalam penelitian ini dapat diakses secara gratis melalui tautan <https://osf.io/84zfw/>.

Word	MorphInd	BaseWord	BaseWordClass	Affix	MorphologicalVariation
ajaran	ajar<v>+an_NSD	ajar	v	-an	
ajaranku	ajar<v>+an_NSD+aku<p>_PS1	ajar	v	-an	Poss1stSing
ajaranmu	ajar<v>+an_NSD+kamu<p>_PS2	ajar	v	-an	Poss2ndSing
ajarannya	ajar<v>+an_NSD+dia<p>_PS3	ajar	v	-an	Poss3rdSing
pengajar	peN+ajar<v>_NSD	ajar	v	peN-	
pengajaran	peN+ajar<v>+an_NSD	ajar	v	peN-/ -an	
pengajaranku	peN+ajar<v>+an_NSD+aku<p>_PS1	ajar	v	peN-/ -an	Poss1stSing
pengajarannya	peN+ajar<v>+an_NSD+dia<p>_PS3	ajar	v	peN-/ -an	Poss3rdSing
pengajarku	peN+ajar<v>_NSD+aku<p>_PS1	ajar	v	peN-	Poss1stSing
pengajarlaha	peN+ajar<v>_NSD+lah<t>_T--	ajar	v	peN-	Particle
pengajarnya	peN+ajar<v>_NSD+dia<p>_PS3	ajar	v	peN-	Poss3rdSing

Gambar 1 Sampel data kata benda yang ada di korpus
Tata Bahasa Indonesia Kontemporer yang berasal dari luaran *MorphInd*

Word	FreqKoran	FreqMajalah	FreqCerpren	FreqNovel	FreqBukuteks	FreqJurnal	FreqTA	FreqBiografi	FreqPopuler	FreqU	FreqLaman	FreqSurat
ajaran	66	43	35	36	365	99	257	397	122	4	183	14
ajaranaku	0	0	0	1	0	0	0	0	0	0	0	0
ajaranmu	0	0	0	0	0	0	0	1	0	0	0	0
ajarannya	1	2	2	0	9	1	4	36	2	0	2	0
pengajar	23	46	6	27	43	26	40	67	66	0	180	8
pengajaran	5	9	1	1	117	84	55	63	119	8	54	29
pengajaranaku	0	0	0	2	0	0	0	0	0	0	0	0
pengajarannya	0	0	0	0	0	0	4	2	2	0	1	0
pengajarku	0	0	0	0	0	0	0	2	0	0	0	0
pengajarlah	0	1	0	0	0	0	0	0	0	0	0	0
pengajarnya	0	5	0	0	2	0	2	8	0	0	1	0

Gambar 2 Sampel data frekuensi kata benda berdasarkan genre teks yang ada di korpus Tata Bahasa Indonesia Kontemporer

Word	Freq	Freq2011	Freq2012	Freq2013	Freq2014	Freq2015	Freq2016	Freq2017	Freq2018	Freq2019	Freq2020
ajaran	1541	9	21	10	8	5	8	9	1	6	4
ajaranaku	1	0	0	0	0	0	0	0	0	1	0
ajaranmu	1	1	0	0	0	0	0	0	0	0	0
ajarannya	59	1	1	1	2	3	2	0	1	6	0
pengajar	452	2	3	2	2	4	1	1	3	2	3
pengajaran	545	2	1	1	9	1	57	2	2	1	1
pengajaranaku	2	2	0	0	0	0	0	0	0	0	0
pengajarannya	9	0	1	0	1	1	0	2	1	1	1
pengajarku	2	0	0	0	0	0	0	0	2	0	0
pengajarlah	1	0	0	1	0	0	0	0	0	0	0
pengajarnya	18	1	2	0	1	2	1	1	2	1	2

Gambar 3 Sampel data total frekuensi dan frekuensi kata benda berdasarkan tahun yang ada di korpus Tata Bahasa Indonesia Kontemporer

3. AFIKS PEMBENTUK NOMINA

3.1 Imbuhan *pe-*

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *pe-* adalah sebagai pembentuk agen, atlet, dan pasien. Berikut contoh fungsi tersebut yang ada di dalam korpus:

- pe-* + tugas = petugas
Petugas medis setempat menyatakan, pekan lalu sebanyak 200 orang tewas di Mistrata. (sumber: koran, 2011)
- pe-* + renang = perenang
Perenang putri lainnya, Ressa Kania Dewi merebut medali emas di nomor 400. (sumber: koran, 2014)
- pe-* + ajar = pelajar
Mudah-mudahan semua *pelajar* di Pamekasan yang akan mengikuti UN bisa lulus dengan nilai yang memuaskan. (sumber: koran, 2014)

Selain itu, kami juga menemukan makna imbuhan *pe-* yang lain. Di antaranya instrumen dan makna kiasan 'orang yang berada di posisi X':

- d. *pe-* + layar = pelayar
Ketentuan pidana dalam undang-undang Indonesia berlaku bagi nakhoda dan penumpang alat *pelayar* (kapal, perahu). (sumber: buku teks, 2014)
- e. *pe-* + tinggi = petinggi
Pada saat yang sama, *petinggi* klub justru terkesan memiliki rencana untuk melepas sejumlah pemain bintangnya. (sumber: koran, 2014)

Imbuhan *pe-* dapat melekat pada kata kerja, kata benda, dan kata sifat, misalnya:

- f. *pe-* + lari (kata kerja) = pelari
Sampai *pelari* kedua, jarak antara *pelari* NTB dan *pelari* tim lainnya masih dekat. (sumber: koran, 2012)
- g. *pe-* + golf (kata benda) = pegolf
Kehebatan *pegolf* dari Amerika Serikat, Tiger Woods, sekarang berada di bawah Rory (sumber: koran, 2012)
- h. *pe-* + mabuk (kata sifat) = pemabuk
Setengahnya warganya *pemabuk*, setengahnya lagi penjudi (sumber: cerpen, 2015)

3.2 Imbuhan *peN-*

Imbuhan *peN-* memiliki 6 alomorf atau varian, yaitu *peN_{peng-}*, *peN_{pen-}*, *peN_{pe-}*, *peN_{peny-}*, dan *peN_{peng-}*. Kemunculan alomorf tersebut dapat diprediksi secara fonologi sebagai berikut [17], [18]:

- a. *peN_{peng-}* muncul ketika melekat pada kata dasar yang berawalan dengan fonem vokal atau fonem *velar obstruent* /g/, /k/, /h/, or /kh/,
i. *peN-* + huni = penghuni
i. *peN-* + gemar = penggemar
- b. *peN_{pen-}* muncul ketika melekat pada kata dasar yang berawalan dengan fonem alveolar obstruent atau palatal obstruent /d/, /t/, /c/, /j/, /sy/, or /z/,
i. *peN-* + tonton = penonton
i. *peN-* + jual = penjual
- c. *peN_{pe-}* muncul ketika melekat pada kata dasar yang berawalan dengan fonem konsonan *labial* /b/, /p/, atau /f/,
i. *peN-* + balap = pembalap
i. *peN-* + pantau = pemantau
- d. *peN_{pe-}* muncul ketika melekat pada kata dasar yang berawalan dengan

fonem nasal, *semivowel*, atau *liquid* /m/, /n/, /ng/, /ny/, /w/, /j/, /r/, or /l/,

ĩ *peN-* + makan = pemakan

ĩ *peN-* + luncur = peluncur

- e. *peN_{peny}-* muncul ketika melekat pada kata dasar yang berawalan dengan fonem /s/,

ĩ *peN-* + saring = penyaring

ĩ *peN-* + sedap = penyedap

- f. *peN_{penge}-* muncul ketika melekat pada kata dasar yang memiliki satu suku kata.

ĩ *peN-* + tes = pengetes

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *peN-* adalah sebagai pembentuk agen, instrumen, dan orang dengan sifat X. Berikut contoh fungsi tersebut yang ada di dalam korpus:

- a. *peN-*+ latih = pelatih

Mantan *pelatih* tim nasional Benny Dolo menilai banyak kelemahan dari penampilan tim nasional senior. (sumber: koran, 2011)

- b. *peN-*+ bangkit = pembangkit

Mereka juga diminta memaksimalkan *pembangkit* listrik sendiri selama masa perbaikan ini. (sumber : koran, 2013)

- c. *peN-*+ diam = pendiam

Sikap Fadlun *pendiam*, berbeda dengan istrinya. (sumber : koran, 2013)

Selain itu, kami juga menemukan makna imbuhan *peN-* yang lain. Di antaranya `penyebab X`, `lokasi X`, dan `musim X`:

- d. *peN-*+ sakit = penyakit

Campak konon merupakan *penyakit* yang muncul tiap dua atau tiga tahun sekali. (sumber: koran, 2011)

- e. *peN-*+ ujung = pengujung

Nikmatnya „Berhaji bersama nabi“ umrah di *pengujung* Ramadan bisa dipastikan selalu penuh perjuangan. (sumber: majalah, 2015)

- f. *peN-*+ hujan = penghujan

Pada musim *penghujan*, ia harus banyak bersabar dan berdoa. (sumber: cerpen, 2012)

Imbuhan *peN-* dapat melekat pada kata kerja, kata benda, dan kata sifat, misalnya:

- g. *peN-*+ tumbuh (kata kerja) = penumbuh
Dia akan meneliti apakah kedua spesies ini kehilangan gen *penumbuh* tulang yang sama. (sumber: koran, 2014)
- h. *peN-*+ program (kata benda) = pemrogram
Diusahakannya pembuatan situs tentang Tiongkok, walaupun ia bukan *pemrogram*. (sumber: majalah, 2017)
- i. *peN-*+ busuk (kata sifat) = pembusuk
Ini mengindikasikan bahwa tanin dapat menekan pertumbuhan dari mikroba-mikroba *pembusuk* pada silase tersebut (sumber: buku teks, 2019)

3.3 Imbuhan *-an*

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *-an* adalah sebagai pembentuk kata turunan dengan makna 'hasil meN-X', 'instrumen untuk meN-X', 'lokasi meN-X', dan 'kumpulan X'. Berikut contoh fungsi tersebut yang ada di dalam korpus:

- a. hitung + *-an* = hitungan
Hanya dalam *hitungan* jam, PM Israel Benyamin Netanyahu langsung terbang ke Washington. (sumber: koran, 2011)
- b. angkut + *-an* = angkutan
Pada saat ini banyak daerah tidak memiliki *angkutan* pedesaan lantaran pada mati, kalah bersaing dengan sepeda motor. (sumber: koran, 2012)
- c. landas + *-an* = landasan
Kedua pesawat ini memiliki kelebihan bisa mendarat di *landasan* pendek dengan panjang 2.200-2.300 meter. (sumber: koran, 2011)

Selain itu, kami juga menemukan makna imbuhan *-an* yang lain. Di antaranya 'bilangan kelipatan X', 'bentuk yang menyerupai X', 'penyakit karena X', 'sesuatu yang memiliki rasa X':

- d. puluh + *-an* = puluhan
Di gang itu tampak berjejer *puluhan* lapak pedagang dengan lebar 1,5 meter. (sumber: koran, 2012)
- e. gunung + *-an* = gunungan
Setelah kerangka *gunungan* dari bambu terbentuk, kerangka dimuati dengan hasil bumi. (sumber: majalah, 2012)

- f. cacing + -an = cacingan
 “Dia tidak kurang gizi, Kawan. Dia *cacingan*.”, temannya menimpali.
 (sumber: novel, 2018)
- g. manis + -an = manisan
 Seperti salak, nangka diolah jadi keripik. Produk lain berupa *manisan* carica, *manisan* nangka, *manisan* nanas, dan *manisan* cocktail dalam kemasan kaleng. (sumber: biografi, 2018)

Imbuhan *-an* dapat melekat pada kata kerja, kata benda, dan kata sifat, misalnya:

- h. jinjing (kata kerja) + -an = jinjingan
 Pertama kali yang dilakukan adalah meletakkan tas dan *jinjingan* kecil yang terbuat dari rotan di atas meja bulat di teras dalam. (sumber: biografi, 2014)
- i. dayung (kata benda) + -an = dayungan
 Hal tersebut dikarenakan *dayungan* lengan akan menghasilkan laju tubuh dengan cepat. (sumber: buku teks, 2018)
- j. asin (kata sifat) + -an = asinan
 Ibu masih bercerita tentang apa saja. Tangannya dengan terampil mulai membuat *asinan* belimbing. (sumber: cerpen, 2020)

3.4 Imbuhan *-wan*

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *-wan* adalah sebagai pembentuk kata turunan dengan makna ‘agen laki-laki’ dan ‘laki-laki yang ahli di bidang X’. Berikut contoh fungsi tersebut yang ada di dalam korpus:

- a. wisata + -wan = wisatawan
 Banyaknya *wisatawan* yang menginap dan mendirikan tenda semakin mengotori pantai. (sumber: koran, 2013)
- b. biologi + -wan = biologiwan
 Penemuan nukleus ini telah diperagakan oleh seorang ahli *biologiwan* dari Jerman yang bernama Theodor Boveri. (sumber: buku teks, 2017)

Selain itu, kami juga menemukan makna imbuhan *-wan* yang lain. Di antaranya `laki-laki yang gemar melakukan X`:

c. puasawan

Lebaran tinggal empat hari lagi. Semua wajah *puasawan* tampak lebih berseri. (sumber: cerpen, 2014)

Imbuhan *-wan* dapat melekat pada kata benda dan kata sifat, misalnya:

d. industri (kata benda) + *-wan* = industriwan

Dengan kebijakan baru di bidang energi itu, pemerintah, *industriwan*, pedagang, dan konsumen harus belajar beradaptasi. (sumber: koran, 2015)

e. sukarela (kata sifat) + *-wan* = sukarelawan

Pada awal sejarah vaksin, *sukarelawan* pertama adalah James Phipps. (sumber: majalah, 2020)

3.5 Imbuhan *-wati*

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *-wati* adalah sebagai pembentuk kata turunan dengan makna `agen wanita`. Berikut contoh fungsi tersebut yang ada di dalam korpus:

a. karya + *-wati* = karyawati

“Saya juga akan menganjurkan Anda untuk berpikir-pikir lagi.” *Karyawati* itu berpikir-pikir dan membatalkan niatnya. (sumber: biografi, 2014)

Selain itu, kami juga menemukan makna imbuhan *-wati* yang lain. Di antaranya `wanita yang ahli dalam bidang X`:

b. seni + *-wati* = seniwati

..., karyawan karyawati, pramugara pramugari, peragawan peragawati, seniman *seniwati*, wartawan wartawati, olahragawan olahragawati... (sumber: novel, 2015)

Imbuhan *-wati* dapat melekat pada kata benda dan kata sifat, misalnya:

c. olah raga + *-wati* =

olahragawati

..., karyawan karyawati, pramugara pramugari, peragawan peragawati, seniman seniwati, wartawan wartawati, olahragawan *olahragawati*... (sumber: novel, 2015)

- d. sukarela (kata sifat) + -wati = sukarelawati
Lalu meneruskan pkursus pendidikan guru sambil menjadi guru *sukarelawati*. (sumber: cerpen, 2019)

3.6 Imbuhan *per-/an*

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *per-/an* adalah sebagai pembentuk kata turunan dengan makna `kegiatan meN-X`, `kegiatan ber-X`, `hasil meN-X`, `tempat melakukan X`, `terkait dengan X`, dan `sekumpulan X`. Berikut contoh fungsi tersebut yang ada di dalam korpus:

- a. *per-* + siap + *-an* = persiapan
Presiden meminta laporan *persiapan* SEA Games yang tinggal satu bulan lagi digelar. (sumber: koran, 2011)
- b. *per-* + buru + *-an* = perburuan
Hasil operasi 2013 ini menunjukkan adanya peningkatan ancaman *perburuan* terhadap harimau Sumatera yang sangat signifikan dibanding 2012 dan 2011. (sumber: koran, 2013)
- c. *per-* + tanya + *-an* = pertanyaan
Komunikasi digital mendorong siapa saja mendapat jawaban secepatnya. Orang cenderung mengajukan *pertanyaan* “lebay“. Orang menggapangkan proses komunikasi. (sumber: koran, 2012)
- d. *per-* + belanja + *-an* = perbelanjaan
Sengaja memilih tempat di pusat *perbelanjaan*, mereka mengaku tak ingin karyanya asing di tengah-tengah masyarakat. (sumber: koran, 2013)
- e. *per-* + politik + *-an* = perpolitikan
Anomali tersebut sejatinya merupakan kebiasaan yang telah mengendap dalam urat nadi *perpolitikan* negeri ini. (sumber: koran, 2011)
- f. *per-* + desa + *-an* = perdesaan
... dan sanitasi lingkungan, yang diarahkan untuk meningkatkan sarana dan prasarana lingkungan *perdesaan* dan perkotaan. (sumber: perundangan, 2013)

Imbuhan *per-/an* dapat melekat pada kata kerja, kata benda, dan kata sifat, misalnya:

g. *per-* + istirahat (kata kerja) + *-an* =
peristirahatan

Mungkin hanya rumah *peristirahatan*, semacam villa yang didatangi pemiliknya sekitar sebulan atau dua bulan sekali. (sumber: cerpen, 2015)

h. *per-* + senjata (kata benda) + *-an* =
persenjataan

Dengan persenjataan mutakhir, satu per satu mereka menaklukkan penguasa setempat. (sumber: novel, 2011)

i. *per-* + adil (kata sifat) + *-an* = peradilan

Saat ini belum ada ketentuan yang mengatur pemberlakuan peradilan anak bagi mereka yang mengalami disabilitas mental. (sumber: koran, 2014)

3.7 Imbuhan *peN-/an*

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *peN-/an* adalah sebagai pembentuk kata turunan dengan makna 'proses meN-X', 'tempat meN-X', dan 'hasil meN-X'. Berikut contoh fungsi tersebut yang ada di dalam korpus:

a. *peN-* + pilih + *-an* = pemilihan

... menyebutkan kecil peluang Presiden Barack Obama untuk terpilih lagi dalam *pemilihan* presiden pada 2012. (sumber: koran, 2011)

b. *peN-* + adil + *-an* = pengadilan

Penyidikan, penuntutan, dan pemeriksaan di sidang *pengadilan* terhadap perkara tidak pidana korupsi dilakukan berdasarkan hukum acara pidana. (sumber: koran, 2011)

c. *peN-* + buku + *-an* = pembukuan

Dokumen itu antara lain berupa keterangan badan hukum dan *pembukuan* keuangan klub. (sumber: koran, 2011)

Korespondensi *meN-* dan *peN-/an* juga termasuk gabungan awalan-akhiran *meN/-kan* dan *meN/-i*. Misalnya, kata *pengadilan* bermakna tempat untuk *mengadili*, serta kata *pembukuan* bermakna hasil dari proses *membukukan*. Selain fungsi yang sudah dijelaskan pada 7a sampai 7c, kami juga menemukan makna imbuhan *peN-/an* yang lain. Di antaranya 'proses memberi X', 'menjadikan X':

- d. *peN*-+ cahaya + *-an* = pencahayaan
Pencahayaan harus memiliki intensitas yang cukup untuk fotosintesa tumbuhan. (sumber: majalah, 2012)
- e. *peN*-+ gawat + *-an* = penggawatan
 ... dan dijalin dengan seksama, yang menggerakkan jalan cerita melalui perumitan (*penggawatan* atau komplikasi) ke arah klimaks penyelesaian. (sumber: buku teks, 2020)

Imbuhan *peN*-/*-an* dapat melekat pada kata kerja, kata benda, dan kata sifat, misalnya:

- f. *peN*-+ intai (kata kerja) + *-an* = pengintaian
 Armand mengatakan, setelah memastikan kebenaran informasi tersebut, penyidik akhirnya melakukan *pengintaian* terhadap kamar 22009. (sumber: koran, 2014)
- g. *peN*-+ lelang (kata benda) + *-an* = pelelangan
 Hasil dari acara ini, baik dari *pelelangan* atau tiket donasi Jelajah Tembang Cinta, digunakan untuk pengobatan, perawatan, ... (sumber: majalah, 2016)
- h. *peN*-+ lambat (kata sifat) + *-an* = pelambatan
 Kondisi perekonomian regional dan global yang mengalami tekanan dan *pelambatan*, antara lain disebabkan oleh wabah COVID-19. (sumber: surat resmi, 2020)

3.8 Imbuhan *ke*-/*-an*

Seperti yang telah dijelaskan pada Tabel 2, makna imbuhan *ke*-/*-an* adalah sebagai pembentuk kata turunan dengan makna 'keadaan X', 'berkaitan dengan X', 'wilayah X', 'kumpulan X', dan 'hasil dari melakukan X'. Berikut contoh fungsi tersebut yang ada di dalam korpus:

- a. *ke* + tertib + *an* = ketertiban
 "Kami juga akan menjerat mereka dengan pasal mengganggu *ketertiban* umum", ujarnya. (sumber: koran, 2015)
- b. *ke* + anggota + *an* = keanggotaan
 Kamis pekan lalu, Tempo menyambangi klub kebugaran dengan tarif *keanggotaan* RP. 3,5 juta per bulan. (sumber: majalah, 2011)
- c. *ke* + aceh + *an* = keacehan

Ia membangun kembali kesadaran akan identitas *keacehan*, sekaligus spirit untuk mewujudkannya. (sumber: biografi, 2011)

d. ke + bohong + an = kebohongan

“Masyarakat bisa mengajukan tuntutan kalau terjadi kebohongan publik yang dilakukan oleh lembaga survei“, kata dia. (sumber: koran, 2014)

Kami tidak menemukan data imbuhan *ke-/-an* yang bermakna ‘kumpulan X’ di dalam korpus. Imbuhan *ke-/-an* dapat melekat pada kata kerja, kata benda, dan kata sifat, misalnya:

e. ke + pindah (kata kerja) + an = kepindahan
Arsenal tampaknya masih terpukul oleh *kepindahan* kapten sekaligus *playmaker* mereka, Cesc Fabregas, ke Barcelona pekan ini. (sumber: koran, 2011)

f. ke + tuhan (kata benda) + an = ketuhanan
Setelah membaca bab ini, Anda diharapkan dapat menjelaskan konsep *ketuhanan* dalam Islam. (sumber: buku teks, 2011)

g. ke + gelap (kata sifat) + an = kegelapan
Sonya sangat membenci kegelapan. Tidur pun dia ingin berada di bawah lampu yang menyala terang. (sumber: cerpen, 2012)

4. ANALISIS KUANTITATIF AFIKS PEMBENTUK NOMINA

Bab ini membahas distribusi imbuhan yang ada di dalam korpus, distribusi kata dasar yang melekat pada setiap imbuhan, dan preferensi kata dasar yang melekat pada imbuhan tertentu. Untuk mengetahui imbuhan mana yang paling produktif di antara imbuhan pembentuk kata benda, kami menggunakan pendekatan kuantitatif dengan menghitung frekuensi penggunaan imbuhan di dalam korpus. Perhitungan ini didasarkan pada gabungan seluruh genre. Untuk melengkapi analisis distribusi afiks, kami juga melakukan analisis terhadap kelas-kata kata-dasar yang melekat pada setiap imbuhan. Analisis kuantitatif kami berfokus pada imbuhan *pe-*, *peN-*, *-wan*, *-wati*, *per-/-an*, *peN-/-an*, dan *ke-/an*.

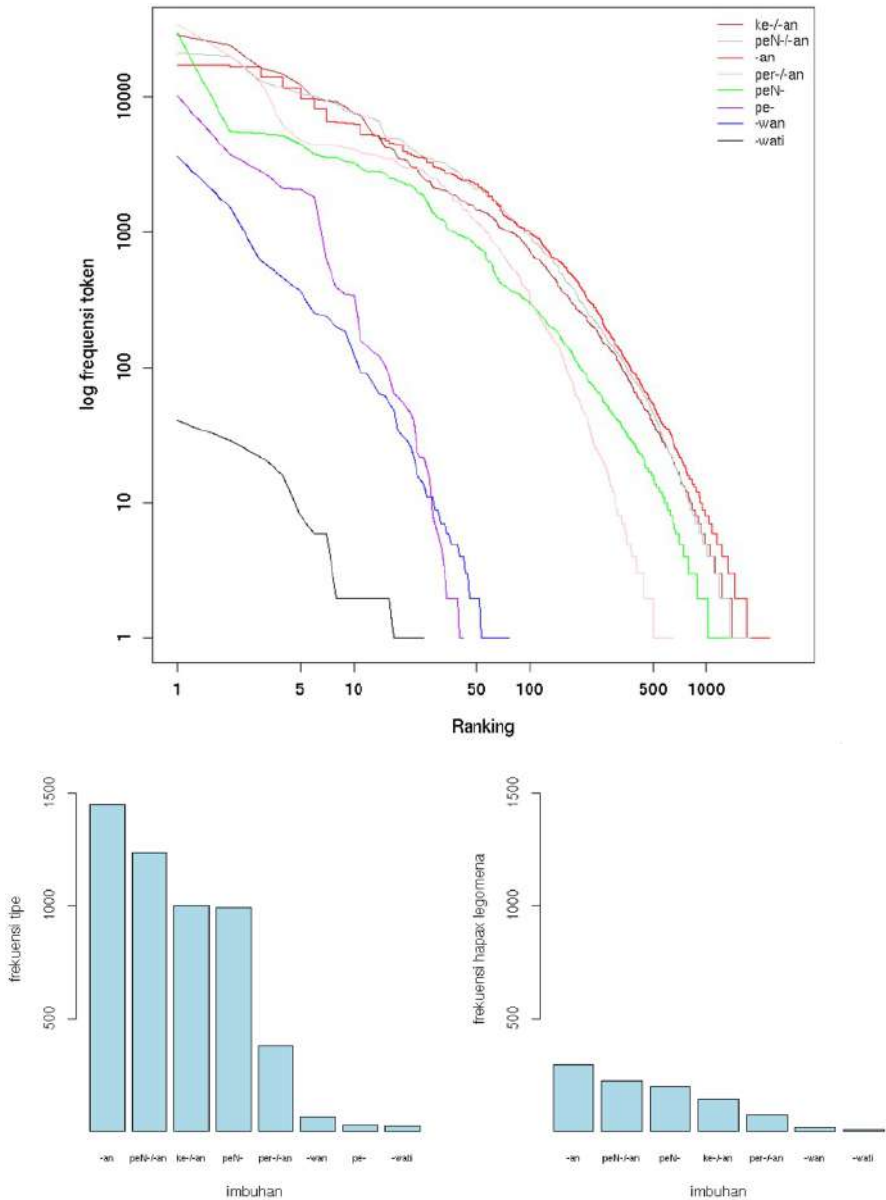
4.1 Distribusi afiks

Setiap imbuhan pembentuk nomina dalam bahasa Indonesia memiliki produktivitas yang berbeda, baik dalam jumlah token, tipe, dan hapax legomena. Gambar 4 menunjukkan bahwa dari perhitungan token dan tipe, frekuensi penggunaan *ke-/an*, *peN-/an*, dan *-an* lebih produktif dibandingkan dengan frekuensi penggunaan imbuhan lainnya. Dari perhitungan *hapax legomena*, imbuhan *peN-* merupakan imbuhan yang paling produktif setelah imbuhan *-an* dan *peN-/an*. Kami menggunakan skala log dalam perhitungan token [19], [20]. Detail perhitungan distribusi setiap imbuhan dapat dilihat pada Tabel 3.

Tabel 3 Perhitungan token, tipe, dan hapax legomena untuk setiap imbuhan pembentuk kata benda dalam bahasa Indonesia

Imbuhan	Token	Tipe	Hapax legomena
<i>peN-/an</i>	436197	1753	406
<i>-an</i>	421011	2298	604
<i>ke-/an</i>	382846	1819	410
<i>per-/an</i>	244593	650	150
<i>peN-</i>	184785	1369	348
<i>pe-</i>	25675	42	3
<i>-wan</i>	8422	77	25
<i>-wati</i>	155	25	9

Untuk mengukur produktivitas imbuhan, kami membagi *hapax legomena* dengan frekuensi token [21]. Tingkat produktivitas menunjukkan besarnya potensi sebuah imbuhan dalam membentuk kata baru. Dengan demikian, semakin tinggi tingkat produktivitas, maka semakin besar pula kemungkinan sebuah kata baru dengan imbuhan tersebut akan terbentuk. Berikut tingkat produktivitas untuk imbuhan pembentuk nomina dalam bahasa Indonesia: *pe-* 0.0001, *per-/an* 0.0006, *peN-/an* 0.0009, *ke-/an* 0.0011, *-an* 0.0014, *peN-* 0.0019, *-wan* 0.0030, dan *-wati* 0.0581. Dilihat dari tingkat produktivitasnya, imbuhan *-wan*, dan *-wati* memiliki potensi yang lebih tinggi untuk menjadi imbuhan yang melekat pada kata dasar tertentu dan membentuk kata baru jika dibandingkan dengan imbuhan *pe-*, dan *per-/an*.



Gambar 4 Kurva produktivitas imbuhan pembentuk nomina dalam hitungan token (panel atas), tipe (panel kiri bawah), dan *hapax legomena* (panel kanan bawah).

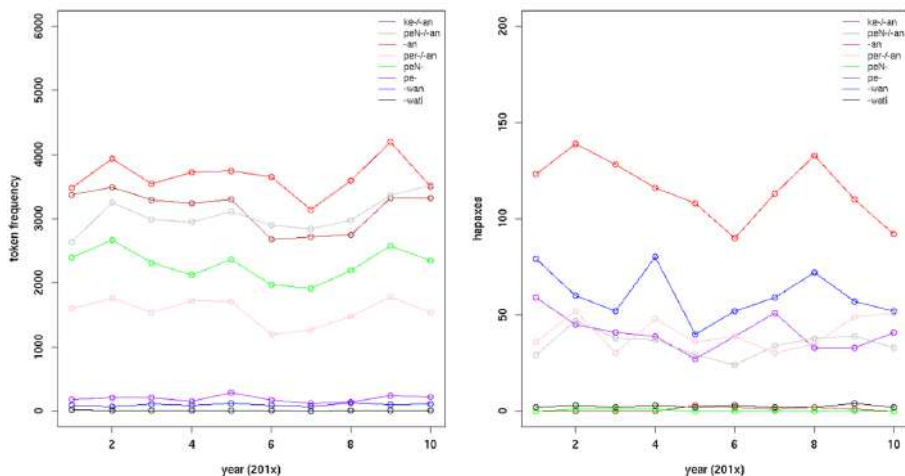
Sebagai gambaran frekuensi imbuhan yang ada di dalam korpus, berikut adalah 10 kata terpilih dengan frekuensi penggunaan tertinggi untuk masing-masing imbuhan:

1. *peN/-an* : penelitian, pendidikan, pelaksanaan, pengembangan, pembelajaran, pelayanan, pembangunan, pembiayaan, peningkatan, penggunaan
2. *-an* : laporan, bagian, lingkungan, tujuan, hubungan, bangunan, jaringan, anggaran, tindakan, makanan
3. *ke/-an* : kegiatan, keuangan, ketentuan, kesehatan, kebijakan, kementerian, kebutuhan, kemampuan, kehidupan, kecamatan
4. *per/-an* : perusahaan, peraturan, perubahan, perkembangan, perbedaan, pertumbuhan, persetujuan, persyaratan, perjalanan, permohonan
5. *peN-* : pemerintah, penyakit, pejabat, penduduk, pemimpin, pemegang, pelaku, pengguna, penyelenggara, peneliti
6. *pe-* : peserta, pekerja, petugas, pelanggan, petani, pedagang, pelajar, pejuang, pesaing, petinggi
7. *-wan* : karyawan, wisatawan, relawan, ilmuwan, bangsawan, wirausahawan, olahragawan, belawan, sukarelawan, budayawan
8. *-wati* : karyawati, santriwati, olahragawati, sukarelawati.

Hapax legomena terpilih untuk setiap imbuhan adalah sebagai berikut:

1. *peN/-an* : pembantingan, pembidanan, pemblokadean, pemonitoran, penabrakan, pendeportasian, pengaspirasian, penggenjotan, pengiprahan, penyetujuan
2. *-an* : ambrukan, seruputan, ringkukan, koyakan, benaman, contekan, zikiran, jentikan, didihan, cabikan
3. *ke/-an* : kekondusifan, kemusliman, kevokalan, kebungkukan, keabaian, kearoganan, kegaiban
4. *per/-an* : perburungan, perjadwalan, perkopian, perlimbahan, persayuran, perguliran, pergundikan, perkencanan
5. *peN-* : penenung, pengaku, pengaudit, pengedit, penginterpretasi, pengorbit, penobat, penyedekah
6. *-wan* : surgawan, biologiwan, santriwan, puasawan, industriwan, firmawan, astrofisikawan
7. *-wati* : seniwati dan belawati.

Di dalam korpus kami, imbuhan *pe-* tidak memiliki *hapax legomena*. Frekuensi terendah ada pada kata ‘pelayar’, yaitu 2. Dalam mengambil contoh kata untuk frekuensi imbuhan ini, kami mengecek penggunaan kata dalam konteks kalimat. Pengecekan kami lakukan secara daring melalui situs cqweb (<https://cqweb.lancs.ac.uk/tbik3/>) [22].



Gambar 5 Perkembangan token dan *hapax legomena* imbuhan pembentuk nomina dalam bahasa Indonesia dari tahun 2011 hingga 2020.

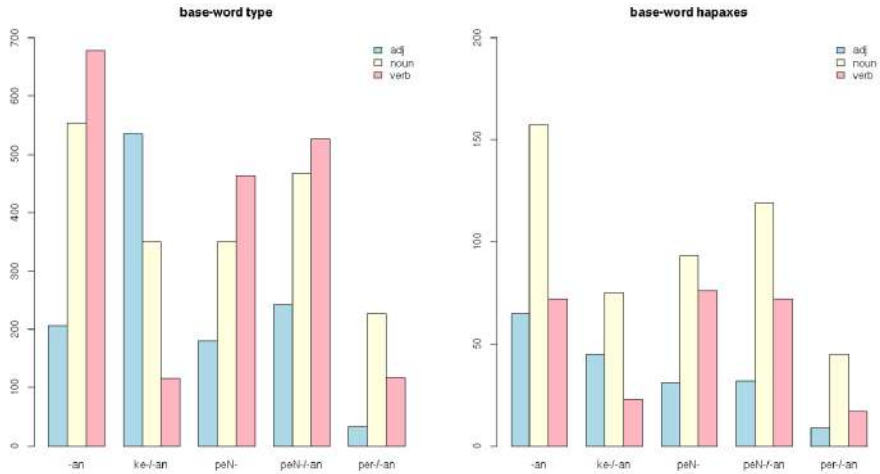
Distribusi imbuhan di dalam korpus yang merupakan produk bahasa tulis tentu saja tidak akan sama dari tahun ke tahun. Bahasa memiliki dinamika, demikian pula dengan dinamika imbuhan yang merupakan salah satu unsur kebahasaan. Gambar 5 menunjukkan perkembangan penggunaan imbuhan pembentuk nomina dalam bahasa Indonesia dari tahun 2011 hingga 2020. Imbuhan *-an*, *ke-/an*, dan *peN-/an* selalu menjadi imbuhan dengan frekuensi penggunaan tertinggi dari tahun ke tahun. Dengan tingkat produktivitas 0.0014, imbuhan *-an* juga memiliki *hapax legomena* tertinggi dibanding imbuhan yang lain. Menariknya, Gambar 5 juga menunjukkan bahwa pertumbuhan token dengan *hapax legomena* cenderung memiliki tren yang berkebalikan. Seperti yang pernah dijelaskan oleh [23], fenomena ini disebut *productivity paradox*, yaitu ketika sebuah formasi yang memiliki frekuensi penggunaannya tinggi justru tidak produktif dalam pembentukan kata baru. Contohnya, imbuhan *ke-/an* yang frekuensi tokennya tinggi justru cenderung tidak produktif dalam pembentukan kata baru (*hapax legomena*). Sedangkan imbuhan *-wati* yang

memiliki token paling sedikit justru cukup produktif dalam pembentukan kata baru (*hapax legomena*) di setiap tahun. Hal ini menjadi bukti pendukung tingkat produktivitas *-wati* yang tertinggi di antara imbuhan lain (0.0581).

4.2 Distribusi Kelas-Kata Kata-Dasar

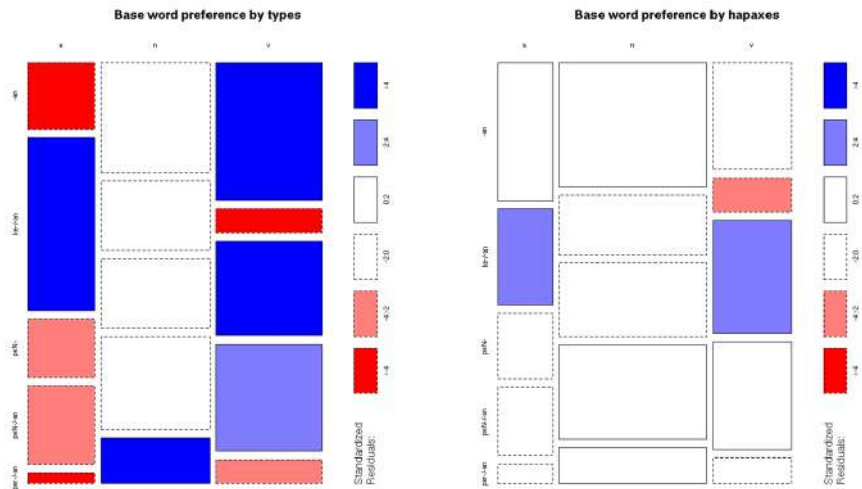
Bagian ini menjelaskan distribusi kelas-kata kata-dasar imbuhan pembentuk kata benda *-an*, *ke-/an*, *peN-*, *peN-/an*, dan *per-/an*. . *Konsep kata dasar dalam penelitian ini adalah kata yang melekat pada imbuhan. Misalnya, kata turunan keterlambatan memiliki kata dasar terlambat.* Kami memilih lima imbuhan tersebut untuk dianalisis secara kuantitatif sebab *-an*, *ke-/an*, *peN-*, *peN-/an*, dan *per-/an* merupakan imbuhan yang paling produktif di antara imbuhan yang lain, baik dari segi frekuensi token maupun tipe (lihat Gambar 4).

Gambar 6 menunjukkan bahwa terdapat variasi distribusi kelas-kata kata-dasar pada perhitungan tipe; verba mendominasi kata dasar untuk imbuhan *-an* (sulap – sulapan, kutuk – kutukan, pikul – pikulan), *peN-* (takluk – penakluk, himpun – penghimpun, ulas – pengulas), dan *peN-/an* (sembunyi – penyembunyian, boleh – pembolehkan, susup – penyusupan). Ajektiva mendominasi kelas-kata kata-dasar untuk imbuhan *ke-/an* (netral – kenetralan, kekal – kekekalan, profesional – keprofesionalan). Nomina mendominasi kelas-kata kata-dasar untuk imbuhan *per-/an* (hewan – perhewan, gula – pergulaan, senyawa – persenyawaan). Berbeda dengan distribusi tipe yang cukup variatif, ternyata kata benda mendominasi seluruh kelas-kata kata-dasar untuk imbuhan *-an* (diskon – diskonan, koreksi – koreksian, cangkok – cangkokan), *ke-/an* (pendeta – kependetaan, hewan – kehewan, popularitas – kepopularitasan), *peN-* (koreksi – pengoreksi, desain – pendesain, survey – penyurvei), *peN-/an* (sifat – penyifatan, bumi – pembumian, asumsi – pengasumsian), dan *per-/an* (burung – perburungan, kopi – perkopian, limbah – perlimbah) dari segi perhitungan *hapax legomena*.



Gambar 6 Distribusi kelas-kata kata-dasar untuk imbuhan pembentuk kata benda dalam bahasa Indonesia untuk sebaran tipe (panel kiri) dan hapax legomena (panel kanan).

Adanya variasi distribusi tipe kelas-kata kata-dasar menimbulkan kemungkinan adanya preferensi kelas-kata kata-dasar yang melekat pada imbuhan *-an*, *ke-/an*, *peN-*, *peN-/an*, dan *per-/an*.



Gambar 7 Plot mozaik yang menggambarkan *cross-classification* antara kelas-kata kata-dasar dengan imbuhan pembentuk nomina dalam bahasa Indonesia, dalam perhitungan tipe (panel kiri) dan perhitungan *hapax legomena* (panel kanan).

Gambar 7 menunjukkan *cross-classification* antara kelas-kata kata-dasar dengan imbuhan pembentuk kata benda. *Cross-classification* adalah klasifikasi satu variabel ke lebih dari satu kategori. Dalam hal ini, kami mengklasifikasi 5 imbuhan (*an*, *ke-/an*, *peN-*, */peN-/an*, dan *per-/an*) dengan 3 kelas kata (kata sifat, kata benda, dan kata kerja). *Cross-classification* ini dilakukan untuk memberikan hasil observasi apakah ada *over-representasi* atau *under-representasi* antara variabel yang diteliti. Warna biru dan ungu menunjukkan bahwa jumlah data yang diobservasi melebihi jumlah data yang diharapkan (*over-representasi*) muncul di dalam korpus. Warna merah dan merah muda menunjukkan *under-representasi*, bahwa jumlah data yang diobservasi kurang dari jumlah data yang diharapkan di dalam korpus ($X^2_{(8)} = 785.3$, $p < 0,0001$ untuk perhitungan tipe dan $X^2_{(8)} = 38.516$, $p < 0,0001$ untuk perhitungan *hapax legomena*). Plot mozaik ini menunjukkan bahwa kata sifat memiliki representasi yang berlebih dalam korpus sebagai kata dasar untuk imbuhan *ke-/an* baik dari perhitungan tipe maupun *hapax legomena*, kata benda memiliki representasi berlebih sebagai kata dasar imbuhan *per-/an* berdasarkan perhitungan tipe, dan kata kerja memiliki representasi berlebih sebagai kata dasar imbuhan *-an*, *peN-*, dan *peN-/an*.

4.3 Imbuhan *pe-* dan *peN-*

Imbuhan *pe-* dan salah satu alomorf *peN-*, *peN_{pe-}-*, memiliki kemiripan dalam hal bentuk dan makna. Contoh kemiripan ini ada ketika imbuhan *pe-* dan *peN-* melekat pada kata dasar berawalan /l/, /w/, dan /r/. Tabel 2 menunjukkan kata dasar berawalan /l/, /w/, dan /r/ yang melekat pada imbuhan *pe-* dan *peN-*. Orang yang tidak memiliki pengetahuan linguistik tidak akan dapat membedakan imbuhan *pe-* dan *peN_{pe-}-* yang terdapat pada kata *pelari* dan *pelukis*.

Tabel 4 Contoh kata berimbuhan *pe-* dan *peN-* yang memiliki kemiripan bentuk dan makna

Kata dasar	Kata turunan	Awalan	Makna
lukis	pelukis	<i>peN-</i>	agen
lari	pelari	<i>pe-</i>	agen
wawancara	pewawancara	<i>peN-</i>	agen
wisata	pewisata	<i>pe-</i>	agen
rintis	perintis	<i>peN-</i>	agen
runding	perunding	<i>pe-</i>	agen

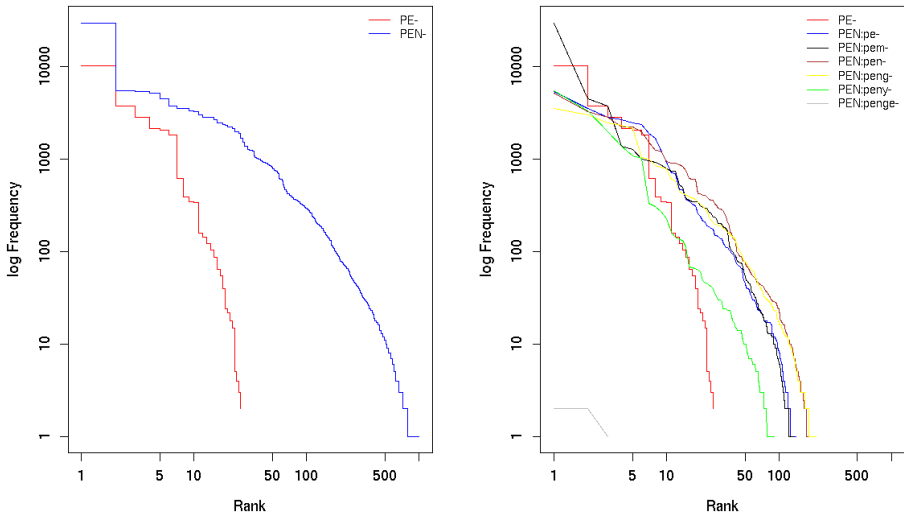
Kemiripan bentuk *pe-* dan *peN-* membuat linguist memiliki dua pendapat yang berbeda mengenai status imbuhan *pe-* dan *peN-* [24]. Sneddon et al. [3] dan Ramlan [13] berpendapat bahwa *pe-* dan *peN-* adalah dua imbuhan yang berbeda, sedangkan Dardjowidjojo [25] dan Kridalaksana [26] berpendapat bahwa *pe-* merupakan alomorf dari *peN-*. Moeliono [2] dan Benjamin [27] percaya bahwa *pe-* bisa merupakan alomorf dari *peN-*, dan bisa juga merupakan versi modern dari *per-*. Hingga sekarang, belum ada kesepakatan secara kualitatif mengenai status morfologi kedua imbuhan ini.

Oleh sebab itu, penelitian terbaru mengenai *pe-* dan *peN-* dilakukan dengan menggunakan analisis kuantitatif dan perhitungan statistik untuk menjawab apakah *pe-* dan *peN-* merupakan dua imbuhan yang berbeda, atau mereka hanyalah alomorf [8], [9], [28]. Data penelitian tersebut berasal dari korpus yang bernama *Leipzig Corpora Collection* [29]. Studi kuantitatif dari perhitungan produktivitas imbuhan menyimpulkan bahwa *pe-* dan *peN-* merupakan dua imbuhan yang berbeda karena (1) kedua imbuhan ini memiliki produktivitas yang berbeda, yaitu *peN-* lebih produktif dibandingkan dengan *pe-*, (2) kedua imbuhan ini memiliki kecenderungan makna yang berbeda, yaitu bahwa *peN-* produktif dalam pembentukan makna agen dan instrumen, sedangkan *pe-* produktif dalam pembentukan makna agen dan pasien. Menariknya, studi tersebut juga menyebutkan bahwa selain sebagai pembentuk agen dan instrumen, imbuhan *peN-* juga memiliki makna 'lokasi X' (*penghujung*), 'penyebab X' (*penyakit*), dan 'sesuatu yang di-X' (*pencelup*).

Dalam penelitian ini, kami mencoba menganalisis produktivitas imbuhan *pe-* dan *peN-* dengan menggunakan metode yang sama untuk diaplikasikan ke dalam korpus yang saat ini kami gunakan. Dari total 15.352 kata benda yang ada di dalam korpus, 1.021 di antaranya memiliki imbuhan *pe-* dan *peN-*. Dari 1.021 data *pe-* dan *peN-*, kami mengambil kata turunan yang tidak memiliki infleksi (misalnya klitik *-ku*, *-mu*, *-nya*). Penyaringan data menghasilkan 26 kata berimbuhan *pe-* dan 1.369 kata berimbuhan *peN-*.

Secara bentuk, produktivitas *peN-* memang lebih tinggi dibandingkan dengan produktivitas *pe-*. Gambar 1 menyajikan plot frekuensi penggunaan kata berimbuhan *pe-* dan *peN-* (beserta alomorfnya -lihat panel kanan-) yang terdapat di dalam korpus. Kami menggunakan skala logaritmik untuk plot frekuensi-peringkat ini [19], [20]. Data menunjukkan bahwa kata-kata berimbuhan *peN-* memiliki peringkat penggunaan tertinggi, yang secara umum jumlahnya selalu berada di atas frekuensi penggunaan kata-kata berimbuhan *pe-*. Selain itu, kurva frekuensi-peringkat *pe-* juga berada di posisi terendah

jika dibandingkan dengan alomorf *peN-* yang lain (lihat panel kanan). Hal ini menunjukkan bahwa dalam analisis menggunakan korpus yang berbeda, *peN-* tetap lebih produktif daripada *pe-*.



Gambar 8 Kurva produktivitas imbuhan *pe-* dan *peN-* (panel kiri), serta imbuhan *pe-* dan alomorf imbuhan *peN-* (panel kanan). Frekuensi penggunaan *peN-* lebih produktif dibandingkan dengan frekuensi penggunaan *pe-*, jika dibandingkan dari segi imbuhan, maupun dari segi alomorf (pengecualian untuk alomorf *penge-* yang memang paling tidak produktif jika dibandingkan dengan alomorf *peN-* lain).

Analisis kuantitatif menggunakan distribusi semantik juga mengkonfirmasi bahwa *pe-* dan *peN-* merupakan dua imbuhan yang berbeda sebab mereka memiliki distribusi semantik tersendiri. Distribusi semantik merupakan metode pemetaan berdasarkan konteks kalimat [30], [31]. Berdasarkan teori [32, p. 11], bahwa “*You shall know a word by the company it keeps*” (konteks menentukan makna kata), distribusi semantik imbuhan *pe-* dan *peN-* menunjukkan bahwa *pe-* memiliki cakupan konteks kalimat yang lebih spesifik jika dibandingkan dengan imbuhan *peN-*. Imbuhan *pe-* lebih spesifik muncul dalam konteks olahraga, seperti misalnya pada kata *petenis*, *pegolf*, *perenang*, *petinju*, *pelayar*, *pelari*, dan *pegulat*. Dengan demikian, teori ini mendukung pernyataan Chaer [33] bahwa pembeda *pe-* dan *peN-* adalah adanya spesifikasi makna dalam bidang keolahragaan. Hal ini juga menjadi landasan teoretis untuk membedakan *penembak* (*peN-*) dan *petembak* (*pe-*),

peninju (*peN-*) dan *petinju* (*pe-*), serta *pendayung* (*peN-*) dan *pedayung* (*pe-*). *Penembak*, *peninju*, dan *pendayung* adalah orang yang menembak, meninju, dan mendayung. Sedangkan *petembak*, *petinju*, dan *pedayung* adalah atlet tembak, tinju, dan dayung.

Perbedaan *pe-* dan *peN-* juga terdapat pada apakah imbuhan ini memiliki korelasi dengan verba berimbuhan *meN-* [8]. Korelasi yang dimaksud adalah bahwa verba berimbuhan *meN-* berpotensi memiliki korespondensi dengan agen atau instrumen berimbuhan *peN-* (misalnya *membaca-pembaca*, *menulis-penulis*, *menyumbang-penyumbang*, *melukis-pelukis*, *mengaduk-pengaduk*, dan *mengecek- pengecek*). Secara kuantitatif, seluruh alomorf *peN-* berkorelasi erat dengan seluruh alomorf *meN-*, sedangkan imbuhan *pe-* tidak berkorelasi dengan imbuhan *meN-*; dengan nilai uji korelasi antara imbuhan *peN-* dan *meN-* sebesar 44%. Angka tersebut berarti bahwa setiap verba yang memiliki imbuhan *meN-*, ada 44% kemungkinan bahwa verba ini memiliki agen atau instrumen yang berimbuhan *peN-*.

5. KESIMPULAN DAN SARAN

Pembentukan kata benda dalam bahasa Indonesia melalui proses afiksasi terbagi dalam awalan, akhiran, serta awalan dan akhiran. Makalah ini membahas lebih detail mengenai bentuk dan makna yang tercipta dari proses afiksasi dengan imbuhan *pe-*, *peN-*, *-wan*, *-wati*, *per-/an*, *peN-/an*, dan *ke-/an*. Contoh kalimat dan analisis data yang disajikan dalam makalah ini berasal dari data korpus bahasa Indonesia tulis yang diambil dari berbagai ragam teks dalam rentang tahun 2011-2020.

Pendekatan kualitatif yang dilakukan dalam penelitian ini memungkinkan kami untuk mengidentifikasi beberapa makna baru yang muncul dalam proses afiksasi pembentukan nomina dalam bahasa Indonesia. Misalnya, makna ‘proses memberi X’ pada kata “pencahaya-an” yang ada pada imbuhan *peN-/an* dapat digolongkan sebagai makna baru. Makalah ini juga menjawab persoalan imbuhan *pe-* dan *peN-* yang secara kualitatif masih belum diputuskan apakah kedua imbuhan ini alomorf atau bukan.

Selain itu, pendekatan kuantitatif yang merupakan kebaruan dari penelitian ini memungkinkan kami untuk menghitung produktifitas afiks dan preferensi kelas-kata kata-dasar dari setiap imbuhan. Data menunjukkan bahwa imbuhan pembentuk nomina yang produktif adalah imbuhan *an*, *ke-/*

an, dan *peN-/an*. Sedangkan imbuhan *pe-*, *-wan*, dan *-wati* merupakan tiga imbuhan yang jarang digunakan di dalam korpus. Akan tetapi, imbuhan yang jarang digunakan di dalam korpus ternyata memiliki potensi penambahan kata baru yang lebih besar dibanding imbuhan yang lebih produktif. Analisis kuantitatif berbasis korpus juga menunjukkan bahwa kata sifat memiliki representasi yang berlebih sebagai kata dasar untuk imbuhan *ke-/an*, kata benda memiliki representasi berlebih sebagai kata dasar imbuhan *per-/an*, dan kata kerja memiliki representasi berlebih sebagai kata dasar imbuhan *-an*, *peN-*, dan *peN-/an*.

Kami menyadari bahwa perlu ada verifikasi manual yang dilakukan oleh peneliti untuk memastikan bahwa data yang merupakan luaran MorphInd bersifat valid. Selain itu, makalah ini belum menyinggung mengenai imbuhan yang homonim, misalnya imbuhan *ke-/an* sebagai pembentuk kata benda dan kata sifat (contoh, *ke-/an* ‘kebesaran’ pada kalimat “Aku tunduk pada kebesaran-Nya.” vs. “Baju itu kebesaran.”). Perbedaan yang signifikan antara imbuhan *peN-/an* dan *per-/an* seperti pada kata ‘penggantian’ dan ‘pergantian’, atau ‘pencampuran’ dan ‘percampuran’ juga akan menjadi topik lanjutan yang menarik untuk dibahas. Selain itu, penelitian mengenai pemodelan untuk menjelaskan keterkaitan antara produktivitas, ragam teks, dan tahun juga bisa dilakukan.

Akhirnya, kami berharap penelitian berbasis korpus ini dapat memberikan pandangan baru mengenai proses pembentukan kata benda dalam bahasa Indonesia.

Referensi

- [1] H. Kridalaksana, *Kelas Kata dalam Bahasa Indonesia*, Second. Jakarta: Gramedia Pustaka Utama, 2007.
- [2] A. M. Moeliono, H. Lipoliwa, H. Alwi, S. S. W. Sasangka, and Sugiyono, *Tata Bahasa Baku Bahasa Indonesia*, 4th ed. Jakarta: Badan Pengembangan dan Pembinaan Bahasa, 2017.
- [3] J. N. Sneddon, A. Adelaar, D. N. Djenar, and M. C. Ewing, *Indonesian: A Comprehensive Grammar*, Second. New York: Routledge, 2010.
- [4] S. Dardjowidjojo, “Nominal Derivation in Indonesian,” in *Second International Conference on Austronesian Linguistics: Proceedings*, 1978, pp. 503–528.

- [5] G. Widiyanto, “Kajian morfosemantis kehomoniman sirkumfiks ke-an dalam bahasa Indonesia,” presented at the Seminar akademik PPPPTK Bahasa, Jan. 30, 2006.
- [6] N. H. Jalaluddin and A. H. Syah, “Penelitian Makna Imbuhan Pen-Dalam Bahasa Melayu: Satu Kajian Rangka Rujuk Silang,” *GEMA Online J. Lang. Stud.*, vol. 9, no. 2, pp. 57–72, 2009.
- [7] P. Grangé, “Aspect in Indonesian: free markers versus bound markers,” *NUSA Linguist. Stud. Lang. Indones.*, vol. 55, pp. 57–79, 2013.
- [8] K. Denistia and H. Baayen, “The Indonesian prefixes PE- and PEN-: A study in productivity and allomorphy,” *Morphology*, vol. 29, no. 3, pp. 384–407, 2019, doi: <https://doi.org/10.1007/s11525-019-09340-7>.
- [9] K. Denistia and H. Baayen, “Affix substitution in Indonesian: A computational modeling approach,” *Linguist. Interdiscip. J. Lang. Sci.*, 2020.
- [10] K. Denistia, E. Shafaei-Bajestan, and R. H. Baayen, “Exploring semantic differences between the Indonesian prefixes PE- and PEN- using a vector space model,” *Corpus Linguist. Linguist. Theory*, pp. 1–26, 2021.
- [11] K. Denistia and H. Baayen, “The morphology of Indonesian: Data and quantitative modeling,” in *The Routledge Handbook of Asian Linguistics*, C. Shei and S. Li, Eds. London: Routledge, 2022, pp. 605–634.
- [12] G. P. W. Rajeg, K. Denistia, and S. Musgrave, “Vector Space Models and the usage patterns of Indonesian denominal verbs: A case study of verbs with meN-, meN-/kan, and meN-/i affixes,” *NUSA Linguist. Stud. Lang. Indones.*, vol. 67, no. 1, pp. 35–76, 2019.
- [13] M. Ramlan, *Morfologi: Suatu Tinjauan Deskriptif*. Yogyakarta: CV Karyono, 2009.
- [14] R. H. Baayen, “Word frequency distributions,” in *Quantitative Linguistics. An international handbook*, R. Köhler and G. Altmann, Eds. Berlin: Walter de Gruyter, 2005, pp. 397–409.
- [15] S. D. Larasati, V. Kuboň, and D. Zeman, “Indonesian morphology tool MorphInd: Towards an Indonesian corpus,” in *Systems and Frameworks for Computational Morphology*, 2011, vol. 100, pp. 119–129.
- [16] S. R Team, *RStudio: Integrated Development for R*. RStudio. Boston, MA: RStudio, Inc., 2015. [Online]. Available: <http://www.rstudio.com/>
- [17] Sugerman, *Morfologi Bahasa Indonesia: Kajian ke Arah Linguistik Deskriptif*. Yogyakarta: Penerbit Ombak, 2016.
- [18] Sukarno, “The Behaviours of the General Nasal /N/ in Indonesian

- Active Prefixed Verbs,” *Int. J. Lang. Linguist.*, vol. 4, no. 2, pp. 48–52, Jun. 2017.
- [19] G. K. Zipf, *The Psycho-Biology of Language*. Boston: Houghton Mifflin, 1935.
- [20] G. K. Zipf, *Human Behaviour and the Principle of the Least Effort. An Introduction to Human Ecology*. New York: Hafner, 1949.
- [21] R. H. Baayen, “Corpus Linguistics. An International Handbook,” A. Lüdeling and M. Kyto, Eds. Berlin: Mouton De Gruyter, 2009, pp. 900–919.
- [22] A. Hardie, “CQPweb—combining power, flexibility and usability in a corpus analysis tool,” *Int. J. Corpus Linguist.*, vol. 17, no. 3, pp. 380–409, 2012.
- [23] A. Krott, S. Robert, and R. H. Baayen, “Complex words in complex words,” *Linguistics*, vol. 37, pp. 905–926, 1999.
- [24] K. Denistia, “Revisiting the Indonesian Prefixes PEN-, PE2-, and PER-,” *Linguist. Indones.*, vol. 36, no. 2, pp. 145–159, 2018.
- [25] S. Dardjowidjojo, *Some Aspects of Indonesian Linguistics*. Jakarta: Djambatan, 1983.
- [26] H. Kridalaksana, *Kamus Linguistik*, 4th ed. Jakarta: PT Gramedia Pustaka Utama, 2008.
- [27] G. Benjamin, “Affixes, Austronesian and iconicity in Malay,” *Bijdr. Tot Taal- Land- En Volkenkd.*, vol. 165, no. 2–3, pp. 291–323, 2009.
- [28] K. Denistia, E. Shafaei-Bajestan, and H. Baayen, “Exploring semantic differences between the Indonesian prefixes PE- and PEN- using a vector space model,” *Corpus Linguist. Linguist. Theory*, vol. 18, no. 3, pp. 573–598, 2022, doi: 10.1515/cllt-2020-0023.
- [29] U. Quasthoff, M. Richter, and C. Biemann, “Corpus Portal for Search in Monolingual Corpora,” Genoa, 2006, pp. 1799–1802.
- [30] T. Landauer and S. Dumais, “A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge,” *Psychol. Rev.*, vol. 104, no. 2, pp. 211–240, 1997.
- [31] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” *ArXiv Prepr. ArXiv13013781*, 2013.
- [32] J. R. Firth, “Studies in Linguistic Analysis,” Oxford: Basil Blackwell, 1957, pp. 1–32.
- [33] A. Chaer, *Morfologi Bahasa Indonesia (Pendekatan Proses)*. Jakarta: PT Rineka Cipta, 2008.

AFIKS PEMBENTUK ADJEKTIVA DALAM BAHASA INDONESIA

Dewi Puspita dan Dora Amalia

Badan Pengembangan dan Pembinaan Bahasa

Abstrak

Adjektiva dalam bahasa Indonesia dapat dibentuk melalui afiksasi. Proses pembentukannya dapat dilakukan secara infleksif maupun derivatif. Kajian terkait afiks pembentuk adjektiva dalam bahasa Indonesia telah menghasilkan beragam pendapat. Makalah ini disusun untuk memeriksa kembali pendapat-pendapat itu melalui pendekatan korpus linguistik dan menemukan hal lain yang belum tercakup sebelumnya berdasarkan data korpus yang berasal dari tahun 2011 hingga 2020. Dengan menggunakan pendekatan kuantitatif dan kualitatif, kajian ini telah berhasil mengonfirmasi jenis, bentuk, dan makna afiks pembentuk adjektiva dalam bahasa Indonesia serta memetakan penggunaan afiks tersebut dalam korpus. Hasil kajian yang menyajikan gambaran penggunaan bahasa, khususnya adjektiva berafiks, masa kini diharapkan dapat menjadi bahan kajian yang mutakhir untuk penyusunan tata bahasa kontemporer.

Kata kunci: afiksasi, adjektiva, afiks, infleksional, derivasional

1. PENDAHULUAN

1.1 Latar Belakang

Dalam bahasa Indonesia, kata dapat dibentuk melalui penambahan afiks atau afiksasi. Afiksasi mengacu pada runtutan perubahan yang dialami oleh suatu bentuk dasar atau leksem sehingga leksem itu menjadi kata [1]. Runtutan perubahan yang dimaksud adalah perubahan bentuk, kelas kata, dan makna. Perubahan bentuk yang terjadi adalah dari bentuk dasar menjadi bentuk berimbuhan.

Contoh:

bentuk dasar

baik

cantik

bentuk berafiks

membaik, sebaik, terbaik

kecantikan, secantik, tercantik

Kelas kata dapat terbentuk secara infleksif maupun derivatif. Infleksi merupakan proses morfologis yang melibatkan tataran sintaksis, bersifat sistematis, dapat diprediksi, teratur, otomatis, bersifat konsisten, dan tidak mengubah identitas leksikal. Adapun derivasi secara sintaksis tidak dapat diramalkan, tidak otomatis, tidak sistemik, bersifat opsional atau sporadis, serta secara morfologis dapat mengubah identitas leksikal [2]. Secara singkat dapat dikatakan bahwa proses pembentukan kata baru secara infleksif tidak mengubah kelas kata awalnya, sedangkan proses pembentukan kata secara derivatif membentuk kelas kata yang berbeda dari kelas kata awalnya. Afiksasi dapat membentuk beberapa kelas kata, di antaranya adalah nomina, verba, adjektiva, dan adverbial.

Contoh perubahan secara infleksif:

baik (adj)	terbaik (adj)
cantik (adj)	secantik (adj)

Contoh perubahan secara derivatif:

baik (adj)	membaik (v)
cantik (adj)	kecantikan (n)

Proses perubahan selanjutnya, yang terbentuk sebagai hasil dari afiksasi, adalah perubahan makna. Tiap-tiap afiks mengandung makna tersendiri. Contoh perubahan makna yang terbentuk dari hasil afiksasi dapat dilihat dalam diagram berikut.

<i>ubah</i>	→	<i>mengubah</i>	→	<i>pengubah</i>	→	<i>pengubahan</i>	→	
<i>ubahan</i>								

dasar	→	kegiatan	→	pelaku	→	proses	→	
hasil								

Diagram 1 Perubahan makna dari hasil afiksasi

Afiksasi dapat membentuk verba, nomina, dan adjektiva. Fokus dari penelitian ini adalah afiks pembentuk adjektiva, baik secara infleksif maupun derivatif, dari segi jenis, bentuk, dan maknanya. Hingga saat ini masih terdapat beberapa perbedaan pemahaman di kalangan ahli bahasa terkait

afiksasi, baik dari jenis afiksnya, kelas kata yang dibentuknya, maupun makna yang dikandungnya. Makalah ini akan merangkum pendapat-pendapat dari para ahli tersebut dengan mengonfirmasikannya pada data korpus. Dengan menggunakan data korpus, akan terlihat bagaimana masyarakat membentuk kata dengan afiksasi untuk mengungkapkan makna dalam komunikasinya secara tertulis.

1.2 Tujuan

Makalah ini disusun dengan tujuan memeriksa kembali pendapat-pendapat yang dikemukakan oleh peneliti sebelumnya terkait afiks pembentuk adjektiva melalui pendekatan korpus linguistik, memeriksa apakah data korpus mengonfirmasi pendapat tersebut atau ada hal lain yang ditemukan seiring dengan perkembangan zaman. Data korpus yang berasal dari tahun 2011 hingga 2020 diharapkan dapat menyajikan penggunaan bahasa masa kini dan menghasilkan bahan kajian yang mutakhir untuk penyusunan tata bahasa kontemporer.

1.3 Tinjauan Pustaka

Artikel penting tentang pembentukan kata melalui afiksasi dalam bahasa Indonesia ada dalam buku Tata Bahasa Baku Bahasa Indonesia [3]. Dalam buku itu dijelaskan secara rinci perihal adjektiva dalam satu bab khusus, termasuk tentang adjektiva berafiks. TBBI menyebutkan bahwa adjektiva berafiks dapat diperinci lebih lanjut menjadi 1) adjektiva berprefiks, 2) adjektiva berinfiks, 3) adjektiva bersufiks, dan 4) adjektiva berkonfiks. Rincian afiks pembentuk adjektiva menurut TBBI dapat dilihat dalam Tabel 1.

Tabel 1 Afiks Pembentuk Adjektiva

prefiks	<i>se-, ter-</i>	<i>sebesar, setinggi, semerah, senyaman termahal, terpanjang, termegah, termiskin</i>
infiks	<i>-em-</i>	<i>gemetar, gemuruh, kemilap, kemilau, temaram</i>
sufiks	<i>-i, -wi, -iah, -wiah -if, -er, -al, -is, -us</i>	<i>alami, duniawi, alamiah administratif, parlementer, prosedural, hierarkis, religius</i>
konfiks	<i>ke-...-an</i>	<i>kesempitan, kehausan, kesakitan, kegirangan</i>

Beberapa catatan penting terkait afiks pembentuk adjektiva dalam Tata Bahasa Baku Bahasa Indonesia adalah:

1. dalam pembentukan kata dengan infiks *-em-*, bentuk dasar nomina dapat berdiri sendiri tanpa penyisipan *-em-* terlebih dahulu, sementara itu, bentuk dasar verba dan adjektiva tidak dapat berfungsi sebagai unsur sintaksis sebelum disisipi *-em-*,
2. adjektiva bersufiks *-i*, *-wi*, atau *-iah*, *-wiah* memiliki bentuk dasar nomina yang pada umumnya berasal dari bahasa Arab. Sufiks-sufiks tersebut sering juga diterapkan pada nomina serapan yang berasal dari bahasa lain, seperti

contoh:

Nomina	Adjektiva	Adjektiva
<i>alam</i>	<i>alami</i>	<i>alamiah</i>
<i>islam</i>	<i>islami</i>	<i>islamiah</i>
insan	insani	insaniah

1. secara umum, sufiks *-i*, *-iah* muncul di belakang kata yang berakhir dengan konsonan, sedangkan sufiks *-wi*, *-wiah* di belakang kata yang berakhir dengan vokal /a/,
2. sufiks *-if*, *-er*, *-al*, *-is*, dan *-us* diserap dari bahasa Belanda atau bahasa Inggris dan kata dasarnya biasanya berupa nomina.

Kajian lain terkait afiksasi dalam bahasa Indonesia memperlihatkan bahwa ada afiks lain, selain yang disebutkan oleh TBBI, yang juga dapat membentuk adjektiva, seperti prefiks *meng-* [4] dan sufiks *-an* [5]. Dalam kajian itu disebutkan bahwa prefiks *meng-* dapat membentuk adjektiva denominal seperti *merakyat*. Prefiks *meng-* yang melekat pada kata *merakyat* memiliki makna ‘bersifat atau berlaku seperti...’. Dengan demikian, makna dari *merakyat* adalah ‘bersifat atau berlaku seperti rakyat’ [4]. Sementara itu, Putra dan Utami [5] dalam makalahnya menyebutkan bahwa sufiks *-an* juga dapat membentuk adjektiva walaupun kata yang dibentuknya merupakan bentuk cakapan atau bentuk yang tidak baku. Bentuk adjektiva takbaku lainnya dapat dibentuk juga oleh konfiks *ke-* *-an* seperti yang mereka tunjukkan dalam tabel berikut.

Tabel 2 Afiks Pembentuk Adjektiva Takbaku

Bentuk	Bentuk Nonformal	Bentuk Formal	Akar Kata
sufiks <i>-an</i>	duluan	lebih dulu	dulu
	gendutan	lebih gendut	gendut
	karatan	berkarat	karat
	(gak) sabaran	tidak sabar	sabar
konfiks <i>ke- -an</i>	kebagusan	terlalu bagus	bagus

Tinjauan selanjutnya terkait proses derivasi dan infleksi dari afiksasi adjektiva [2], [6], [7], [8], [9], [10]. Adjektiva berafiks dapat dibentuk secara infleksif maupun derivatif. Adriani dkk [10] menyatakan bahwa dalam bahasa Indonesia, afiks infleksional hanya terbatas pada sufiks saja, sedangkan afiks derivasional dapat berupa prefiks, sufiks, maupun konfiksa. Sementara itu, penelitian tentang prefiks *ter-* dalam bahasa Indonesia [7] menyebutkan bahwa prefiks *ter-* dapat digunakan untuk membentuk verba, adjektiva, dan nomina. Bentuk yang paling banyak ditemukan adalah verba. Untuk membentuk verba, adjektiva, dan nomina, prefiks *ter-* dapat digabungkan pada kata dasar berkelas kata sama ataupun berbeda. Dengan kata lain, prefiks *ter-* dapat membentuk kata baru secara infleksional maupun derivasional. Secara lengkap, Ermanto [8], [9] menyajikan tabel afiks infleksional dan derivasional pembentuk adjektiva sebagai berikut.

Tabel 3 Afiks Infleksional Pembentuk Adjektiva

No.	Afiks Infleksi	Proses dengan D	Fungsi Menurunkan	Contoh
1	<i>ter-</i>	<i>ter-</i> + Adj	Adj superlatif (paling)	tertinggi, terbaik
2	<i>ke--an</i>	<i>ke--an</i> + Adj	Adj ekksesif (berlebih)	kebesaran, kejauhan
3	<i>ke--an</i>	<i>ke--an</i> + Adj	Adj atenuatif (agak)	kemerahan, kekuningan
4	<i>se-</i>	<i>se-</i> + Adj	Adj ekuatif (sama)	secantik, semahal

Tabel 4 Afiks Derivasional Pembentuk Adjektiva

No.	Afiks Derivasi	Proses dengan D	Fungsi Menurunkan	Contoh
1	<i>-em-</i>	<i>-em-</i> + N	Adjektiva	gemetar, gemuruh
2	<i>-em-</i>	<i>-em-</i> + akar terikat	Adjektiva	gemerlap, gemilang

3	-i	-i + N	Adjektiva	alami, abadi, insani, hewani
4	-iah	-iah + N	Adjektiva	alamiah, insaniah, amaliah
5	-if	-if + N	Adjektiva	agresif, kompetitif
6	-er	-er + N	Adjektiva	komplementer, parlementer
7	-al	-al + N	Adjektiva	moral, struktural, ideal
8	-is	-is + N	Adjektiva	praktis, teknis
9	ke--an	ke--an + N	Adjektiva	keibuan, kebapakan

Pendapat-pendapat di atas akan dikonfirmasi kebenarannya melalui kajian berbasis korpus ini.

2. METODE

Penelitian ini adalah penelitian yang menggunakan pendekatan kuantitatif dan kualitatif berbasis korpus. Secara kuantitatif, data dianalisis berdasarkan frekuensi kemunculan, ragam teks yang memuat data itu, serta tahun kemunculannya. Hal ini dilakukan untuk melihat produktivitas afiksasi adjektiva dalam bahasa Indonesia. Secara kualitatif, data dianalisis berdasarkan jenis, bentuk, dan maknanya. Metode pencarian dan penganalisisan data dijelaskan sebagai berikut.

2.1 Metode Pencarian Data

Korpus yang digunakan untuk menjaring adjektiva berafiks adalah korpus TBIK di CQPWeb. Korpus ini memuat sekitar 25 juta kata. Pencarian dilakukan melalui gabungan simbol *regular expression* (regex) dan POS tag. Simbol yang digunakan adalah

1. me+_JJ untuk menjaring adjektiva yang diawali *ter-*
2. se+_JJ untuk menjaring adjektiva yang diawali *se-*
3. ter+_JJ untuk menjaring adjektiva yang diawali *ter-*
4. +em+_JJ untuk menjaring adjektiva yang mengandung *-em-*
5. ke+an_JJ untuk menjaring adjektiva yang diawali dengan *ke-* dan diakhiri dengan *-an*
6. +an_JJ untuk menjaring adjektiva yang diakhiri dengan *-an*
7. +i_JJ untuk menjaring adjektiva yang diakhiri dengan *-i*
8. +wi_JJ untuk menjaring adjektiva yang diakhiri dengan *-wi*

9. +iah_JJ untuk menjarang adjektiva yang diakhiri dengan *-iah*
10. +is_JJ untuk menjarang adjektiva yang diakhiri dengan *-is*

Hasil yang diperoleh dari pencarian dengan simbol di atas tidak hanya menjarang adjektiva berafiks, tetapi juga menjarang bentuk lain seperti kata dasar. Selain itu, karena POStagging yang belum sempurna, banyak kelas kata yang tertukar atau salah. Hal ini menyebabkan banyak data yang tidak diperlukan turut terjaring sementara data yang diperlukan malah tidak terjaring. Untuk itu, perlu dilakukan penyeleksian data. Data yang bukan bentuk adjektiva berafiks dikeluarkan dari daftar.

2.2 Metode Analisis Data

Analisis dilakukan pada seluruh data yang terseleksi dari hasil pencarian pada korpus TBIK. Data dianalisis berdasarkan frekuensi, perhitungan statistik pada kolokasi, dan baris konkordansi. Metode penghitungan statistik yang digunakan adalah Mutual Informaton (MI). Skor Mutual Informaton (MI) biasanya digunakan sebagai kriteria dalam menetapkan bahwa kemunculan bersama dari kata-kata tertentu yang diamati adalah kolokasi yang dapat dipercaya atau penting dalam struktur leksikal bahasa [11]. Mutual Informaton (MI) mengukur kekuatan kolokasi, seberapa kuat atau seberapa eksklusif dua item terikat satu sama lain. Analisis kolokasi dan konkordansi dilakukan untuk melihat makna yang dikandung atau dibentuk oleh afiks yang dikaji.

3. HASIL

Ditemukan beberapa jenis afiks pembentuk adjektiva dalam korpus, yang secara etimologis dapat dibagi ke dalam dua kelompok, yaitu afiks dari bahasa Indonesia dan afiks dari bahasa asing. Afiks serapan dari bahasa asing berasal dari bahasa Arab, Belanda, dan Inggris. Afiks pembentuk adjektiva dari bahasa Indonesia terdiri atas prefiks *meng-*, *se-*, dan *ter-*; infiks *-em-*; konfiks *ke-* *-an*, dan sufiks *-an*. Afiks serapan pembentuk adjektiva dari bahasa Arab adalah *-i*, *-wi*, dan *-iah*. Afiks serapan pembentuk adjektiva dari bahasa Belanda atau Inggris adalah *-is*.

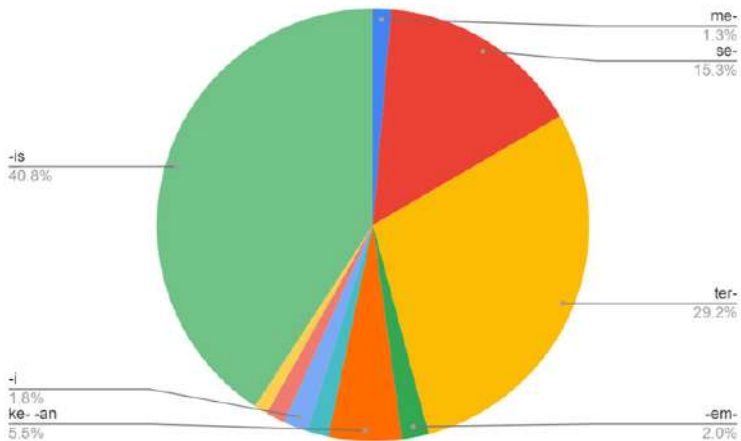
Selain kesepuluh afiks di atas, masih ada beberapa sufiks lain, yaitu *-wiah* (Arab), *-al*, *-ik*, *-if*, *-er*, dan *-us* (Belanda dan/atau Inggris). Namun, pencarian

kata berakhiran *-wiah* yang berkelas kata adjektiva tidak ditemukan dalam data korpus TBIK. Selain itu, data korpus juga menunjukkan bahwa kata dengan akhiran *-al*, *-ik*, *-if*, *-er*, dan *-us* seluruhnya merupakan kata serapan. Hal ini sesuai dengan kaidah penyerapan afiks dalam Pedoman Umum Pembentukan Istilah [12] sebagai berikut.

<i>-al</i> (Inggris)	menjadi	<i>-al</i>
<i>-eel, aal</i> (Belanda), <i>-al</i> (Inggris)	menjadi	<i>-al</i>
<i>-iek</i> (Belanda), <i>-ic, -ique</i> (Inggris)	menjadi	<i>-ik</i>
<i>-ief</i> (Belanda), <i>-ive</i> (Inggris)	menjadi	<i>-if</i>
<i>-air</i> (Belanda), <i>-ary</i> (Inggris)	menjadi	<i>-er</i>
<i>-eus</i> (Belanda)	menjadi	<i>-us</i>
<i>-ious</i> (Inggris)	menjadi	<i>-us</i>

Dengan demikian, *-wiah*, *-al*, *-ik*, *-if*, *-er*, dan *-us* dikecualikan dari pembahasan dalam artikel ini.

Sebelum masuk pada penjelasan masing-masing afiks, terlebih dulu disajikan frekuensi kemunculan adjektiva berafiks dalam korpus seperti yang disajikan dalam grafik berikut.



Grafik 1 Frekuensi Kemunculan Adjektiva Bersufiks dalam Korpus

Grafik menunjukkan bahwa adjektiva bersufiks *-is* paling banyak ditemukan dalam korpus (40,8%) disusul adjektiva berprefiks *ter-* (29,2%), dan adjektiva berprefiks *se-* (15,3%). Selebihnya, frekuensi kemunculan adjektiva berprefiks *meng-*, berkonfiks *ke-* *-an*, berinfix *-em-*, dan bersufiks *-an*, *-i*, dan *-wi* berjumlah kurang dari 5%. Banyaknya jumlah adjektiva bersufiks *-is* dalam data disebabkan oleh banyaknya bentuk serapan asing yang turut terjaring dalam proses pencarian data. Dari jumlah itu hanya beberapa saja yang merupakan kata turunan berafiks.

Dalam subbab-subbab selanjutnya dijelaskan jenis, bentuk, dan makna yang dihasilkan oleh afiks pembentuk adjektiva dalam bahasa Indonesia berdasarkan data yang diperoleh dari korpus Tata Bahasa Indonesia Kontemporer.

3.1 Prefiks *meng-*

Beberapa literatur tidak mendaftarkan prefiks *meng-* sebagai prefiks pembentuk adjektiva dalam bahasa Indonesia. KBBI edisi V [13] hanya mencatat prefiks *meng-* sebagai prefiks pembentuk verba. Namun, data korpus memperlihatkan adanya beberapa adjektiva berprefiks *meng-* seperti *menengah*, *melingkar*, *meriah*, *mencuat*, *menggigil*, dan *merinding*. Frekuensi kemunculan kata-kata itu di atas seratus kali. Kata *menengah* bahkan muncul hingga 3.017 kali. Daftar adjektiva berprefiks *meng-* yang ditemukan dalam korpus beserta frekuensi kemunculannya dapat dilihat pada tabel di bawah.

Tabel 5 Daftar Adjektiva Berprefiks *meng-*

Bentuk Berprefiks	Bentuk Dasar	Jumlah Kemunculan	Persentase
menengah (adj)	tengah (n)	3017	19.63%
melingkar (adj)	lingkar (n)	313	2.04%
meriah (adj)	riah (?)	245	1.59%
mencuat (adj)	cuat (?)	236	1.54%
menggigil (adj)	gigil (adj)	218	1.42%
merinding (adj)	rinding (?)	130	0.85%
mendunia (adj)	dunia (n)	76	0.49%
meradang (adj)	radang (?)	74	0.48%
menegak (adj)	tegak (adj)	21	0.14%
menyemak (adj)	semak (n)	7	0.05%
melintir (adj)	lintir (?)	5	0.03%
mengganda (adj)	ganda (adj)	4	0.03%
melunglai (adj)	lunglai (adj)	2	0.01%
menggeledak (adj)	geledak (n)	2	0.01%

Dilihat dari bentuk dasar kata-kata dalam daftar, proses pembentukan adjektiva dengan prefiks *meng-* bersifat infleksional dan derivasional. Prefiks *meng-* dapat membentuk adjektiva dari nomina, adjektiva, dan prakategorial.

Prefiks <i>meng-</i>	+	nomina	→	adjektiva
meng-	+	tengah	→	menengah
meng-	+	lingkar	→	melingkar
meng-	+	dunia	→	mendunia
meng-	+	semak	→	menyemak
meng-	+	geledek	→	menggeledek
Prefiks <i>meng-</i>	+	adjektiva	→	adjektiva
meng-	+	gigil	→	menggigil
meng-	+	tegak	→	menegak
meng-	+	ganda	→	mengganda
meng-	+	lunglai	→	melunglai
Prefiks <i>meng-</i>	+	prakategorial	→	adjektiva
meng-	+	riah	→	meriah
meng-	+	cuat	→	mencuat
meng-	+	rinding	→	merinding
meng-	+	radang	→	meradang
meng-	+	lintir	→	melintir

Beberapa sumber menjelaskan bahwa adjektiva dapat bergabung dengan partikel tidak, menerangkan nomina, dan secara umum dapat bergabung dengan kata lebih, sangat, dan agak [13], [14]. Adjektiva-adjektiva dalam daftar di atas keseluruhannya menerangkan nomina, tetapi sebagian besar tidak berkolokasi dengan kata tidak, lebih, sangat, atau agak.

Makna yang dibentuk dari prefiks *meng-* terhadap daftar adjektiva di atas terlihat berbeda-beda berdasarkan kelas kata pembentuknya. Makna yang terbentuk adalah sebagai berikut.

meng- + nomina

1. ‘berada di posisi atau urutan yang tertera dalam kata dasar’: menengah
2. ‘menyerupai atau membentuk’: melingkar, menyemak, menggeledak
3. ‘terkenal’: mendunia

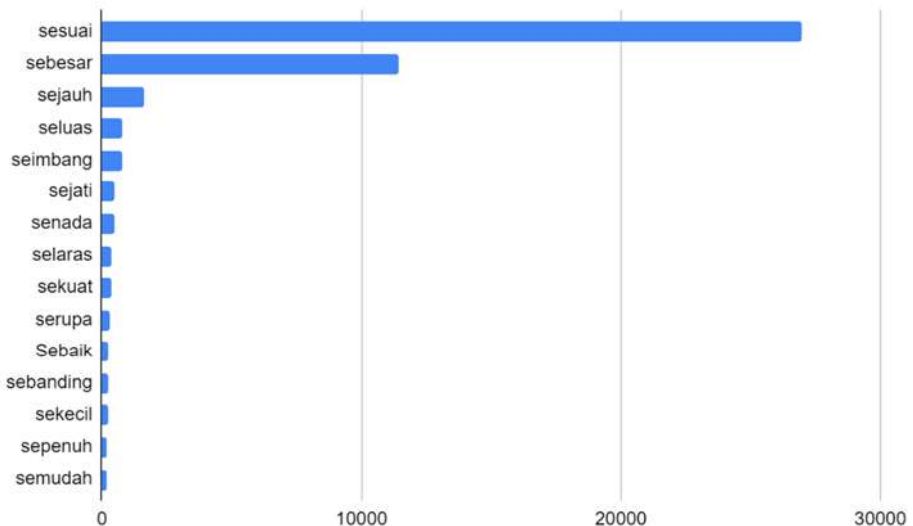
meng- + adjektiva

‘menjadi seperti yang tertera dalam kata dasar’: menggigil, menegak, mengganda, melunglai

meng- + prakategorial membentuk kata sifat yang tidak dapat didefinisikan oleh kata dasarnya.

3.2 Prefiks *se-*

Jumlah adjektiva berprefiks *se-* yang terdapat dalam korpus TBIK cukup banyak, 15,3% dari total adjektiva berafiks yang menjadi data atau sekitar 182 kata dengan frekuensi kemunculan tertinggi 26.970 kali. Hal ini menunjukkan bahwa prefiks *se-* adalah prefiks yang produktif untuk membentuk adjektiva. Lima belas kata dengan frekuensi kemunculan tertinggi dapat dilihat dalam grafik di bawah.



Grafik 2 Daftar kata berprefiks *se-* terbanyak dalam korpus

Dari grafik terlihat bahwa adjektiva berprefiks *se-* dapat terbentuk dari kata dasar berupa adjektiva maupun nomina. Kata sesuai malah terbentuk dari prefiks *se-* + prakategorial. Dengan demikian, proses pembentukan adjektiva dengan prefiks *se-* bersifat infleksional dan derivasional.

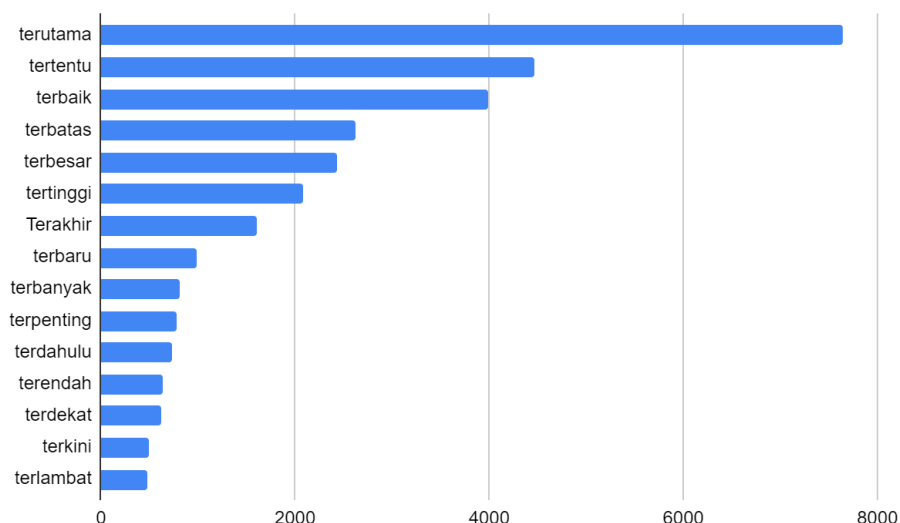
Makna yang terbentuk dari proses afiksasi dengan prefiks *se-*, jika dilihat dari daftar yang ada, hanya satu, yaitu ‘sama’, seperti *sekuat* ‘sama kuat’, *serupa* ‘sama rupa’, *sebaik* ‘sama baik’, dan *sebesar* ‘sama besar’. Seluruh kata dalam daftar di atas dapat dengan mudah dinegasikan dengan membubuhkan kata tidak di depannya.

Uniknya, KBBI [13] tidak mencatat *se-* sebagai prefiks pembentuk adjektiva. *Se-* dianggap sebagai bentuk terikat untuk menyatakan ‘satu’, seperti pada *sekamar*, *sekelas*, *serumah*, dan bentuk terikat untuk menyatakan ‘sama’, seperti pada *sependai*, *setinggi*, *secerdas*. Selain itu, KBBI [13] mencatat *se-* sebagai prefiks pembentuk adverbia.

Se- yang mengandung makna ‘satu’ boleh jadi merupakan bentuk terikat. Namun untuk yang bermakna ‘sama’, *se-* merupakan prefiks. Selain itu, Sasangka dkk [16] menyatakan bahwa salah satu ciri adjektiva adalah dapat bergabung dengan kata pengingkar *tidak*. Kata-kata seperti *sesuai*, *sebesar*, *sejauh*, dan *seluas* dapat dinegasikan dengan kata tidak.

3.3 Prefiks *ter-*

Pencarian dengan simbol *ter+_JJ* menghasilkan 530 tipe dan 41.548 token. Akan tetapi, dari 530 tipe itu, jumlah bentuk berprefiks hanya 348, sisanya adalah kata dasar. Ke-348 bentuk berprefiks *ter-* itu tidak seluruhnya berkelas kata adjektiva. Lima belas adjektiva berprefiks *ter-* dengan frekuensi kemunculan tertinggi dapat dilihat dalam grafik di bawah.



Grafik 3 Daftar kata berprefiks *ter-* terbanyak dalam korpus

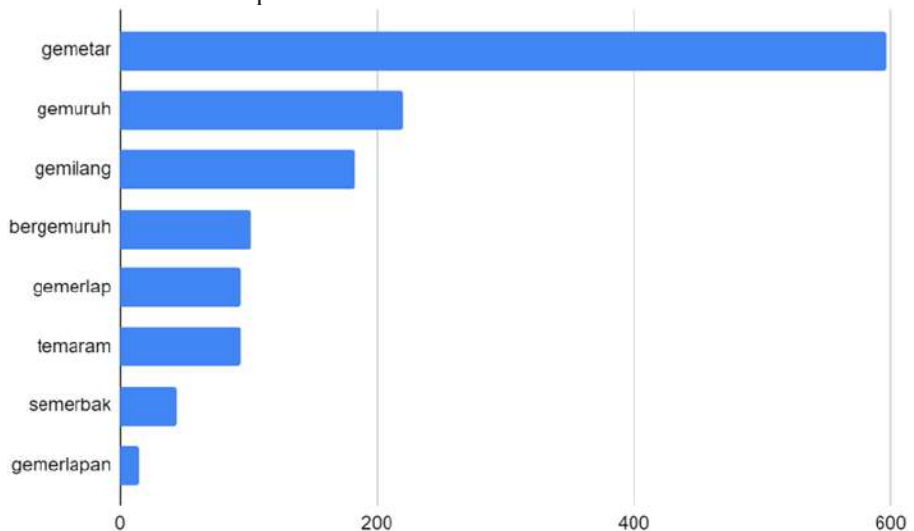
Untuk membentuk adjektiva, prefiks *ter-* dapat digabungkan pada kata dasar berupa nomina, verba, dan adjektiva. Ini membuktikan bahwa proses pembentukan adjektiva dengan penambahan prefiks *ter-* bersifat infleksional dan derivasional. Hal ini ditemukan dalam data yang diperoleh dari korpus seperti berikut.

Prefiks <i>ter-</i>	+	nomina	→	adjektiva
<i>ter-</i>		<i>akhir</i>		<i>terakhir</i>
<i>ter-</i>		<i>kini</i>		<i>terkini</i>
<i>ter-</i>		<i>depan</i>		<i>terdepan</i>
Prefiks <i>ter-</i>	+	verba	→	adjektiva
<i>ter-</i>		<i>percaya</i>		<i>terpercaya</i>
<i>ter-</i>		<i>gesa</i>		<i>tergesa</i>
Prefiks <i>ter-</i>	+	adjektiva	→	adjektiva
<i>ter-</i>		<i>utama</i>		<i>terutama</i>
<i>ter-</i>		<i>tentu</i>		<i>tertentu</i>
<i>ter-</i>		<i>baik</i>		<i>terbaik</i>

Prefiks *ter-* membentuk adjektiva superlatif. Ada pendapat yang menyebutkan bahwa prefiks *ter-* dapat menggantikan bentuk superlatif ‘paling’ jika ditambahkan pada kata dasar adjektiva. Jika ditambahkan pada kata dasar lain, seperti verba dan nomina, prefiks *ter-* biasanya akan menghasilkan verba yang menunjukkan keadaan atau kondisi akhir atas suatu perbuatan yang tidak mementingkan subjeknya, seperti kata *tercatat* atau *terbatas* [15]. Pendapat tersebut tidak sesuai dengan data, karena bentuk superlatif ‘paling’ dapat dihasilkan dari kata dasar berkelas adjektiva, nomina, ataupun verba seperti dalam daftar di atas.

3.4 Infiks *-em-*

Infiks adalah afiks yang kurang produktif dalam bahasa Indonesia. Hal ini ditunjukkan dengan tidak ditemukannya kata baru yang dibentuk dari infiks. Dari beberapa jenis infiks, satu-satunya infiks yang membentuk adjektiva adalah *-em-*. Grafik berikut menunjukkan adjektiva berinfiks *-em-* yang ditemukan dalam korpus TBIK.



Grafik 4 Daftar kata berinfiks *-em-* terbanyak dalam korpus

Tidak ada kata baru yang terjaring dalam korpus untuk bentuk berinfiks *-em-*. Frekuensi penggunaan katanya pun tidak terlalu banyak. Frekuensi tertinggi tidak mencapai 600 kali.

Informasi yang dapat diambil dari data adalah adjektiva dalam data terbentuk dari proses penambahan infiks *-em-* terhadap kata dasar berupa nomina, verba, atau adjektiva. Seperti yang telah dijelaskan dalam TBBI [3], dalam pembentukan kata dengan infiks *-em-*, bentuk dasar nomina dapat berdiri sendiri tanpa penyisipan *-em-* terlebih dahulu, sementara itu, bentuk dasar verba dan adjektiva tidak dapat berfungsi sebagai unsur sintaksis sebelum disisipi *-em-*.

adjektiva	+ <i>-em-</i>	adjektiva
<i>gilang</i>	+ <i>-em-</i>	<i>gemilang</i>
<i>gerlap</i>	+ <i>-em-</i>	<i>gemerlap</i>
nomina	+ <i>-em-</i>	adjektiva
<i>getar</i>	+ <i>-em-</i>	<i>gemetar</i>
<i>guruh</i>	+ <i>-em-</i>	<i>gemuruh</i>
<i>taram</i>	+ <i>-em-</i>	<i>temaram</i>
verba	+ <i>-em-</i>	adjektiva
<i>serbak</i>	+ <i>-em-</i>	<i>semerbak</i>

Infiks *-em-* yang disisipkan pada kata dasar di atas membentuk adjektiva yang mengandung makna kondisi yang terjadi berulang atau terus menerus, seperti pada *gemerlap*, *gemetar*, dan *gemuruh*; kondisi yang luar biasa, seperti *gemilang*; dan menggambarkan suasana (berdasarkan pancaindra), seperti *temaram* dan *semerbak*.

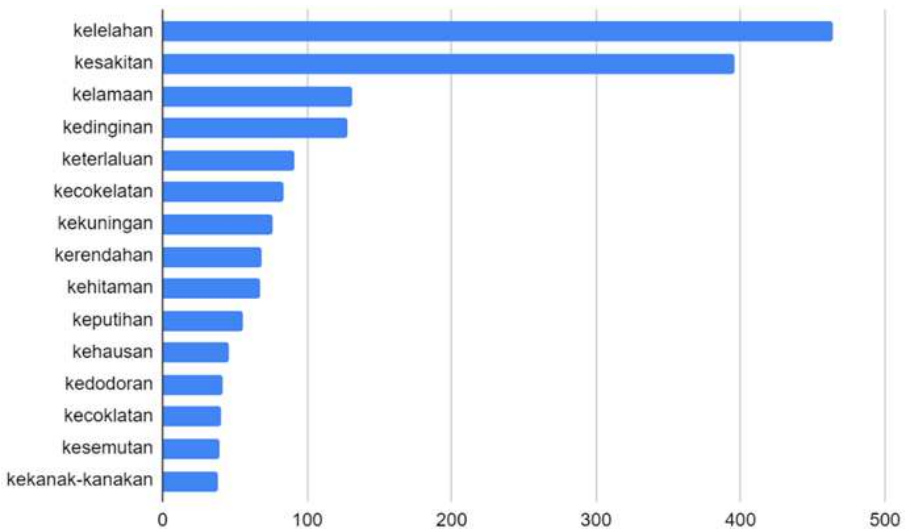
Nuansa makna kata *gemilang* dapat dilihat melalui penggunaan kata itu dalam konteks dari baris konkordansi berikut.

pembenahan, meski mengawali permainan dengan	gemilang	, " kata pelatih Italia, Cesare Prandelli.
Persela lewat tendangan Nzekou . Namun, secara	gemilang	Choirul Huda mampu menggagalkan peluang
hingga 1990-an. Saat itu sektor pertanian secara	gemilang	menunjukkan peningkatan produktivitas
Venus Williams, memulai tampil lagi dengan	gemilang	di Sony Ericsson Open, yang berlangsung
adalah melalui lifter Mela Eka Rahayu yang tampil	gemilang	di Invitation Powerlifting Championship
Sebaliknya , buat Oldham , ini adalah prestasi	gemilang	. " Kami memulai pertandingan dengan baik
kiper Julio Cesar, yang malam itu tampil	gemilang	. " Cesar luar biasa. Ia pemain kelas dunia
mencuat setelah berhasil mencetak prestasi	gemilang	di nomor ganda putra. Pada SEA Games 1991,
nasional, dapat menuju masyarakat yang	gemilang	, dengan pendapatan yang tinggi disertai
2015 Timnas Indonesia bermain cukup	gemilang	. Dari lima pertandingan dengan sistem
sedangkan Benny Wijaya, yang tampil	gemilang	bagi NSH, mengatakan mereka sudah
Nainggolan dan Mauro Icardi . Bermain	gemilang	bersama klub sepanjang musim, bukan
dari satu tahun terakhir setelah tampil	gemilang	di timnas U-19 asuhan Indra Sjafri. Meski
timnya. Radovic menyebut Arema tampil	gemilang	walaupun bermain di depan pendukung Persib.
generasi mdatang dapat menorehkan prestasi	gemilang	di banyak cabor. Secara khusus, Eko ingin

Gemilang berkolokasi dengan nomina *prestasi* dan verba *tampil*. Subjek yang diterangkan adalah *seorang* individu atau satu tim yang kebanyakan berada di bidang olahraga. Fitur distribusi dalam CQPWeb menunjukkan bahwa kata *gemilang* banyak muncul di laman resmi, koran, dan majalah.

3.5 Konfiks *ke-* *-an*

Selain membentuk nomina, konfiks *ke-* *-an* juga dapat membentuk adjektiva ketika ditambahkan pada kata dasar adjektiva. Data korpus menunjukkan bahwa selain pada kata dasar adjektiva, konfiks *ke-* *-an* juga dapat ditambahkan pada kata dasar berupa nomina. Hasil pencarian adjektiva berkonfiks *ke-* *-an* pada korpus menghasilkan sekitar 67 kata, yang kata dasarnya berupa adjektiva dan nomina. Grafik berikut menunjukkan lima belas kata teratas dalam daftar, berdasarkan jumlah frekuensi kemunculan.



Grafik 5 Daftar kata berkonfiks *ke-* *-an* terbanyak dalam korpus

Sebagian besar adjektiva di atas terbentuk dari kata dasar berupa adektiva juga, seperti *lelah*, *sakit*, *lama*, *dingin*, *terlalu*, *rendah*, dan *haus*. Kosakata warna dalam daftar seperti *cokelat*, *kuning*, *hitam*, dan *putih* dapat berkelas kata nomina maupun adjektiva. Sisanya adalah adjektiva dengan kata dasar nomina, yaitu *semut*, dan *kanak-kanak*. Tidak ditemukan adjektiva berkonfiks *ke-* *-an* yang terbentuk dari kata dasar verba dalam data korpus. Walaupun

demikian, dapat dikatakan bahwa konfiks *ke-* *-an* membentuk adjektiva secara infleksional dan derivasional.

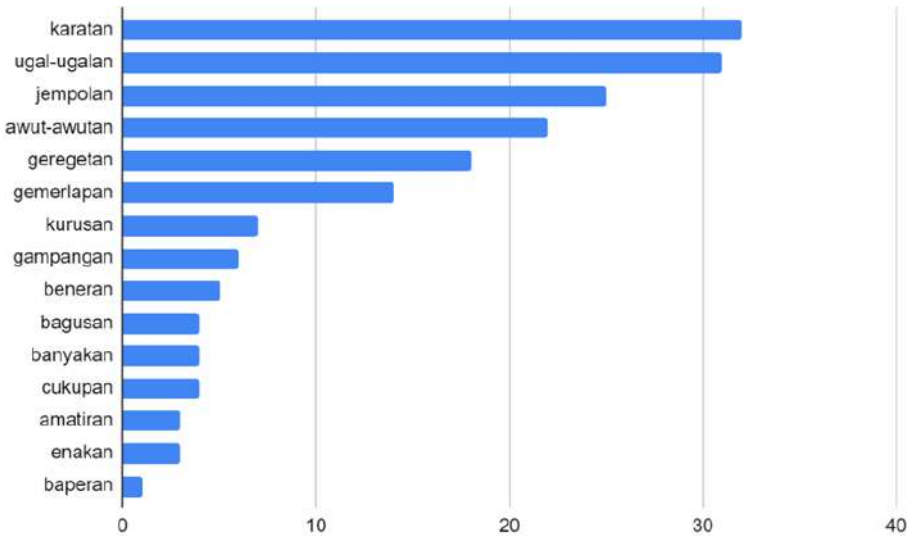
KBBI menyebutkan bahwa makna yang dibentuk dari konfiks *ke-* *-an* terhadap adjektiva adalah ‘terlalu’, seperti *kelamaan*, *kerendahan*, *kebesaran*, *kesempitan*, dan *kepagian*. Namun, data menunjukkan ada beberapa makna lain, yaitu:

- | | |
|--------------------------------------|---|
| 1. merasa sangat ... | <i>kehausan, kedinginan, kesakitan,</i>
<i>kecapean, kegerahan</i> |
| 2. agak | <i>kehitaman, kekuningan, kemerahan,</i>
<i>keputihan</i> |
| 3. bersifat seperti | <i>kekanak-kanakan, keibuan</i> |
| 4. bertingkah laku seperti orang ... | <i>kebarat-baratan, kebelanda-</i>
<i>belandaan.</i> |

Dengan demikian ada setidaknya lima makna yang dapat dikandung oleh konfiks *ke-* *-an* ketika membentuk adjektiva.

3.6 Sufiks *-an*

Tidak banyak literatur yang mencatat sufiks *-an* sebagai sufiks pembentuk adjektiva. Data korpus menunjukkan bahwa adjektiva bersufiks *-an* memang ada dan digunakan oleh masyarakat. Grafik 6 di bawah menunjukkan beberapa adjektiva bersufiks *-an* dengan frekuensi kemunculan tertinggi.



Grafik 6 Daftar kata bersufiks *-an* terbanyak dalam korpus

Putra dan Utami [10] menyebutkan bahwa adjektiva yang terbentuk dari penambahan sufiks *-an* pada kata dasar menghasilkan kata cakapan atau kata takbaku. Namun hal itu tidak seluruhnya tepat. Kata-kata dalam daftar memang sebagian besar adalah kata cakapan, seperti *awut-awutan*, *beneran*, *bagusan*, *banyakan*, *enakan*, dan *baperan* yang kemungkinan berasal dari pengaruh bahasa daerah seperti bahasa Jawa atau bahasa Betawi. Namun, pencarian data dalam korpus juga menghasilkan kata *jempolan*. *Jempolan* tidak hanya digunakan dalam ragam cakapan, tetapi juga dalam ragam standar. Distribusi penggunaan kata *jempolan* dalam korpus menunjukkan kata ini muncul di enam jenis teks, yaitu koran, majalah, cerpen, novel, biografi, dan teks populer.

Dilihat dari proses pembentukannya, adektiva bersufiks *-an* dibentuk dari kata dasar berkelas kata adjektiva dan nomina. Ini membuktikan bahwa proses pembentukan adjektiva dengan penambahan sufiks *-an* bersifat infleksional dan derivasional. Perinciannya dapat dilihat sebagai berikut.

adjektiva	+ -an	adjektiva
<i>gereget</i>	+ -an	<i>geregetan</i>
<i>gemerlap</i>	+ -an	<i>gemerlapan</i>
<i>kurus</i>	+ -an	<i>kurusan</i>
<i>gampang</i>	+ -an	<i>gampang</i>

<i>bener</i>	+ -an	<i>beneran</i>
<i>banyak</i>	+ -an	<i>banyakan</i>
<i>cukup</i>	+ -an	<i>cukupan</i>
<i>amatir</i>	+ -an	<i>amatiran</i>
<i>enak</i>	+ -an	<i>enakan</i>
<i>baper</i>	+ -an	<i>baperan</i>
nomina	+ -an	adjektiva
<i>karat</i>	+ -an	<i>karatan</i>
<i>jempol</i>	+ -an	<i>jempolan</i>
<i>panu</i>	+ -an	<i>panuan</i>

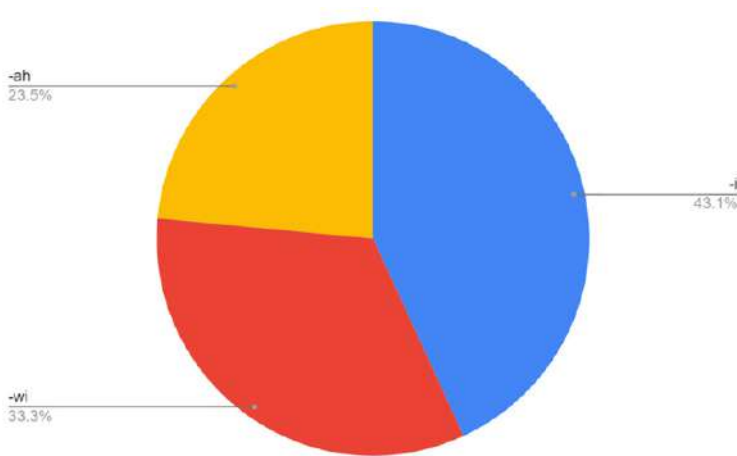
Makna dari adjektiva yang terbentuk dari penambahan sufiks *-an* pada nomina sama dengan verba *ber-* yang artinya ‘mempunyai’: *karatan* = *berkarat*, *panuan* = *berpanu*. Namun, *jempolan* ≠ *berjempol*. *Jempolan* memiliki makna kias ‘hebat’ atau ‘bagus sekali’. Sementara itu, makna dari adjektiva yang terbentuk dari penambahan sufiks *-an* pada adjektiva adalah

1. agak atau lebih
kurusan ‘agak kurus’ atau ‘lebih kurus’
banyakan ‘agak banyak’ atau ‘lebih banyak’
enakan ‘agak enak’ atau ‘lebih enak’
gampang ‘lebih gampang’
2. mudah
baperan ‘mudah baper’
3. secara
amatiran ‘secara amatir’
4. makna lain
beneran ‘betul-betul’, *gemerlapan* ‘berkilauan’,
geregetan ‘geram, kesal, jengkel’.

3.7 Sufiks serapan *-i*, *-wi*, *-iah*

Tata Bahasa Baku Bahasa Indonesia [3] juga KBBI Edisi V [13] menyatakan bahwa sufiks serapan dari bahasa Arab *-i*, *-wi*, dan *-iah* memiliki fungsi dan yang sama dalam bahasa Indonesia, yaitu membentuk adjektiva yang ‘berkenaan dengan kata dasarnya atau berdasar pada’. Untuk mengonfirmasi pernyataan tersebut, telah dilakukan pemeriksaan data korpus pada masing-masing sufiks dengan hasil sebagai berikut.

Adjektiva bersufiks *-i* ditemukan lebih banyak dalam korpus daripada yang bersufiks *-wi*, dan *-iah* dengan perbandingan seperti dalam grafik 7.



Grafik 7 Perbandingan jumlah adjektiva bersufiks *-i*, *-wi*, dan *-iah*

Sebagian besar adjektiva yang terjaring dari korpus adalah kosakata serapan Arab yang memang mendapat imbuhan sufiks *-i*, *-wi*, dan *-iah*, seperti *alami*, *islami*, *hayati*, *hewani*; *manusiawi*, *duniawi*, *kimiawi*, *surgawi*; *ilmiah*, *harfiah*, *jasmaniah*, *rohaniah*. Namun, dalam data korpus ditemukan juga bentukan kata yang bukan serapan langsung dari bahasa Arab, seperti berikut.

sufiks	kata dasar	kata bentukan
<i>-i</i>	<i>kristian</i>	<i>kristiani</i>
<i>-wi</i>	<i>manusia</i>	<i>manusiawi</i>
<i>-wi</i>	<i>surga</i>	<i>surgawi</i>
<i>-wi</i>	<i>indra</i>	<i>indrawi</i>
<i>-iah</i>	<i>alami</i>	<i>alamiah</i>
<i>-iah</i>	<i>naluri</i>	<i>naluriah</i>

Beberapa dari kata di atas sepertinya terbentuk dari hasil analogi. *Kristiani* terbentuk karena beranalogi pada kata *islami*. *Naluriah* bisa jadi merupakan

analogi dari *jasmaniah* dan *rohaniah*. Kata *alamiah* tidak ditemukan dalam bahasa Arab. Dalam bahasa Arab, adjektiva dari *alam* adalah *alami*. *Alami* juga digunakan dalam bahasa Indonesia. Penambahan afiks *-iah* pada kata *alamiah* bisa jadi merupakan bentuk analogi.

Seluruh adjektiva bersufiks *-i*, *-wi*, dan *-iah* terbentuk dari nomina. Dengan demikian, sufiks serapan dari bahasa Arab ini merupakan sufiks derivatif karena dapat membentuk adjektiva dari nomina.

KBBI Edisi V [13] menakrifkan sufiks *-iah* sama dengan sufiks *-i*, dan *-wi*, yaitu sufiks pembentuk adjektiva yang menghasilkan makna ‘berkenaan dengan’ atau ‘berdasar pada’. Makna itu dapat dikonfirmasi melalui pemakaian kata dalam konteks melalui analisis konkordansi dan kolokasi. Dalam hal ini, konfirmasi dilakukan melalui kata *alamiah*. Distribusi dari kata *alamiah* dan sepuluh kolokasi teratas dari kata *alamiah* dapat dilihat dalam tabel 6 dan tabel 7 berikut.

Tabel 6 Distribusi Kata *alamiah* dalam Korpus

Distribution of hits for query "alamiah" returned 437 matches in 222 different texts Currently displaying distribution across text classifications.				
Display distribution of:	Text categories	Show as:	Distribution table	
Cross-tabulating against:		Other actions:	Choose action	
Based on classification: <i>number</i>				
Category [.]	Words in category	Hits in category	Dispersion (no. texts with 1+ hits)	Frequency [.] per million words in category
A_koran	2,823,992	23	21 out of 6,689	8.14
B_majalah	2,439,621	26	22 out of 2,835	10.66
C_cerpen	2,504,297	14	13 out of 1,457	5.59
D_novel	2,622,758	28	13 out of 38	10.68
E_buku_teks	2,467,943	69	24 out of 63	27.96
F_jurnal	2,410,112	91	50 out of 490	37.76
G_disertasi_tesis_skripsi	2,703,612	71	29 out of 85	26.26
H_biografi	2,450,463	46	13 out of 53	18.77
I_populer	2,290,231	28	14 out of 66	12.23
J_perundangan	2,500,251	31	15 out of 196	12.40
K_jamari_resmi	2,440,866	10	8 out of 4,698	4.10
Total:	27,654,146	437	222 out of 16,670	15.80

No.	Word	Total no. in whole corpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Mutual information
1	jurnal-jurnal	34	0.013	11	8	9.736
2	orasi	79	0.030	21	15	9.456
3	Forum-forum	52	0.020	7	2	8.473
4	KIR	48	0.018	5	4	8.102
5	artikel	2,123	0.805	192	22	7.899
6	Musababah	81	0.031	7	3	7.833
7	mempraktekkan	93	0.035	8	1	7.828
8	Publikasi	1,014	0.384	84	45	7.772

Kata *alamiah* terdistribusi ke seluruh jenis teks yang ada dalam korpus TBIK, yaitu koran, majalah, cerpen, novel, buku teks, jurnal, disertasi, tesis, skripsi, biografi, populer, perundangan, dan laman resmi. Kata ini juga muncul setiap tahun. Hal ini menandakan bahwa kata *alamiah*, yang merupakan kata bentukan, berterima di masyarakat Indonesia, tidak kalah dari bentuk serapan aslinya, *alami*.

Tabel 7 Kolokasi Kata *alamiah* dalam Korpus

No.	Word	Total no. in whole corpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Mutual information
1	jurnal-jurnal	34	0.013	11	8	9.736
2	orasi	79	0.030	21	15	9.456
3	Forum-forum	52	0.020	7	2	8.473
4	KIR	48	0.018	5	4	8.102
5	artikel	2,123	0.805	192	22	7.899
6	Musababah	81	0.031	7	3	7.833
7	mempraktekkan	93	0.035	8	1	7.828
8	Publikasi	1,014	0.384	84	45	7.772
9	Jurnal	1,055	0.400	83	44	7.698
10	metode-metode	64	0.024	5	4	7.685
11	tulis	1,851	0.702	132	39	7.556
12	DEBAT	555	0.210	39	6	7.535

Sebagian besar kolokasi dari ilmiah berkelas kata nomina sehingga membentuk frasa nomina seperti *jurnal-jurnal ilmiah*, *orasi ilmiah*, *forum-forum ilmiah*, *KIR ilmiah*, *artikel ilmiah*, dan *musabaqah ilmiah*.

Selain berkolokasi dengan nomina, adjektiva berafiks *-i*, *-wi*, *-iah* juga berkolokasi dengan partikel secara dan verba bersifat seperti yang dapat dilihat dari baris konkordansi kata *lahiriah* dan *naluriah* berikut.

tahap ini ialah yang bersifat	lahiriah	saja dan bisa mengalami banyak
atau kepasrahan dan ketaatan secara	lahiriah	kepada hukum Tuhan serta mewajibkan
, dan penampilan mereka secara	lahiriah	. Kepala sekolah profesional harus
) hubungan manusia dengan kemajuan	lahiriah	dan kepuasan rokhaniah . Menurut
dalam tahanan . Kasimo secara	lahiriah	termasuk yang lebih beruntung karena
sifatnya hanya semu . Secara	lahiriah	ia memang mempunyai seorang isteri
, tetapi semua itu hanya	lahiriah	. Pertemuan itu tidak menghasilkan
Atau tidak sekadar dengan penampilan	lahiriah	, termasuk bentuk pakaian atau
SWT , termasuk setelah usaha	lahiriah	dalam mencari rizki . Ketentuan-ketentuan
burhaniah (dalil-dalil yang bersifat	lahiriah) . Isyraqiyah menggunakan media
lantaran mereka hanya melihat secara	lahiriah	ayat Al-Qur'an . Karena itu
melihat bahwasanya ciri fisik atau	lahiriah	lebih mengidentifikasi mereka sebagai suatu
ideologi bermuatan materialis atau bersifat	lahiriah	adalah dengan mengimplementasikan nilai-ni

. Bagi Miqdad , secara	naluriah	, orang terus berproses dalam
dan investasi . Bank secara	naluriah	akan mengejar dan mencari kemanapun
seks bukan sekadar memenuhi hasrat	naluriah	semata . Ada makna yang
, dan tersengal . Secara	naluriah	, ia meraih kaki perempuan
adalah bentuk pertahanan diri secara	naluriah	. Gray mengambil selebar foto
dari pohon , sang beruk— secara	naluriah	dan ajaib—menuntunnya melangkah ke makam
tanpa komando , ada perasaan	naluriah	manusia yang diam-diam tumbuh di
untuk pertama-kalinya , mereka secara	naluriah	, akan jatuh cinta dengan
telah menjadi suatu hal yang	naluriah	serta tak perlu dipermasalahkan .
rasa penasaran , hanya keinginan	naluriah	belaka . Tiap kali badannya

3.8 Sufiks Serapan -is

Adjektiva bersufiks *-is* sebagian besar merupakan adjektiva serapan dari bahasa Belanda yang berakhiran *-isch* seperti *dramatisch* atau dari bahasa Inggris yang berakhiran *-ic* dan *-ical* seperti *academic* dan *academical*. Dalam bahasa Indonesia, *-isch* (Belanda), dan *-ic*, *-ical* (Inggris) menjadi *-is* [12].

Kata bentukan dengan penambahan sufiks ini yang ditemukan dalam korpus bahasa Indonesia hanya sedikit sekali. Dari data adjektiva bersufiks *-is* hanya ditemukan bentuk kata *agamis*, *agamaais*, dan *salafis* dalam konteks berikut.

itu orang yang mengaku sangat	agamis	dan sepat terjangnya juga selalu
diperoleh dari statusnya sebagai seorang	agamis	yang saleh tetapi juga dari
demonstran . ”Ikhwatul Muslimin dan	Salafis	memang tidak banyak , tetapi
sebuah kelompok dan antara kaum	agamis	sebagai sebuah kelompok , dan
kita baca Alkitab , secara	agamis	maupun sekuler , kita tahu
Novita Siswayanti menekankan nilai-nilai	agamaais	dalam diri seorang guru seperti
yang lebih baik , lebih	agamis	. Jilbab menjadi sebuah pembentuk
ini pakaian seragam menjadi makin	agamaais	(jilbab atau rok panjang
Tapin Mandiri , Sejahtera dan	Agamaais	(E-Tamasa) yang dipresentasikan
profesi guru untuk mengedepankan nilai	agamaais	dalam segenap aktivitas pembelajaran teruta

Bentuk afiksasi yang terjadi melalui sufiks *-is* dari data di atas adalah bentuk afiksasi derivatif, yaitu

nomina	+	<i>-is</i>	→	adjektiva
<i>agama</i>	+	<i>-is</i>	→	<i>agamis, agamaais</i>
<i>salaf</i>	+	<i>-is</i>	→	<i>salafis</i>

Makna yang dikandung oleh sufiks *-is* dalam membentuk adjektiva dalam bahasa Indonesia adalah menerangkan sesuatu yang berkaitan dengan kata dasarnya, misal *agamaais* ‘berkaitan dengan agama’, *salafis* ‘berkaitan dengan manhaj salaf’.

Sedikitnya jumlah kata baru yang dapat dibentuk dengan penambahan sufiks *-is* menandakan bahwa sufiks ini tidak cukup produktif untuk membentuk kata baru di luar kata serapan. Sementara itu, untuk kata serapan sendiri, bentuk bersufiks *-is* masih bersaing dengan bentuk bersufiks *-ik*. Hal ini akan dijelaskan dalam tulisan yang lain.

4. SIMPULAN

Dari hasil analisis yang didapatkan dapat disimpulkan beberapa hal sebagai berikut.

1. Afiks pembentuk adjektiva dalam bahasa Indonesia adalah
 - prefiks *meng-, se-, dan ter-*
 - infiks *-em-*
 - konfiks *ke- -an*
 - sufiks *-an, -i, -wi, -iah, dan -is.*

2. Afiks-afiks pembentuk adjektiva itu bersifat infleksional dan derivasional kecuali sufiks serapan *-i*, *-wi*, *-iah*, dan *-is* yang hanya bersifat derivasional.
3. Makna yang terbentuk dari penambahan afiks lebih beragam daripada makna yang telah tertera dalam literatur sebelumnya.

Penggunaan data yang bersumber dari korpus mengonfirmasi hampir seluruh pendapat yang telah ada dalam literatur terkait pembentukan adjektiva melalui afiksasi, sekaligus menambah dan melengkapi beberapa hal yang belum ada, terutama terkait dengan penghitungan statistik. Penggunaan dua pendekatan sekaligus, kuantitatif dan kualitatif dapat lebih menguatkan hipotesis sehingga hasilnya lebih reliabel.

Referensi

- [1] D. Suparno, *Morfologi Bahasa Indonesia*, Jakarta: UIN Press, 2015.
- [2] Bagiya, "Infleksi dan Derivasi dalam Bahasa Indonesia," *Journal of Language Learning and Research*, vol. I, no. 1, pp. 1-11, 2018.
- [3] H. L. H. A. S. W. S. S. Anton M. Moeliono, *Tata Bahasa Baku Bahasa Indonesia*. Edisi keempat, Jakarta: Badan Pengembangan dan Pembinaan Bahasa, 2017.
- [4] D. Resnita, "Pembentukan Adjektiva Denominal dalam Bahasa Indonesia.," *JURNAL KIP*, vol. IV, no. 2, 2015.
- [5] R. B. S. P. d. E. Utami, "Non-formal Affixed Word Stemming in Indonesian Language," in *International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, 2018.
- [6] D. Purnanto, "Kajian Morfologi Derivasional dan Infleksioanal dalam Bahasa Indonesia," *Kajian Linguistik dan Sastra*, vol. 18, no. 35, pp. 136-152, 2006.
- [7] K. A. Ni Wayan Suastini, "Indonesian prefix *ter-*," *Journal of Applied Studies in Language*, vol. 2, no. 2, pp. 192-196, 2016.
- [8] Ermanto, "Proses Morfologi Infleksi pada Adjektiva Bahasa Indonesia," *Humanus*, vol. XV, no. 1, pp. 41-52, 2016.
- [9] Ermanto, "Proses Morfologi Derivasi pada Adjektiva Afiksasi Bahasa Indonesia," in *Bahasa, Sastra Indonesia, dan Pembelajarannya*, Padang, 2016.
- [10] M. N. B. A. J. T. S. M. M. W. H. E. Adriani, "Stemming Indonesian:

A confix-stripping approach,” *ACM Transactions on Asian Language Information Processing*, vol. 6, no. 4, 2007.

- [11] O. J. Ballance, “Methodological considerations for the use of mutual information: Examining the role of context in collocation research,” *Research Methods in Applied Linguistics*, vol. 1, no. 3, 2022.
- [12] M. A. R. Anton M. Moeliono, *Pedoman Umum Pembentukan Istilah*. Edisi Ketiga, Jakarta: Pusat Bahasa, 2007.
- [13] T. Penyusun, “Kamus Besar Bahasa Indonesia,” Badan Pengembangan dan Pembinaan Bahasa, 30 April 2022. [Online]. Available: <https://kbbi.kemdikbud.go.id/Beranda>. [Accessed Oktober 2022].
- [14] H. Kridalaksana, *Kelas Kata dalam Bahasa Indonesia*. Edisi Kedua, Jakarta: PT. Gramedia Pustaka Utama, 2005.
- [15] T. R. W. Serly Natalia, “Identifying types of affixes in English and Bahasa Indonesia,” *Holistic Journal*, vol. 9, no. 17, pp. 8-22, 2017.
- [16] Sry Satriya Tjatur Wisnu Sasangka; Nantje Harijati Widjaja; Titik Indiyastini. “Adjektiva dan adverbial dalam bahasa Indonesia”. Jakarta : Proyek Penelitian Bahasa dan Sastra Indonesia dan Daerah Pusat Pembinaan dan Pengembangan Bahasa Departemen Pendidikan dan Kebudayaan, 1999.

REDUPLIKASI UTUH DAN SEBAGIAN DALAM BAHASA INDONESIA

Dora Amalia dan Dewi Puspita

Badan Pengembangan dan Pembinaan Bahasa

Abstrak

Penelitian tentang reduplikasi berbasis korpus ini merupakan penelitian yang menguak penggunaan reduplikasi secara empiris. Dalam data korpus ditemukan banyak bentuk dan pola reduplikasi yang belum sepenuhnya dideskripsikan dalam tata bahasa baku bahasa Indonesia. Hal ini terjadi karena data yang dikumpulkan berasal dari berbagai macam genre sehingga meliputi bahasa Indonesia ragam formal maupun nonformal. Penelitian dilakukan dengan metode gabungan, antara kuantitatif dan kualitatif. Secara kuantitatif, data dihitung berdasarkan frekuensi kemunculan dan distribusinya. Hasil penghitungan tersebut kemudian diinterpretasikan secara kualitatif. Dari hasil penelitian ini ditemukan ada beberapa bentuk reduplikasi yang secara kreatif diciptakan oleh pengguna bahasa untuk mengungkapkan makna yang hanya dapat dipahami jika diletakkan pada konteks penggunaannya. Selain itu, penggunaan reduplikasi sebagai bentuk simbolisme bunyi juga dibahas. Oleh sebab itu, penelitian ini tidak hanya dilakukan pada tingkat morfologi, melainkan juga pada tingkat sintaksis atau gramatikal.

Kata kunci: korpus linguistik, reduplikasi morfologi, reduplikasi sintaksis

1. PENDAHULUAN

1.1 Latar Belakang

Pembahasan tentang reduplikasi sudah banyak dilakukan, baik dalam konteks umum, maupun dalam kasus khusus untuk bahasa tertentu. Sebagai suatu fenomena bahasa yang universal, reduplikasi merupakan salah satu ciri yang cukup menonjol dari bahasa-bahasa yang termasuk rumpun Austronesia (Rubino, 2005).

Reduplication is very common throughout Austronesian (Pacific islands, Philippines, Indonesia, Madagascar), Australia, South Asia, and many parts of Africa, the Caucasus, and Amazonia. In the Western Hemisphere, some language families are particularly

amenable to reduplication (Salishan, Pomoan, Uto-Aztecan, Algonquian, Yuman, Sahaptian, Siouan, etc.) while others are not, such as Athabaskan and Eskimo-Aleut. Western Europe is one area where reduplication does not play a critical role in the morphology. However, many Indo-European languages in the east, which are in contact with other language families, do have reduplicative morphemes.

Gejala yang hampir sama dengan reduplikasi ini adalah repetisi. Gil (2011) menyatakan bahwa keduanya dibedakan dari satuan apa yang mengalami peristiwa itu. Dikatakan bahwa reduplikasi hanya terjadi pada tataran kata (morfologi), sedangkan repetisi dapat terjadi pada tataran di atasnya, bahkan klausa. Ada pula yang membedakannya dari cara mengucapkan. Jika perulangan itu diucapkan dalam satu nafas, tanpa jeda, berarti itu tergolong reduplikasi. Namun, jika perulangan tersebut diucapkan dengan jeda, dapat diduga hal itu merupakan ciri repetisi. Tentang hal ini Gil (2011) mengupasnya secara terperinci dengan mengambil contoh kasus bahasa Melayu Riau. Karena beragamnya pendekatan terhadap reduplikasi ini, dalam penelitian ini hanya akan dibahas tiga hal utama, yaitu kategori, pola pembentukan, dan makna reduplikasi. Dalam penelitian ini juga dibatasi hanya pada reduplikasi utuh dan reduplikasi berimbuhan, bukan reduplikasi berubah bunyi atau sejenisnya.

1.2 Tujuan

Artikel ini disusun dengan tujuan untuk mendeskripsikan kategori, pola pembentukan, dan makna reduplikasi kata dalam bahasa Indonesia berdasarkan korpus bahasa Indonesia kontemporer. Data yang terkumpul di dalam korpus meliputi bahasa Indonesia formal dan nonformal dalam kurun waktu sepuluh tahun terakhir. Oleh sebab itu, dalam penjelasan tulisan ini terdapat beberapa hal yang belum diatur dalam tata bahasa baku bahasa Indonesia.

1.3 Tinjauan Pustaka

1.3.1 Penelitian Terdahulu

Pembahasan tentang reduplikasi sudah banyak dilakukan, baik dalam bentuk tulisan umum tentang tata bahasa, khususnya pembentukan kata, maupun dalam bentuk penelitian khusus. Kelompok pertama meliputi referensi

klasik yang selama ini menjadi teori dasar dalam tata bahasa, khususnya reduplikasi dalam hal ini. Ada empat orang ahli yang perlu disebut dalam pembahasan ini, yaitu Kridalaksana (1985; 1989), Ramlan (1979), Parera (1988), dan Samsuri (1980).

Kelompok yang kedua terdiri atas pembahasan reduplikasi dalam konteks bahasa asing, bahasa Indonesia, dan bahasa daerah. Dalam konteks bahasa asing, Kallergi (2015) membahas khusus tentang reduplikasi dalam bahasa Yunani. Untuk konteks bahasa Indonesia, reduplikasi sudah diteliti oleh Anggraini (2019); Nita (2019) dan Fatimah (2017) yang keduanya meneliti reduplikasi dengan data dari media massa. Adapun untuk konteks bahasa daerah, reduplikasi sudah diteliti oleh Prasetiawan (2014) untuk bahasa Sasak; Firman (2014) untuk bahasa Moronene; dan Gil (2011) yang membahas tentang repetisi dan reduplikasi dalam bahasa Melayu Riau. Selain itu, penelitian reduplikasi dengan membandingkan bahasa Indonesia dengan bahasa daerah atau bahasa asing juga dilakukan oleh Falah (2016) dan Ong (2020).

1.4 Definisi Konseptual

Ada beberapa definisi konseptual yang dipakai untuk analisis dalam artikel ini. Konsep-konsep yang dimaksud adalah sebagai berikut.

1.4.1 Bentuk Dasar dan Pangkal

Di dalam analisis ini digunakan istilah *bentuk dasar* dan *pangkal*. Yang dimaksud dengan bentuk dasar adalah bentuk apa pun yang menjadi dasar reduplikasi, baik berupa kata dasar maupun kata turunan atau kata berimbuhan. Adapun pangkal adalah morfem atau kata yang bergabung dengan atau ditambahkan pada afiks. Oleh karena itu, kategorisasi reduplikasi secara formal atau berdasarkan bentuk ini pun mengikuti pembagian bentuk dasarnya, yaitu reduplikasi yang bentuk dasarnya berupa kata dasar dan reduplikasi yang berbentuk dasar kata turunan.

1.4.2 Reduplikasi

Beberapa ahli bahasa Indonesia memberikan batasan tentang reduplikasi. Di antaranya adalah Keraf (1991: 149) yang menyatakan bahwa reduplikasi merupakan sebuah bentuk gramatikal yang berujud penggandaan sebagai atau seluruh bentuk dasar sebuah kata. Adapun Ramlan (1979: 38) menyatakan

bahwa proses pengulangan atau reduplikasi merupakan pengulangan bentuk, baik seluruhnya maupun sebagian, baik dengan variasi fonem maupun tidak. Hasil reduplikasi itu berupa kata, dan bentuk yang diulang merupakan bentuk dasar. Samsuri (1988: 14) menyatakan bahwa reduplikasi merupakan pengulangan bentuk kata, yang dapat utuh atau sebagian, sedangkan Matthews (1978: 127) mengungkapkan bahwa reduplikasi merupakan repetisi yang dapat parsial tetapi dapat pula keseluruhan.

Dari pendapat beberapa ahli tersebut dapat disimpulkan bahawa reduplikasi adalah perulangan bentuk yang dapat terjadi secara utuh maupun sebagian yang memiliki fungsi tertentu secara semantis maupun sintaksis.

1.4.3 Kategori Reduplikasi

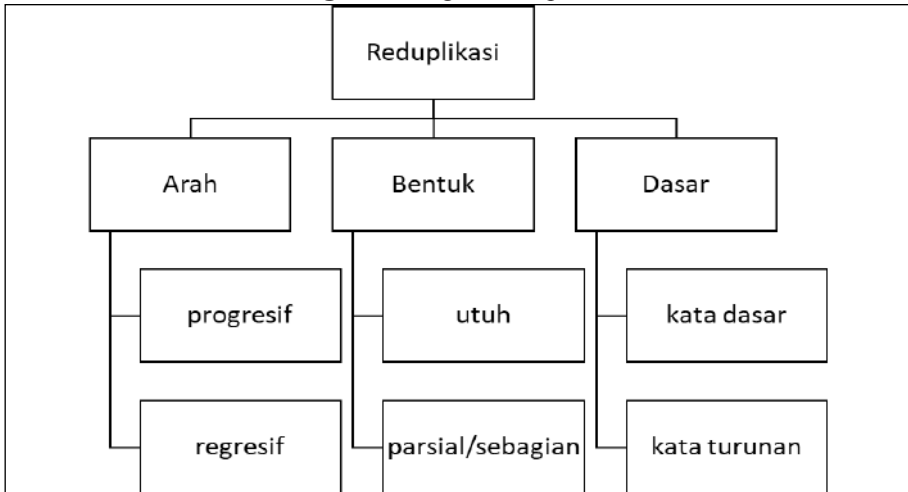
Berdasarkan arahnya, reduplikasi dapat terjadi sebelum bentuk dasar (disebut reduplikasi arah belakang atau regresif) dan setelah bentuk dasar (disebut arah depan atau progresif). Untuk reduplikasi yang setiap unsurnya sama sehingga tidak dapat ditentukan bentuk dasarnya, dikategorisasikan menurut arah ujaran, sehingga termasuk ke dalam reduplikasi progresif.

Berdasarkan bentuknya, reduplikasi dapat dikelompokkan menjadi tiga, yaitu (1) reduplikasi utuh dan (2) reduplikasi sebagian atau parsial. Reduplikasi dikatakan merupakan utuh bila bentuk reduplikasinya sama dengan bentuk dasarnya. Reduplikasi sebagian atau parsial adalah bentuk reduplikasi sebagian dari bentuk dasarnya.

Dilihat dari bentuk dasarnya, reduplikasi dapat dikelompokkan menjadi dua, yaitu (1) bentuk dasar asal dan (2) bentuk dasar turunan atau jadian. Reduplikasi yang memiliki bentuk dasar berupa bentuk asal disebut reduplikasi bentuk asal. Contohnya, *pipa-pipa*, *lari-lari*, *gadis-gadis* yang masing-masing mempunyai bentuk dasar *pipa*, *lari*, dan *gadis*. Reduplikasi yang memiliki bentuk dasar berupa bentuk turunan atau jadian disebut reduplikasi turunan atau jadian. Misalnya, *petani-petani*, *masukan-masukan*, dan *menangis-nangis* dengan bentuk dasarnya *petani*, *masukan*, dan *menangis* yang merupakan bentuk turunan atau jadian.

Jika disimpulkan, kategorisasi reduplikasi dalam analisis ini dapat digambarkan dengan bagan berikut ini.

Bagan 1 Kategori Reduplikasi



Kategorisasi inilah yang akan digunakan untuk mendeskripsikan hasil analisis dalam artikel ini.

1.5 Fungsi Reduplikasi

Fungsi reduplikasi adalah fungsi semantis yang muncul karena adanya perulangan kata atau morfem. Fungsi tersebut dapat terjadi pada tataran leksikal dan tataran sintaksis. Fungsi semantis pada tataran sintaksis dapat dikenali dengan menghubungkan bentuk reduplikasi dengan kata lain yang menduduki fungsi lain dalam kalimat atau klausa. Tentang hal ini Inkelas (2015) menyatakan sebagai berikut.

Morphological reduplication, both total and partial, is associated with a wide range of syntactic and semantic functions. Reduplication is often semantically iconic, expressing meanings that are impressionistically related to its duplicative nature, like pluralization, emphasis, and frequency/repetition.

Dalam bahasa Indonesia, khususnya verba, telah teridentifikasi sebelas fungsi reduplikasi pada tataran sintaksis tersebut. Fungsi-fungsi tersebut adalah fungsi (1) kejamakan; (2) keanekaan; (3) keserempakan; (4) kesalingan; (5) perulangan; (6) keberlanjutan; (7) intensitas; (8) kesantiaan; (9) ketakterdugaan; (10) kekonsesifan; dan (11) kesertamertaan (Moeliono, 2017). Pada artikel ini pembagian tersebut akan ditinjau kembali karena

disesuaikan dengan data yang ada dalam korpus.

1.6 Jenis

Beberapa ahli membuat kategorisasi bentuk reduplikasi. Samsuri (1988) menyebutkan ada tiga macam reduplikasi, yaitu reduplikasi atau reduplikasi utuh, reduplikasi parsial, dan reduplikasi semu. Keraf (1984) menyebutkan ada empat macam reduplikasi atau pengulangan, yaitu pengulangan dwipurwa, dwilingga, dwilingga salin suara, dan reduplikasi atau ulangan berimbuhan. Selain itu, ia menyebutkan pula istilah reduplikasi semu. Kridalaksana (1985) menyebutkan lima macam reduplikasi, yaitu reduplikasi penuh, berinfiks, dengan variasi fonem, dengan reduplikasi suku pertama dengan atau tanpa pelemahan vokal, dan antisipatoris. Selain itu disebutkan juga lima macam reduplikasi yang lain, yaitu dwipurwa, dwilingga, dwilingga salin suara, dwiwasana, dan trilingga. Parera (1988) menyebutkan reduplikasi (menggunakan istilah bentuk ulang) simetris, regresif, progresif, konsonan, vokal, dan reduplikasi atau bentuk ulang reduplikasi. Ramlan (1979) menyebutkan ada empat macam pengulangan dilihat dari cara mengulang bentuk dasarnya, yaitu pengulangan seluruh, pengulangan sebagian, pengulangan yang berkombinasi dengan proses pembubuhan afiks, dan pengulangan dengan perubahan fonem.

Secara umum reduplikasi dapat dikategorikan berdasarkan bentuk, makna. Dan proses pembentukannya. Dari segi bentuk, reduplikasi dapat menghasilkan beberapa kategori leksikal dan oleh sebab itu disebut dengan reduplikasi leksikal. Adapun reduplikasi yang menghasilkan beberapa kategori makna disebut dengan reduplikasi semantis. Pada jenis reduplikasi ini, bentuk kata dapat saja sama, tetapi jika diletakkan pada kalimat atau konteksnya, makna yang ditimbulkan dapat saja berbeda. Oleh sebab itu, reduplikasi ini juga disebut dengan reduplikasi morfosintaksis.

Dari proses pembentukannya, reduplikasi dapat dibagi menjadi reduplikasi utuh dan reduplikasi sebagian, baik melalui bentuk berimbuhan maupun majemuk. Secara bentuk, reduplikasi dapat menghasilkan beberapa bentuk kata baru, baik itu melalui reduplikasi secara utuh atau keseluruhan, maupun reduplikasi secara sebagian atau parsial.

2. METODE

Dalam penelitian ini digunakan metode bauran (*mixed method*) yang menggunakan pendekatan kuantitatif dan kualitatif. Metode kuantitatif terutama digunakan pada tahap pengumpulan dan pengolahan data, sedangkan metode kualitatif digunakan dalam tahap analisis dan interpretasi data. Kedua metode ini digunakan sedemikian rupa sehingga lebih bersifat kontinum, bukan berupa urutan yang kaku.

2.1 Metode Pengumpulan Data

Data untuk kata ulang utuh dan kata ulang berimbuhan diperoleh dari korpus tata bahasa Indonesia kontemporer (TBIK) di CQPWeb dengan menggunakan pencarian melalui *regular expression* (regex). Simbol regex yang digunakan adalah $+-+$. Simbol itu digunakan untuk menjaring semua kata yang mengandung tanda hubung (-). Bentuk ulang dalam Bahasa Indonesia ditulis dengan menggunakan tanda hubung [1]. Dalam korpus terdapat juga bentuk ulang yang tidak ditulis dengan menggunakan tanda hubung. Namun, pencarian bentuk tersebut dalam CQPWeb tidak mudah. Oleh karena itu, data dibatasi hanya pada bentuk ulang yang penulisannya sesuai dengan kaidah ejaan Bahasa Indonesia.

Hasil yang diperoleh dari pencarian tanda hubung (-) dengan simbol $+-+$ dalam korpus TBIK di CQPWeb adalah 389.077 kemunculan pada 15.130 teks yang berbeda (korpus TBIK mengandung 30.110.481 kata dan 17.472 teks dengan frekuensi kemunculan 12.921.647 per juta kata. Ada 50.693 tipe berbeda dari hasil pencarian tersebut. Akan tetapi, tidak semua tipe merupakan bentuk ulang. Terdapat juga bentuk lain seperti nomor surat, nomor keputusan, bentuk terikat, nama surat dalam Al-Qur'an, dan penanggalan. Untuk itu, dilakukan penyeleksian data. Data yang bukan bentuk ulang dihapus dari daftar. Karena data yang diperlukan adalah bentuk ulang utuh dan bentuk ulang berimbuhan, data berupa bentuk ulang sebagian dan salin suara juga dikecualikan. Jumlah akhir data setelah proses penyeleksian adalah lebih dari 11.300. Jumlah tersebut bukan angka pasti karena masih ada kemungkinan data yang tak terpakai luput dari proses pembersihan.

Pada awalnya data dibatasi pada data dengan frekuensi kemunculan 100 tertinggi. Namun, data yang frekuensinya lebih rendah daripada 100 memperlihatkan keunikan yang menarik untuk dianalisis lebih lanjut. Oleh karena itu, pada akhirnya data yang diolah meliputi seluruh data reduplikasi

yang terseleksi dari hasil pencarian pada korpus TBIK.

Data yang sudah dijarah melalui pencarian dalam korpus kemudian dikategorisasi menurut bentuknya menjadi reduplikasi utuh dan reduplikasi sebagian. Reduplikasi utuh dapat terdiri atas perulangan akar kata dan perulangan kata berimbuhan. Adapun reduplikasi sebagian dapat dibagi menjadi perulangan dengan prefiks, sufiks, infiks, dan konfiks, serta klitik. Selain itu, kata ulang yang muncul dalam ragam cakapan (informal) dan bentuk asing juga dikelompokkan tersendiri. Kategorisasi data tersebut menghasilkan pembagian sebagai berikut.

Kategori	Jumlah Tipe
Bentuk ulang dasar utuh,	3.836
Bentuk ulang berimbuhan utuh	1.284
Bentuk ulang utuh dari kata asing	108
Pola me- + D	563
Pola me- + D + -kan	215
Pola me- + D + -i	42
Pola ke- + D + -an	63
Pola pe- + D	19
Pola pe- + D + -an	24
ter- + D	195
Pola ter- + D + -kan	1
Pola di- + D	265
Pola di- + D + -kan	195
Pola di- + D + -i	28
Pola se- + D	97
Pola se- + D + -nya	231
Pola D + -an	375
Pola D + -i	18
Pola D + -kan	49
D + me-	126
lain-lain	24

Data-data yang sudah dikategorisasi tersebut kemudian dianalisis berdasarkan arah, bentuk, dan jenis bentuk dasarnya. Ada tiga kriteria pemerian reduplikasi yang diterapkan dalam hal ini. Ketiga kriteria tersebut adalah (1) arah reduplikasi, (2) bentuk reduplikasi, dan (3) bentuk dasar.

2.2 Metode Analisis Data

Data yang sudah dikategorisasi pada tahap pengumpulan data kemudian dianalisis secara kualitatif. Metode analisis data yang digunakan adalah metode padan referensial. Metode ini digunakan untuk mengidentifikasi jenis bentuk reduplikasi dan makna yang timbul karena pembentukan tersebut. Dalam metode ini digunakan analisis unsur langsung. Dalam hal ini bentuk reduplikasi dipilah berdasarkan unsur dasar dan unsur perulangannya. Unsur-unsur tersebut kemudian dikategori menurut kesamaannya dan dibuatkan pola pembentukannya. Tahap selanjutnya adalah mengidentifikasi makna yang muncul dari pola-pola pembentukan tersebut.

3. HASIL

Dari hasil analisis data dengan menggunakan metode tersebut di atas, diperoleh hasil analisis menurut arah, bentuk, dan dasar reduplikasi. Selain berdasarkan kategori bentuk, hasil analisis juga meliputi makna yang timbul dalam pembentukan reduplikasi tersebut. Berikut ini struktur penyajian hasil analisis yang dimaksud.

3.1 Kategori Arah Reduplikasi

Kategori arah reduplikasi ditentukan berdasarkan arah pembentukannya. Berdasarkan hal ini, reduplikasi dapat dibagi menjadi reduplikasi progresif dan reduplikasi regresif. Berikut adalah penjelasan keduanya.

3.1.1 Reduplikasi Progresif

Reduplikasi progresif adalah reduplikasi yang arah perulangannya sesuai dengan arah ujaran, yaitu ke kanan. Penentuan dasar reduplikasi didasarkan pada kata yang menjadi inti dari reduplikasi tersebut. Misalnya, kata *menangis-nangis* mempunyai unsur inti *menangis*, sedangkan *nangis* merupakan unsur yang diulang. Karena arah pembentukannya ke kanan, reduplikasi ini digolongkan ke dalam reduplikasi progresif. Begitu pula halnya dengan reduplikasi utuh yang unsur inti dan tambahannya sama. Dalam hal ini, arah pembentukannya menurut arah ujaran, jadi digolongkan ke dalam jenis reduplikasi progresif.

Jika disimpulkan, reduplikasi progresif ini meliputi pola (1) D+D; (2) ber-+D; (3) meng-+D; (4) ter-+D; (5) di-+D; dan (6) se-+D. Berikut ini adalah

contoh untuk pola-pola tersebut.

- a. Serikat Buruh turun ke **jalan-jalan** di luar Bank Sentral Australia.
- b. Untuk menampung **usulan-usulan** yang bersifat perorangan, dibentuklah panitia kecil.
- c. Seorang bocah tampak **bermain-main** menggunakan boks di sebuah layar televisi.
- d. Dalih ini terlalu **mengada-ada**.
- e. Mereka yang awalnya **terkatung-katung** nasibnya, kini mendapat kejelasan.
- f. Seniman jangan **didesak-desak**, bisa mati daya khayalnya.
- g. Pada zaman yang semuanya **seakan-akan** mengusung paham serba global ini...

3.1.2 Reduplikasi Regresif

Reduplikasi regresif adalah reduplikasi yang dibentuk dengan mengulang unsur intinya ke arah belakang atau berlawanan arah dengan arah ujaran. Misalnya, kata *bayar-membayar* mempunyai unsur inti *membayar* dengan unsur tambahan *bayar* yang terletak di sebelah kiri atau berlawanan dengan arah ujaran. Data data korpus yang teranalisis, pola reduplikasi regresif ini meliputi (1) D+meng-; (2) D+meng-/i; (3) D+meng-/kan. Adapun contohnya adalah sebagai berikut.

- a. Jari-jariku cukup jitu dalam **bidik-membidik**.
- b. Meniup berbarengan atau bergantian, **unggul-mengungguli**.
- c. Ya , barangkali dilupakan saja , langsung rekonsiliasi dengan **maaf-memaafkan**.

3.2 Kategori Bentuk Reduplikasi

3.2.1 Reduplikasi Utuh

Reduplikasi utuh diperoleh dengan mengulang bentuk pangkalnya secara utuh. Bentuk pangkal dapat berupa kata dasar maupun kata berimbuhan. Oleh sebab itu, reduplikasi utuh dapat dibagi lagi menjadi reduplikasi utuh dengan pangkal kata dasar (contoh 1) dan 2)) dan dengan pangkal kata berimbuhan (contoh 3) dan 4)) sebagai berikut.

Contoh:

- a. Sekarang mereka mengambil karang dari **pulau-pulau** kecil di depannya.
- b. ... aku mendapat tempat latihan dan **alat-alat gratis**.
- c. Kita juga bisa berlatih melakukan **gerakan-gerakan** menantang nyali itu.
- d. Pada **kesempatan-kesempatan** berikutnya aku kembali mengikuti seleksi dengan tujuan utama.

Bentuk reduplikasi yang mirip dengan perulangan utuh ini adalah reduplikasi fonologis dan reduplikasi semu. Walaupun Zamzani (1993) menyatakan bahwa reduplikasi semu atau reduplikasi fonologis merupakan peristiwa reduplikasi yang terjadi pada suku kata atau sebagian kata, tetapi keduanya dapat dibedakan dari dapat tidaknya dikembalikan kepada bentuk pangkalnya. Misalnya, kata *sumsum*, *waswas*, *ubun-ubun*, *ogoh-ogoh*, dan *ondel-ondel*. Untuk kasus-kasus tersebut perlu dianalisis lebih lanjut apakah tidak ada bentuk yang lebih kecil yang digunakan sebagai sebuah morfem. Jika tidak ada, dapat dipastikan bentuk reduplikasi itu merupakan bentuk reduplikasi fonologis yang tidak menghasilkan leksem baru atau bukan merupakan peristiwa reduplikasi karena secara struktur dan makna tidak dapat dikembalikan bentuk pangkalnya.

Berbeda halnya dengan reduplikasi semu yang masih dapat ditelusuri dan dikembalikan kepada pangkalnya. Di dalam korpus ditemukan beberapa contoh lain yang menunjukkan adanya dua bentuk yang sama-sama lazim digunakan, baik sebagai bentuk dasarnya, maupun sebagai bentuk reduplikasinya. Di dalam korpus kata *paru* muncul sebanyak 560 kali dan *paru-paru* sebanyak 216 kali. Dari segi distribusi, walaupun frekuensi penggunaannya lebih tinggi, kata *paru* hanya muncul dalam 98 teks yang berbeda. Pola distribusi kata ini tidak merata dengan buku teks dan jurnal sebagai dua jenis teks terbanyak yang menggunakan kata ini. Adapun kata *paru-paru* muncul dalam 134 teks yang berbeda dengan distribusi yang merata untuk semua teks. Hal ini menunjukkan bahwa kata *paru* lebih banyak digunakan sebagai istilah teknis di buku teks dan jurnal dibandingkan dengan *paru-paru*. Hal yang hampir sama terjadi pada kata *pori* dan *pori-pori*. Kata *pori* lebih banyak muncul di buku teks dan disertasi/tesis/skripsi (dengan frekuensi 24 dan 11 dari 72), sedangkan *pori-pori* muncul lebih banyak dalam cerpen dan novel (dengan frekuensi 24 dan 10 dari 60).

- a. *Organ lain selain **paru-paru** yang dapat diserang oleh kuman TBC ini adalah kelenjar di leher.*
- b. *Kanker **paru** merupakan penyumbang insiden kanker pada laki-laki tertinggi di Indonesia...*
- c. *Belalang bisa menghambat bakteri penyebab jerawat dan penyumbat **pori-pori**.*
- d. *Masker tanah liat dapat membantu mengeluarkan minyak dan kotoran yang menyumbat **pori**.*

Kata lain yang cukup bersaing adalah *oseng* (8 kali) dan *oseng-oseng* (11 kali). Contoh lain adalah variasi *laki-laki* (5.993 kali) dan *laki* (291 kali) serta *rawa-rawa* (61 kali) dan *rawa* (259 kali).

3.2.2 Reduplikasi Sebagian

Reduplikasi sebagian adalah reduplikasi yang dihasilkan dari mengulang sebagian dari unsur-unsurnya, baik itu pengulangan pada dasar atau pangkalnya. Pengulangan tersebut terutama sekali terjadi pada reduplikasi yang salah satu unsurnya berubah kata turunan. Akan tetapi, selain itu juga pengulangan dapat juga terjadi pada konstruksi majemuk kata.

Ada 18 pola reduplikasi yang dapat diidentifikasi dalam data. Kedelapan belas pola tersebut adalah sebagai berikut.

3.2.2.1 Pola [D + meng-]

Contoh:

- 1) *Kasak-kusuk terjadi, **ancam-mengancam** dilakukan beberapa partai*
- 2) *Suara terompet **susul-menyusul** memekakkan telinga.*

3.2.2.2 Pola [D + -an]

Contoh:

- 1) *Kita harus **pegang-pegangan** lagi kayak upacara api unggun?*
- 2) *Lalu dijamunya makan minum dan **suka-sukaan** malam siang*

3.2.2.3 Pola [D + -nya]

Contoh:

- 1) ***Tiga-tiganya** langsung dihukum berdiri di depan kelas.*
- 2) *Kenapa kamu **bisa-bisanya** pergi tanpa (meng)ajak aku?*

3.2.2.4 Pola [ber- + D]

Contoh:

- 1) *Mungkin mereka tak berani kelihatan **bersantai-santai** saja di depan Pelatih Toharun.*
- 2) *Pantai curam biasanya **bergunung-gunung**.*

3.2.2.5 Pola [ber- + D + -an]

Contoh:

- 1) *Marilah kita **bersuka-sukaan** makan minum karena perburuan terlalu amat banyak itu.*
- 2) *Airnya kuning keemasan, **berkilau-kilauan** tertimpa cahaya.*

3.2.2.6 Pola [meng- + D]

Contoh:

- 1) *Ada yang **mencubit-cubit** pipiku.*
- 2) *Genduk keluar dari kamar dengan lincah sambil **menyanyi-nyanyi** kecil.*

3.2.2.7 Pola [meng- + D + -kan]

Contoh:

- 1) *Teddy tersenyum sambil **menggeleng-gelengkan** kepala.*
- 2) *Nurmi berupaya keras **menguat-nguatkan** dirinya.*

3.2.2.8 Pola [meng- + D + -i]

Contoh:

- 1) *Aku akan menggunakannya untuk **menakut-nakuti** seseorang.*
- 2) *Mereka **mengiming-imingi** hak politik rakyat dengan uang.*

3.2.2.9 Pola [ke- + D + -an]

Contoh:

- 1) *... kalimat-kalimat yang saya susun penuh **kehati-hatian**.*
- 2) *Barangnya itu dulang tembaga busuk **kehijau-hijauan** peninggalan neneknya.*

3.2.2.10 Pola [pe- + D]

Contoh:

- 1) *Memberi sumbangan sesuai kemampuan kepada **peminta-minta** yang datang.*
- 2) *Artinya **perlahan-lahan** kita akan kehilangan.*

3.2.2.11 Pola [peng- + D + -an]

Contoh:

- 1) ... menyelamatkan duit rakyat dari praktek **penghambur-hamburan** anggaran.
- 2) Tidak ada **pembeda-bedaan** di antara mereka.

3.2.2.12 Pola [ter- + D]

Contoh:

- 1) Kami tak mau **terburu-buru** menyimpulkan.
- 2) pengajian yang hendak dimulai menjadi **tertunda-tunda** beberapa lama.

3.2.2.13 Pola [ter- + D + -kan]

Contoh:

Dua tahun yang **tersia-siakan**.

3.2.2.14 Pola [se-- + D]

Contoh:

- 1) ... kepentingan dalam perilaku **sehari-hari** mereka.
- 2) Padahal, **sewaktu-waktu** ia bisa rusak juga.

3.2.2.15 Pola [se-- + D]

Contoh:

- 1) kepentingan dalam perilaku **sehari-hari** mereka.
- 2) Padahal, **sewaktu-waktu** ia bisa rusak juga.

3.2.2.16 Pola [se-- + D + -nya]

Contoh:

- 1) Setiap stadion harus dilengkapi **sekurang-kurangnya** dua ruang ganti.
- 2) Aku melolong **sejadi-jadinya** karena rasa sesal itu begitu tak tertahankan.

3.2.2.17 Pola [D + -em-/ -el-/ -er-]

Contoh:

- 1) ... mengenakan satu atau dua cincin di **jari-jemari** mereka.
- 2) Tubuhnya diikat **tali-temali**.

3.2.2.18 Pola Bentuk Majemuk

Majemuk kata yang ditulis serangkai diulang seluruhnya, sedangkan majemuk kata yg terdiri atas dua kata atau lebih, hanya diulang unsur pertamanya.

Contoh:

- 1) ***Kacamata-kacamata** yang dibuat Kerin selalu ada sentuhan handmade-nya.*
- 2) *Ia menjadi pengunjung tetap **rumah-rumah sakit** Roma.*

3.3 Kategori Jenis Dasar Reduplikasi

Kategori reduplikasi berdasarkan jenis dasarnya dibagi menjadi reduplikasi dengan dasar kata dasar dan reduplikasi dengan dasar kata turunan. Berikut ini penjelasan beserta contoh untuk setiap jenis reduplikasi tersebut.

3.3.1 Reduplikasi Kata Dasar

Reduplikasi dengan dasar berupa kata dasar dapat dibagi menjadi beberapa kategori berdasar kelas kata seperti berikut ini.

3.3.1.1 Nomina

Contoh:

- 1) *Di **bangku-bangku** yang berbaris antara tiang lampu itu, selalu terdengar sayup percakapan.*
- 2) *Di **waktu-waktu** tertentu, si gurita mengibaskan tentakelnya ke jendela.*

3.3.1.2 Verba

Contoh:

- 1) *Jangan ada yang **main-main** lagi.*
- 2) *Kafanya bisa-bisa sudah tutup kalau kita tidak **pergi-pergi** juga.*

3.3.1.3 Adjektiva

Contoh:

- 1) *... memakai bau-bauan yang **harum-harum** dan mengambil segala bunga.*
- 2) *Telapak kakinya perih, **lecet-lecet** akibat lari tadi.*

3.3.1.4 Adverbia

Contoh:

- 1) *Jarang-jarang* aku mendapat kesempatan ini.
- 2) Ia memberi pelatihan secara *cuma-cuma*.

3.3.1.5 Pronomina

Contoh:

- 1) Sudah saya sampaikan ke *beliau-beliau*, secepat-cepatnya.
- 2) Katanya *kapankapankapankapan* kau boleh datang memancing bersamanya.

3.3.1.6 Numeralia

Contoh:

- 1) Aku melompati anak-tangga *dua-dua*.
- 2) *Satu-satunya* yang dibutuhkan sejumlah besar pikiran ini adalah pengarahan.

3.3.2 Reduplikasi Kata Berimbuhan/Turunan

Reduplikasi ini dibentuk dari pengulangan kata berimbuhan, baik sebagian maupun seluruhnya. Pengulangan seluruhnya masuk dalam kategori reduplikasi utuh seperti contoh berikut ini.

Contoh:

- 1) Ada *lapisan-lapisan* pada tanah yang tersusun sejak zaman purba.
- 2) Orang yang berpedoman pada prinsip ekonomi selalu melakukan *pertimbangan-pertimbangan*.

Reduplikasi dengan dasar berupa kata berimbuhan secara parsial dibentuk seperti halnya reduplikasi parsial yang sudah disebutkan pada bagian sebelumnya. Sebagai dasar reduplikasi, kata berimbuhan dapat berada di sebelah kiri atau sebelah kanan unsur tambahannya yg berupa kata dasar menurut ola pembentukan reduplikasi parsial yang sudah dibahas sebelumnya.

3.4 Analisis Makna Reduplikasi

Melalui analisis data korpus, setidaknya terdapat enam makna yang dapat diidentifikasi dari bentuk reduplikasi. Selain keenam kategori tersebut, ada satu kategori makna yang meliputi variasi makna yang belum dapat ditentukan dengan jelas kategorinya, sehingga dikumpulkan dalam satu kategori tertentu. Kategori ketujuh ini sebagian besar merupakan kategori

makna secara pragmatis yang harus diidentifikasi dari tujuan atau fungsi kalimat atau ujaran itu digunakan. Adapun keenam kategori makna yang dapat diidentifikasi tersebut adalah sebagai berikut.

3.4.1 Kejamakan/Keanekaan

Makna kata ulang yang menunjukkan kejamakan timbul dari reduplikasi kategori nomina, adjektiva, dan adverbial. Bentuk reduplikasinya dapat berubah bentuk dasar maupun bentuk berimbuhan.

Contoh:

- 1) *Hujan lebat diikuti angin kencang yang sering menumbangkan **pohon-pohon*** ('aneka/banyak pohon')
- 2) *Badanku malah sudah mulai **bengkak-bengkak***. ('bengkak di beberapa tempat')
- 3) *Seperti yang **sudah-sudah**, riuh rendah komentar bermunculan* ('yang sudah terjadi/berlalu')
- 4) *Harapannya, **bangunan-bangunan** di Yogyakarta akan tetap kokoh seandainya nanti gempa datang lagi.*
- 5) *Aku masih temukan **kesenangan-kesenangan** di dunia, masih melihat banyak keindahan di sini.*

3.4.2 Intensitas

Makna kata ulang yang menunjukkan kejamakan timbul dari reduplikasi kategori verba, adjektiva, dan adverbial.

- 1) *Luka yang tidak **sembuh-sembuh** akibat infeksi dapat menunjukkan adanya DM* ('tidak kunjung sembuh')
- 2) *Kenapa ia bermain **subuh-subuh** di pemakaman umum* ('subuh sekali')
- 3) *Waktu itu aku punya bayi, rumah terasa sempit dan tentunya sulit **pergi-pergi***. ('pergi berkali-kali')
- 4) *Mula-mula Fitri mengeluh sakit perut, lalu **muntah-muntah***. ('muntah berkali-kali')
- 5) *Dindingnya hanya anyaman daun kelapa yang sudah tua dan **jarang-jarang***. ('sangat jarang')

Pada kelas kata verba, selain dengan reduplikasi utuh, makna intensitas ini juga dapat ditunjukkan melalui bentuk reduplikasi berimbuhan sebagian seperti contoh berikut ini.

- 6) ... aku berteriak ketakutan, **memanggil-manggil** ibu.
- 7) Para penonton **tertawa-tawa** ketika penghibing itu menari dengan gerakan-gerakan lucu.
- 8) ... tubuh kerempeng Tanjil dihela, **dibanting-banting**, kemudian diempaskan ke atas badan jalan.
- 9) Masyarakat pun **berbondong-bondong** mengurus e-KTP ke kantor-kantor kelurahan dan kecamatan.

3.4.3 Kualifikasi

Makna kata ulang dapat dilihat dari penggunaan kata tersebut di dalam konteks kalimat. Oleh sebab itu, kata ulang jenis ini disebut juga kata ulang sintaksis. Reduplikasi sintaksis merupakan reduplikasi gramatikal yang masukannya berupa leksem (ada yang menyebut morfem), dan keluarannya berupa klausa.

Contoh:

Tua-tua, masih mampu naik sepeda orang itu.

Bentuk *tua-tua* dalam konteks itu dapat diparafrasakan menjadi *meskipun sudah tua* atau *walaupun sudah tua*. Bentuk kata ulang yang sama dapat mempunyai makna yang berbeda jika diletakkan ke dalam konteks kalimat yang berbeda.

Contoh:

- 1) Mereka minta **bapak-bapak** kami menaruh tanda jempol di surat. ('jamak')
- 2) Nah waktu asyik merokok, tiba-tiba datang seorang **bapak-bapak**. ('laki-laki yang kualifikasinya seperti seorang bapak')

Pada klausa 1) bentuk *tua-tua* mempunyai acuannya orang tua laki-laki dalam bentuk jamak, sedangkan pada klausa 2) maknanya mengacu kepada laki-laki yang mempunyai kualifikasi seorang bapak (agak tua) dalam bentuk tunggal. Perbedaan ini sangat bergantung pada konteks penggunaannya dalam kalimat dan kategorisasi maknanya tidak dapat hanya bersifat leksikal.

Makna ‘berkualifikasi seperti’ dalam reduplikasi hanya muncul pada kelas kata nomina. Dalam identifikasinya, makna ‘berkualifikasi seperti’ harus secara jeli dibedakan dalam makna ‘jamak’ karena sangat ditentukan oleh konteks dalam kalimat.

- 1) *Setahuku dia memang bukan tipe **nenek-nenek** tua yang lemah.*
(‘wanita yang tingkah lakunya seperti seorang nenek’)
- 2) *cowok yang (me)ngaku coverboy itu malah seorang **om-om** gendut.*
(‘laki-laki setengah tua’)

Bandingkan dengan contoh 3) dan 4) berikut ini yang bermakna jamak.

- 3) *Banyak **nenek-nenek** dan kakek-kakek yang masih hidup karena terus merokok.*
- 4) *... namun para **om-om** pun tak mau ketinggalan.*

Termasuk dalam kategori ini adalah simbolisme bunyi. Reduplikasi sering digunakan untuk menyimbolkan bunyi (Kallergi, 2015). Simbolisme bunyi (*sound symbolism*) ini menghubungkan antara bunyi yang mencerminkan keadaan atau kondisi dan makna yang timbul.

Di dalam korpus, reduplikasi berupa simbol bunyi ini juga muncul dengan frekuensi yang rendah (1 kali) untuk setiap kata.

Contoh:

- 1) *Paling jadinya ya **jeng-jeng** close-up melulu kayak sinetronmu itu.*
- 2) *... mereka juga ikut-ikut bertanya kanan-kiri apa benar **kaok-kaok** burung gagak itu.*
- 3) *Potongan hijau **kres-kres** berenang di tengah genangan gulai santan yang menguning*
- 4) *... kentang gorengnya yang **kriuk-kriuk** atau potongan kol yang segar.*
- 5) *Bunga merah yang meregang itu pun terbuka dengan sempurna dan mengeluarkan bunyi **ting-ting** berkas-berkas cahaya.*
- 6) *Suara **tong-tong** bertalu-talu di pagi buta , membuat warga segera berduyun-duyun naik ke bukit.*

Jika diperhatikan, contoh reduplikasi tersebut semuanya menyerupai simbol bunyi dari benda atau perbuatan yang dilakukan. Oleh sebab itu, reduplikasi yang mengandung makna simbolisme bunyi ini dapat terjadi pada kelas kata nomina, verba, dan adjektiva.

Pada contoh 7) di bawah ini, walaupun bentuk reduplikasinya adalah sama, yaitu *tong-tong*, tetapi maknanya berbeda. Contoh 6) bermakna tiruan bunyi benda sepereti kaleng yang dipukul-pukul, sedangkan contoh 7) bermakna beberapa tong atau kaleng besar.

- 7) *Kedua teman Deren sempat menyimpan **tong-tong** plastik berwarna hitam di dalam mobil itu.*

3.4.4 Kesalingan (Keresiprokalan)

Makna kesalingan atau berbalas-balasan pada reduplikasi ini sangat khas terjadi pada kelas kata verba, khususnya verba taktransitif.. Bentuk reduplikasi yang paling lazim adalah reduplikasi sebagian yang berupa kombinasi bentuk berimbuhan dan bentuk berimbuhan atau dalam ragam informal dengan pola bentuk dasar dengan bentuk berimbuhan. Adapun contohnya adalah sebagai berikut.

- 1) *Bagaimana rekonsiliasi kalau tak dibarengi dengan **bermaaf-maafan**? ('saling memaafkan')*
- 2) *Habis Lebaran, kita **maaf-maafan**.*
- 3) *Suaranya **bersahut-sahutan** dengan Nat King Cole dan semakin lama semakin tak keruan. ('saling menyahut')*
- 4) *Sahri melaju dengan perasaan yang bergurindam **sahut-sahutan**.*
- 5) *Proses pengecoran bendung dan beton **berkejar-kejaran** dengan naiknya muka air Danau Menjer ('saling mengejar')*
- 6) *Maka, terjadilah **kejar-kejaran** mobil, antara mobil Heldy dengan mobil Bung Karno.*

3.4.5 Kemiripan

Untuk kelas kata nomina yang sangat terbatas, terdapat bentuk reduplikasi yang menyatakan makna kemiripan dengan makna pangkalnya. Kata-kata ini tidak sama dengan reduplikasi perulangan semu, karena jika digunakan tidak dalam bentuk perulangannya, akan menyebabkan makna yang sangat berbeda. Contoh:

- 1) ***Kuda-kuda** atap pabrik menempa dua kamar kontrakan yang berada tepat di belakang pabrik.*
- 2) ****Kuda** atap pabrik menempa dua kamar kontrakan yang berada tepat di belakang pabrik.*

- 3) *Cara membuat kubah dibuat pola dengan bambu sebagai **jari-jari** lingkaran.*
- 4) **Cara membuat kubah dibuat pola dengan bambu sebagai **jari** lingkaran.*
- 5) *... membersihkan ruangan secara khusus dan mengganti **langit-langit** dan lantai ruangan.*
- 6) **... membersihkan ruangan secara khusus dan mengganti **langit** dan lantai ruangan.*

3.4.6 Perihal

Makna ‘perihal’ atau ‘seluk-beluk tentang’ ini dapat dihasilkan melalui perulangan sebagian dengan pangkal verba. Reduplikasi ini dapat membentuk leksem baru yang berkelas kata nomina. Adapun contohnya adalah sebagai berikut.

- 1) *Selain menggambar, passion Naela Ali tampaknya juga di dunia **tulis-menulis**.*
- 2) *Pagi harinya ia dibangunkan dan diberitahu tentang **tembak-menembak** dan kerusakan-kerusakan yang telah terjadi malam itu.*
- 3) *Pada perang dunia ke II rutinitas sehari-hari di kebun karet dan kegiatan **jual-menjual** jadi terhenti.*

3.4.7 Lain-lain

Selain makna yang dapat diidentifikasi dan dikategorisasi dengan jelas seperti tersebut sebelumnya, makna reduplikasi lain juga muncul dengan sangat bervariasi. Reduplikasi jenis ini terutama muncul pada kata-kata seperti berikut ini.

Contoh:

- 1) *Nurbaya mengeratkan pegangannya, khawatir kalau lelaki yang harus dijaganya itu **kenapa-kenapa**. (‘khawatir’)*
- 2) *Gadis salihah itu mau menerimaku yang **begini-begini** saja. (‘hanya begini’).*
- 3) *Katanya yang **begini-begini** ini memang merupakan bagian yang sah dari kehidupan. (‘yang seperti ini’).*
- 4) ***Jangan-jangan** motif rente ekonomi inilah yang mendorong kebijakan impor. (‘barangkali, mungkin’)*
- 5) *“Kami cuma nelayan kecil. **Syukur-syukur** kalau bisa menutupi ongkos melaut,” ucap Darwis. (‘sudah sangat beruntung’)*

- 6) *Tahu-tahu* wajah Camilla muncul juga dari pintu kamarnya.
(‘tiba-tiba’)
- 7) *Siapa mereka? Lagi-lagi* Syamsu enggan menyebut. (‘sekali lagi;
kembali lagi’)

Di samping makna yang sangat beragam tersebut, ada pula bentuk reduplikasi yang secara spesifik tidak mengandung makna leksikal, tetapi secara pragmatis mempunyai fungsi untuk menjadi pengantar untuk memulai percakapan. Fungsi seperti ini tergolong ke dalam fungsi fatis (*phatic communion*) yang konsepnya dicetuskan oleh Malinowski. Dalam data korpus, ungkapan basa-basi semacam ini ditemukan seperti contoh berikut.

- 8) *Dengar-dengar* dia tak cocok dengan Tante Bahar.
- 9) *Omong-omong*, benar juga ucapanmu dulu soal kematian.

4. SIMPULAN

Reduplikasi atau perulangan kata adalah gejala yang sangat lazim dalam bahasa Indonesia. Secara umum, reduplikasi dapat didefinisikan sebagai proses mengulang kata, dapat berupa kata dasar, kata berimbuhan, maupun majemuk. Reduplikasi berfungsi membentuk kategori kata dan makna yang baru dan dapat terjadi pada empat kelas kata utama dalam bahasa Indonesia, yaitu verba, nomina, adjektiva, dan adverbial.

Dari pendekatan yang baru terhadap kategorisasi reduplikasi dalam bahasa Indonesia berdasarkan korpus penggunaan bahasa kontemporer dapat diperoleh kesimpulan sebagai berikut.

1. Bentuk reduplikasi dalam bahasa Indonesia kontemporer tidak banyak berubah secara signifikan.
2. Reduplikasi yang merupakan wujud simbolisme bunyi belum menjadi hal yang diperbincangkan dengan terperinci dalam tata bahasa terdahulu, padahal pembentukan kata ini sangat potensial terjadi. Oleh sebab itu, perlu dimasukkan ke dalam pembahasan khusus.
3. Penggunaan kata ulang semu yang mempunyai variasi, terutama yang digunakan dalam tulisan ilmiah (jurnal, buku teks, disertasi/tesis/skripsi) dapat menjadi dasar pembakuan bentuk reduplikasi.

Referensi

- [1] Alisyahbana, Sutan Takdir. 1953. *Tata Bahasa Baru Bahasa Indonesia II*. Jakarta: Pustaka Rakyat.
- [2] Amalia, Dora (ed.). 2016. *Kamus Besar Bahasa Indonesia Edisi V*. Jakarta: Badan Pengembangan dan Pembinaan Bahasa
- [3] Anggraini, Rizky. 2019. “Analisis Kata Ulang (Reduplikasi) dan Makna dalam Cerpen *Maryam* Karya Afrion”. Skripsi FKIP. Medan: Universitas Muhammadiyah Sumatra Utara.
- [4] Falah, Aimah Nurul. 2016. “Reduplikasi Verba Bahasa Indonesia dan Bahasa Jawa”. Skripsi Fakultas Bahasa dan Seni. Semarang: Universitas Negeri Semarang.
- [5] Fatimah, Nurul. 2017. “Reduplikasi di Harian Kompas dan Implikasinya dengan Pembelajaran Bahasa Indonesia di SMP”. Skripsi pada Program Studi Pendidikan Bahasa Indonesia. Surakarta: FKIP Universitas Muhammadiyah Surakarta.
- [6] Firman, A.D. 2014. “Bentuk dan Makna Reduplikasi dalam Bahasa Moronene”. Artikel dalam *Jurnal Kandai*, Vol. 10, No. 1. Kendari: Kantor Bahasa Sulawesi Tenggara.
- [7] Fitrianti. 2020. “Analisis Bentuk dan Makna Nomina Reduplikasi dalam Novel Kontemporer Indonesia: Tinjauan Morfologi”. Skripsi Departemen Sastra Indonesia, Fakultas Ilmu Budaya, Universitas Hasanuddin. Makassar
- [8] Gil, David. 2011. “From repetition to reduplication in Riau Indonesian” dalam *Studies on Reduplication*. Berlin/Boston: De Gruyter Mouton.
- [9] Inkelas, S. dan Laura J. Downing. 2015. “What is Reduplication?” dalam *Language and Linguistics Compass*. Hlm. 502–515.
- [10] Kallergi, Haritini. 2015. *Reduplication at The Word Level: The Greek Facts of Typological Perspectives*. Berlin/Boston: De Gruyter Mouton.
- [11] Keraf, Gorys. 1984. *Tata Bahasa Indonesia*. Ende-Flores: Nusa Indah.
- [12] Keraf, Gorys. 1991. *Tata Bahasa Rujukan Bahasa Indonesia*. Jakarta: Grasindo.
- [13] Kridalaksana, Harimurti. 1982. *Kamus Linguistik*. Jakarta: PT Gramedia.
- [14] Kridalaksana, Harimurti. 1985. *Tata Bahasa Deskriptif Bahasa Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Depdikbud.
- [15] Kridalaksana, Harimurti. 1989. *Pembentukan Kata dalam Bahasa Indonesia*. Jakarta: PT Gramedia.

- [16] Mattes, Veronica. 2014. *Types of Reduplication: A Case Study of Bikol*. Berlin/Boston: De Gruyter Mouton.
- [17] Matthews, P.H. 1978. *Morphology*. Cambridge: Cambridge University Press.
- [18] Moeliono, Anton M. dkk. 2017. *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Badan Pengembangan dan Pembinaan Bahasa.
- [19] Nita, Lensaf dkk. 2019. “Reduplikasi pada Harian Rakyat Bengkulu”. Artikel dalam *Jurnal Ilmiah Korpus*. Bengkulu: Universitas Bengkulu.
- [20] Ong, Shyi Nian dkk. 2020. “Bentuk, Fungsi, dan Distribusi Reduplikasi Kata Benda Bahasa Melayu dan Bahasa Jepang”. Artikel dalam *Journal of Japanese Language Education & Linguistics*. Volume 4, No. 1
- [21] Parera, Jos Daniel. 1988. *Morfologi*. Jakarta: PT Gramedia.
- [22] Prasetiawan, Deny. 2014. *Mabasan*, Vol. 8 No.2, Juli—Desember 2014 : 100—111
- [23] Ramlan, M. 1979. *Morfologi*. Yogyakarta: UP Karyono.
- [24] Rubino, Carl R. 2005. “Reduplication” dalam Martin Haspelmath et al. (ed.) *The World Atlas of Language Structure*. Oxford: Oxford University Press.
- [25] Samsuri. 1980. *Analisis Bahasa*. Jakarta: Erlangga.
- [26] Sugono, Dendy dkk. 2007. *Pedoman Umum Pembentukan Istilah*. Jakarta: Pusat Bahasa.
- [27] Zamzani. 1993. “Pemerian Wujud Reduplikasi Bahasa Indonesia”. Artikel dalam *Jurnal DIKSI* No.2 Th./ Mei 1993

PEMAJEMUKAN DALAM BAHASA INDONESIA: PENDEKATAN DISTRIBUSIONAL

Adi Budiwiyanto

Badan Pengembangan dan Pembinaan Bahasa

adibudiwiyanto@gmail.com

Abstrak

Meskipun para linguis telah menetapkan ciri kata majemuk, penentuan gabungan kata sebagai majemuk masih terasa intuitif. Perlu suatu pendekatan baru untuk melihat pemajemukan secara empiris sesuai penggunaan alami dalam masyarakat bahasa. Pendekatan empiris tersebut dapat dilakukan dengan linguistik korpus. Penelitian ini mengkaji pemajemukan dalam bahasa Indonesia dengan pendekatan yang berbeda, yaitu pendekatan berbasis korpus dengan berbasis pada data kebahasaan digital kontemporer yang terdiri atas ± 25 juta kata (token) dari 12 ragam tulis bahasa Indonesia. Penelitian ini bertujuan untuk mengidentifikasi majemuk kata/frasa, mendeskripsikan jenis, struktur, makna dan fungsi majemuk dengan menggunakan pendekatan distribusional. Berdasarkan hasil pembahasan di atas, pendekatan distribusional yang berbasis frekuensi kemunculan dapat digunakan sebagai salah satu cara dalam menentukan gabungan kata yang majemuk. Majemuk yang dihasilkan dapat berisifat kata dan bersifat frasa. Majemuk yang bersifat kata disebut majemuk kata, sedangkan majemuk yang bersifat frasa disebut majemuk frasa. Majemuk tersebut dapat terbentuk dari kata dasar bebas, kata berimbuhan, atau kombinasi keduanya. Berdasarkan kategori unsurnya, majemuk dapat dikelompokkan menjadi majemuk verbal, majemuk nominal, majemuk adjektival, majemuk adverbial, majemuk preposisional, dan majemuk numeral. Dari analisis didapatkan bahwa ada majemuk yang memiliki makna idiomatis, baik berupa idiom penuh maupun idiom sebagian. Selain itu, ada juga majemuk yang memiliki makna non-idiomatis. Dalam hal fungsi, khususnya yang berkaitan dengan bagaimana bahasa bekerja untuk mencapai tujuan komunikatif tertentu, majemuk memiliki fungsi referensial (*referential*), fungsi tekstual (*textual*), dan fungsi pendirian (*stance*).

Kata Kunci: *majemuk, korpus, pendekatan distribusional, struktur, makna, fungsi*

1. PENDAHULUAN

Ide menggunakan frekuensi dalam menentukan kata majemuk dicetuskan oleh Parera [1]. Ia menawarkan perhitungan frekuensi pasangan dalam mengidentifikasi kata majemuk dalam sebuah naskah atau tutur sebuah bahasa. Perhitungan pasangan itu berguna untuk penentuan sebuah pasangan, apakah bentuk majemuk atau tidak. Menurut Parera [1], ada dua kemungkinan frekuensi pasangan itu. Pertama ialah pasangan bebas, yaitu bentuk kata yang secara potensial dapat berpasangan dengan kata atau bentuk mana saja. Itu berarti bahwa bentuk dan kata itu bebas dan tidak terikat untuk berpasangan. Dalam contoh yang ia berikan, tampak bahwa kata *bentuk* dan *cara* dapat berpasangan dengan kurang lebih 35 kata yang berbeda. Kedua ialah pasangan terikat, yang berarti kata itu dengan sangat terbatas atau dapat berpasangan dengan kata yang tertentu saja. Dari hasil perhitungan frekuensi, pasangan bahasa seperti *kamar mandi*, *tata buku*, dan *tanda tangan* tergolong sebagai pasangan bebas karena kata *kamar*, *tata*, dan *tanda* secara potensial berfrekuensi tinggi untuk berpasangan dengan kata-kata yang lain. Misalnya, kata *kamar* dapat berpasangan dengan *-mandi*, *-tidur*, *-tunggu*, *-operasi*, *-rapat*, *-gelap*, *-kecil*, *-rias*, *-tamu*, *obat*, *-dokter*; kata *tanda* dapat berpasangan dengan *-tangan*, *-mata*, *-kasih*, *-lalulintas*, *-bahaya*, *-bunyi*, *-baca*, *-laku*; kata *tata* dapat berpasangan dengan *-buku*, *-bahasa*, *-warna*, *-kuliah*, *-rias*, *-cara*, *-negara*, *-bunyi*, *-irama*, *-kalimat*, *-krama*, *-bentuk*, *-laksana*. Berdasarkan pengamatan itu, Parera [1] menyimpulkan bahwa kata-kata yang secara potensial dapat berpasangan dengan sebanyak mungkin kata tidak termasuk dan tidak dapat dikatakan sebagai bentuk majemuk. Kata-kata itu berpasangan secara bebas dan terbuka. Dengan demikian, sistem tata urutan dalam konstruksi morfologis dan sitaksis bahasa Indonesia merupakan sistem yang terbuka dan bebas. Pasangan terikat adalah bentuk majemuk. Pasangan terikat adalah dua bentuk bahasa yang secara khusus, terbatas, dan tetap berpasangan dalam keseluruhan pelaksanaan bahasa. Pasangan terikat tidak mempunyai kemungkinan untuk berpasangan dengan bentuk atau kata lain. Dengan kata lain, kata majemuk bersifat tertutup.

Cara Parera [1] dalam menentukan kata majemuk tersebut menarik saya untuk menggunakan pendekatan yang berbeda dalam mengkaji kata majemuk. Perlu suatu pendekatan baru untuk melihat kata majemuk secara empiris sesuai penggunaan alami dalam masyarakat bahasa. Pendekatan empiris tersebut dapat dilakukan dengan linguistik korpus. Hal itu juga didukung oleh

kemajuan teknologi yang memudahkan para peneliti untuk mengumpulkan data elektronik yang dapat dianalisis menggunakan penganalisis korpus.

Penelitian ini mengkaji pemajemukan dalam bahasa Indonesia dengan pendekatan yang berbeda, yaitu pendekatan berbasis korpus dengan berbasis pada data kebahasaan digital kontemporer yang terdiri atas \pm 25 juta kata (token) dari 12 ragam tulis bahasa Indonesia. Penelitian ini bertujuan untuk mengidentifikasi majemuk, mendeskripsikan jenis, struktur, makna dan fungsi majemuk dengan pendekatan berbasis korpus, khususnya pendekatan distribusional.

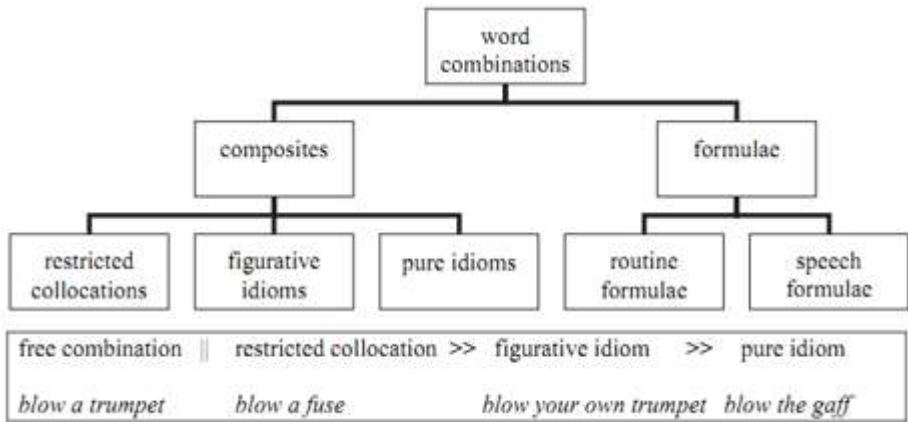
Pada awal 1990-an pada umumnya penelitian yang berkenaan dengan pola fraseologi telah bersifat empiris dengan memanfaatkan analisis korpus. Ada dua pendekatan umum yang telah digunakan untuk mengidentifikasi dan menganalisis satuan multikata, yaitu pendekatan berbasis korpus dan pendekatan tergerakkan korpus. Dalam kajian satuan multikata berbasis korpus, peneliti terlebih dahulu memilih ungkapan multikata yang dianggap penting atau secara teoretis menarik untuk dikaji, kemudian menganalisis korpus untuk mengetahui bagaimana gabungan kata itu digunakan. Dengan kata lain, validitas struktur kebahasaan diasumsikan berasal dari teori linguistik. Tujuan utama penelitian adalah menemukan pola penggunaan yang sistematis untuk fitur kebahasaan yang telah ditentukan sebelumnya. Sebaliknya, dalam kajian yang tergerakkan korpus, korpuslah yang dianalisis secara induktif untuk mengidentifikasi rangkaian multikata yang penting atau patut dicatat. Analisis korpus tersebut meliputi pencatatan semua kombinasi kata dengan karakteristik khas yang muncul di dalam korpus dengan memperhatikan frekuensi kemunculan. Jadi, pada penelitian tergerakkan korpus, konstruksi kebahasaan muncul sebagai hasil analisis korpus.

2. TINJAUAN PUSTAKA

2.1 Ancangan Distribusional

Ada dua pendekatan utama dalam fraseologi, yaitu pendekatan tradisional dan pendekatan distribusional) [2]. Pendekatan tradisional, yang dipelopori oleh linguis Rusia dan negara Eropa bagian timur, seperti Vinogardov dan Amosova, melihat fraseologi sebagai satuan multikata berupa kontinum, satusisi bersifat sangat kabur dan tetap dan di sisi lain bersifat sangat jelas dan

bervariasi. Pendekatan ini terlihat dalam kontinum yang dibuat oleh Cowie [3] [4] yang menarik garis lurus dari gabungan bebas ke idiom melalui kolokasi terbatas dan idiom figuratif. Dalam tradisi ini, satuan paling idiomatis, yang maknanya tidak dapat ditentukan dari unsur-unsur yang menyusun gabungan kata itu, disebut satuan paling inti dan dianggap sebagai prototipe satuan fraseologis. Oleh karena itu, pendekatan ini disebut juga pendekatan fraseologis.



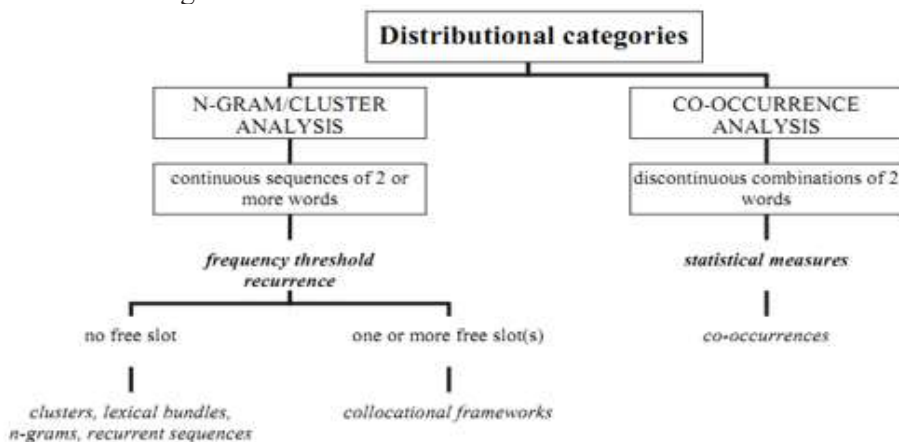
Gambar 1 Klasifikasi dan kontinum fraseologis menurut Cowie [3]

Dalam pembagian gabungan kata, Cowie [3] memisahkan secara jelas antara majemuk (*composite*) yang secara sintaktis berfungsi di bawah tataran kalimat dan formula (*formulae*) yang secara pragmatis sebagai ujaran yang otonom. Kategori formula yang mencakup satuan lirkalimat (*sentence-like*), seperti pepatah, slogan, dan formula percakapan, dibagi lagi menjadi dua, yaitu formula rutin, misalnya *good morning* dan *see you soon*, dan formula ujaran, misalnya *you know what I mean* atau *are you with me?* Tipologi gabungan kata ini merupakan salah satu yang berpengaruh dalam bidang leksikologi dan leksikografi.

Pendekatan yang kedua terhadap fraseologi telah mengubah arah fraseologi. Pendekatan yang dipelopori oleh John Sinclair [5] ini tidak menggunakan pendekatan atas-bawah untuk mengidentifikasi satuan fraseologis dengan mendasarkan pada kriteria linguistik, tetapi pendekatan tergerakkan korpus bawah-atas untuk mengidentifikasi kookuransi (*coocurance*) kata. Pendekatan induktif ini disebut pendekatan distribusional

atau pendekatan berbasis frekuensi. Pendekatan ini menghasilkan beragam gabungan kata yang tidak masuk dalam kategori linguistik yang sebelumnya telah ditentukan dan telah menuntun pada wilayah kemungkinan sintagmatis yang luas, seperti kerangka (*frame*), kerangka kolokasional, koligasi, dan frasa berulang komposisional. Semua jenis rangkaian kata itu menggambarkan prinsip idiom (*idiom principle*) Sinclair [5], yaitu prinsip yang melihat bahasa pada dasarnya tersusun atas untaian kata-kata yang terseleksi bersama yang membentuk pilihan tunggal dan tata bahasa leksikal. Prinsip idiom mengacu pada tendensi bahasa yang bersifat fraseologis, yaitu katakata tidak muncul sendirian, tetapi saling bergabung untuk membentuk makna [6]. Hal tersebut berlawanan dengan prinsip *open-choice* yang mengasumsikan bahwa kata-kata secara individual dipilih untuk mengisi slot atau gatra tertentu dalam sebuah kalimat. Sementara itu, tata bahasa leksikal merupakan sebuah upaya untuk membangun tata bahasa dan kosakata (leksis) pada basis yang setara—makna dan struktur dianggap satu. Untuk menggemakan hal itu, dikembangkan tata bahasa pola (*pattern grammar*).

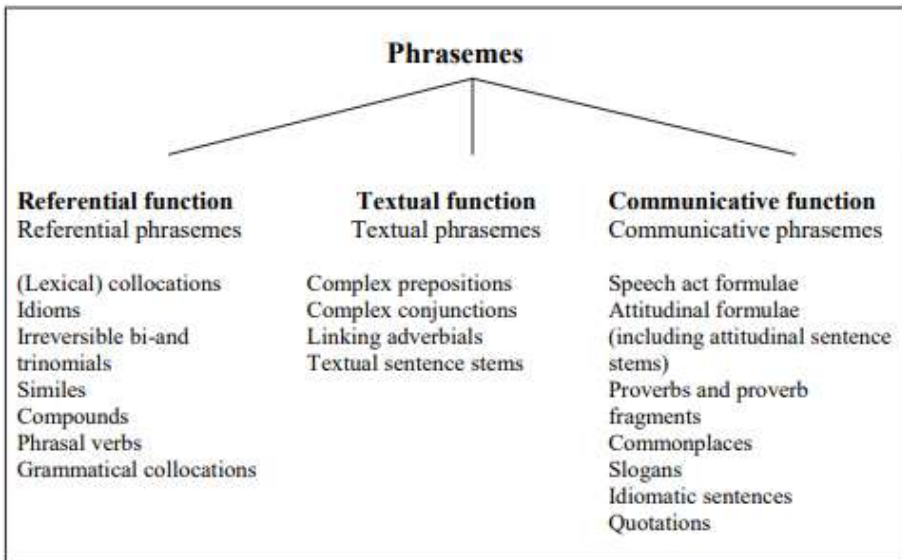
Berakar dari prinsip idiom Sinclair, pendekatan distribusional membawa fraseologi dari pinggiran menjadi pusat analisis bahasa karena satuan fraseologis tersebut pervasif (hadir di banyak tempat) dalam bahasa. Pendekatan tersebut membuat cakupan fraseologi semakin luas, tidak hanya terbatas pada ungkapan tetap, seperti idiom dan peribahasa. Pendekatan distribusional tidak bergantung pada kriteria makna untuk menentukan apa itu satuan fraseologis.



Gambar 2 Kategori distribusional oleh Granger & Paquot [2]

Pada pendekatan distribusional tidak terdapat kategorisasi satuan fraseologis. Namun, pendekatan ini menghasilkan tipologi jenis satuan yang diperoleh dari prosedur ekstraksi yang berbeda. Seperti yang terlihat pada gambar di atas, terdapat dua subdivisi utama yang dihasilkan dari dua metode ekstraksi, yaitu analisis n-gram dan analisis kookuransi (kemunculan bersama). Analisis n-gram adalah metode yang memungkinkan untuk ekstraksi rangkaian sinambung berulang dari dua kata atau lebih, yaitu ungkapan berulang, tanpa melihat keidiomatisannya dan tanpa melihat status strukturalnya [7]. Rangkaian yang terekstrak disebut n-gram (atau bigram, trigram, dst.) [8], *lexical bundle* [7] [9], *cluster* [10], *chain* [11], *reccurent sequence* [12], dan *reccurent word combination* [13]. Contoh rangkaian ekstraksinya, di antaranya, adalah *in the case of*, *on the other hand*, *the use of*, dan *the fact that*.

Sementara itu, analisis kookuransi mengekstraksi rangkaian dua kata yang taksinambung. Analisis ini merupakan pengungkapan kookuransi kata secara statistis. Dengan kata lain, analisis kookuransi berfokus pada hubungan statistis antarbutir leksikal. Kata yang lebih sering muncul bersama dengan kata inti disebut *kolokat*, sementara *kolokasi* mencerminkan peristiwa probabilistik yang muncul dari pemilihan bersama kata-kata yang berulang oleh penutur bahasa dalam bahasa tertentu. Granger dan Paquot [2] mencoba mengintegrasikan pendekatan tradisional dan pendekatan berbasis korpus dengan mengembangkan klasifikasi yang diajukan oleh Burger [14]. Satuan fraseologis dibagi menjadi tiga kategori utama: (1) frasem referensial, (2) frasem tekstual, dan (3) frasem komunikatif. Frasem referensial digunakan untuk menyampaikan pesan isi (*content message*), yaitu yang merujuk pada objek, fenomena, atau fakta kehidupan nyata. Yang tergolong dalam kategori ini adalah kolokasi leksikal dan gramatikal, idiom, simile, bi- dan trinomial, kata majemuk, dan verba frasal. Frasem tekstual digunakan untuk menyusun dan menata isi (informasi referensial) suatu teks atau wacana. Yang termasuk di dalamnya adalah preposisi kompleks, konjungsi kompleks, adverbial penghubung, dan dasar kalimat tekstual. Frasem komunikatif digunakan untuk mengungkapkan perasaan atau keyakinan terhadap isi proposisional atau secara eksplisit menyampaikan kepada mitra/teman wicara, untuk memfokuskan perhatian, melibatkan mitra wicara sebagai pelibat wacana, atau untuk memengaruhi mereka. Spektrum fraseologis tersebut dapat dilihat dalam gambar berikut.



Gambar 3 Kategori distribusional oleh Granger & Paquot [2]

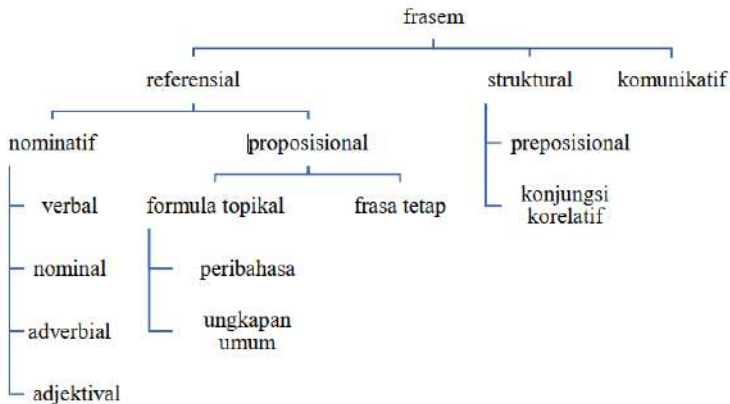
2.2 Fraseologi dalam Bahasa Indonesia

Kridalaksana [15] mendefinisikan fraseologi sebagai “1. kaidah perangkaan kata; 2. cara-cara memakai kata atau frase dalam tulisan atau ujaran; gaya bahasa; 3. Perangkat ungkapan yang dipakai oleh orang atau kelompok tertentu; mis. yang dipakai oleh para nelayan, pedagang, montir, dsb”. Dari definisi tersebut, tampak bahwa selain berkaitan dengan kaidah penggabungan kata, fraseologi juga identik dengan ungkapan (*expression*). Namun, cakupan ungkapan tersebut dapat berupa ujaran fonologis atau grafemis dan juga berupa idiom [15].

Ajie [16] menggunakan istilah frasem (*phrase*) sebagai satuan utama dalam fraseologi. Pengidentifikasian frasem didasarkan pada tiga kriteria, yaitu polileksikalitas (*polylexicality*), keajekan (*fixedness*), dan keidiomatisan (*idiomaticity*). Polileksikalitas artinya bahwa sebuah ungkapan setidaknya terdiri atas dua kata. Kriteria keajekan terpenuhi jika sebuah ungkapan hanya digunakan dalam sebuah bentuk kombinasi kata-kata tertentu (atau kombinasi kata-kata tertentu yang sangat mirip). Adapun keidiomatisan terpenuhi jika makna fraseologisnya tidak dapat ditelusuri secara sintaksis dan semantis dari makna komponen pembentuknya.

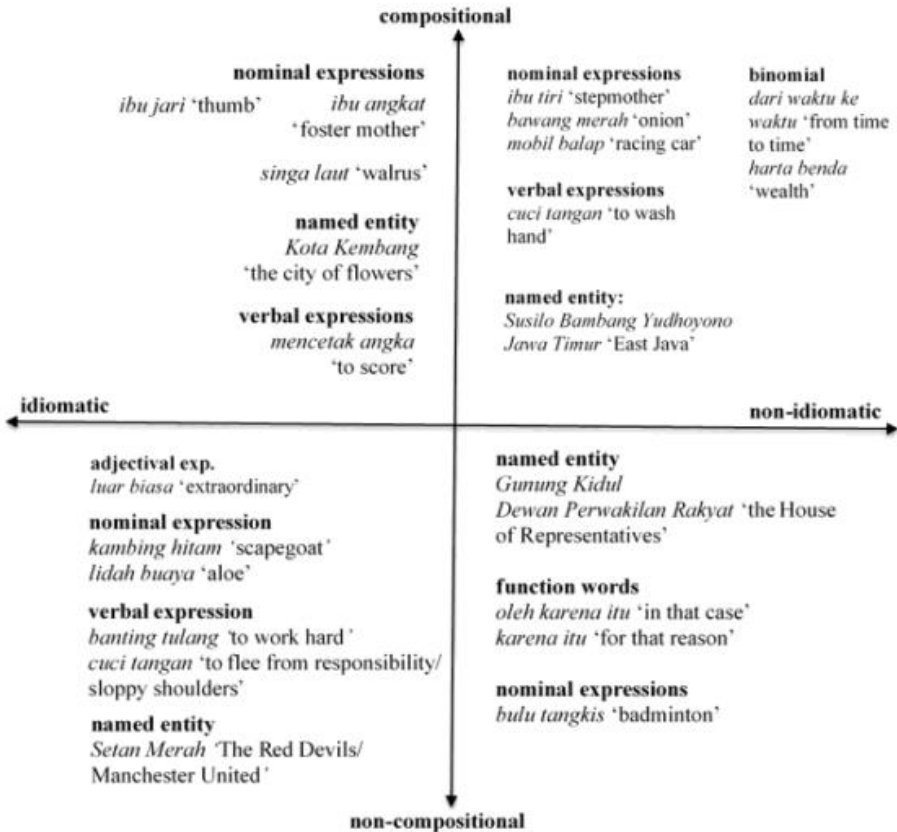
Berdasarkan pembagian frasem yang diajukan oleh Burger [14], Ajie [16]

membagi frasem menjadi tiga kategori, yaitu (1) referensial, (2) struktural, dan (3) komunikatif. Frasem referensial terbagi lagi atas nominatif dan proposisional. Secara sintaksis, frasem referensial nominative memiliki tiga kategori: (i) frasem verbal (misalnya *memejahijaukan*, *naik darah*), (ii) frasem nominal (misalnya *meja hijau*, *harga mati*), (iii) frasem adjektival (misalnya *malu-malu kucing*, *semanis madu*), dan iv) frasem adverbial (misalnya *secepat kilat*, *dengan penuh pertimbangan*), sedangkan frasem proposisional mempunyai dua kategori: (i) formula topikal, terdiri atas peribahasa (misalnya *maksud hati memeluk gunung, apa daya tangan tak sampai*; *tak ada gading yang tak retak*) dan tempat umum/ungkapan umum (misalnya *apa yang terjadi*, *terjadilah*; *hidup itu hanya satu kali*); dan (ii) frasa tetap (misalnya *ketegangan mencair*; *sesuatu bergoyang*). Secara semantis frasem referensial dibagi menjadi tiga jenis: frasem kolokasi (misalnya *gosok gigi*, *mengambil keputusan*), idiom sebagian (misalnya *daftar hitam*, *harga mati*), dan idiom penuh (misalnya *membanting tulang*). Sementara itu, frasem struktural terbagi atas frasem preposisional (misalnya *atas nama*) dan konjungsi korelatif (*tidak hanya...*, *tetapi juga*). Adapun yang termasuk frasem komunikatif di antaranya adalah formula rutin, seperti *selamat pagi* dan *hormat saya*. Pembagian frasem tersebut dapat digambarkan seperti bagan berikut.



Gambar 4 Klasifikasi frasem menurut Ajie [16]

Suhardijanto [17], dengan menggunakan istilah *multi-word expression* (MWE), membagi ungkapan multikata berdasarkan dua poros kategori: komposisionalitas dan idiomatisitas, seperti yang tampak pada gambar berikut.



Gambar 5 Klasifikasi ungkapan multikata menurut Suhardijanto [17]

Berdasarkan dua poros tersebut dihasilkan empat kategori ungkapan multikata, yaitu (1) ungkapan verbal, misalnya *minta izin*, *cuci tangan*, *banting tulang*, dan *tinggal di*; (2) ungkapan nominal, misalnya *kambing hitam*, *ibu jari*, *bulu tangkis*, *harta benda*, dan *waktu ke waktu*; (3) ungkapan adjektival, misalnya *tegak lurus* dan *luar biasa*; (4) ungkapan adverbial, misalnya *lebih kurang* dan *lebih lanjut*; dan (5) kata tugas, misalnya *di dalam*, *di luar*, *akan tetapi*, dan *oleh karena itu*. Pembagian ungkapan yang dilakukan oleh Suhardijanto *et al.* ini belum menyentuh ungkapan yang berupa klausa.

3. METODOLOGI

Penelitian ini menggunakan metode gabungan kuantitatif dan kualitatif dengan anjakan tergerakkan korpus (*korpus driven*). Korpus yang digunakan terdiri atas 25.499.772 kata (token), 261.175 kata unik (tipe), dan 16.453 teks yang berasal dari 12 ranah, yaitu koran, majalah, cerpen, novel, buku teks, jurnal, disertasi/ tesis/ skripsi, biografi, populer, perundang-undangan, surat resmi, dan laman resmi.

Pengumpulan data dilakukan dengan mengekstrak data n-gram dengan menggunakan WordSmith 8.0. Istilah n-gram biasa digunakan dalam bidang linguistik komputasional. Metode n-gram ini menggunakan program komputer untuk menghitung frasa atau rangkaian kata sinambung yang muncul dalam data korpus. beragam panjang n-gram yang dapat dihitung secara terpisah [18]. Dalam penelitian ini, n-gram diekstrak adalah 2—5 kata dengan menggunakan pengolahan korpus WordSmith 8.0 [19].

Dalam identifikasi majemuk (baik kata maupun frasa), satuan kata ortografis digunakan sebagai dasar. Pada penelitian ini, sebuah gabungan kata dikategorikan sebagai majemuk jika gabungan kata itu muncul sekurang-kurangnya 40 kali per juta kata dengan asumsi makin tinggi ambang batas yang digunakan, makin tinggi pula tingkat kepastian bahwa suatu rangkaian kata itu tergolong dalam majemuk. Kemunculan gabungan kata itu harus tersebar setidaknya pada lima teks yang berbeda dalam register itu. Hal itu dilakukan untuk menghindari idiosinkrasi penutur/penulis. Berdasarkan parameter tersebut, terekstraksi 1.094 gabungan kata. Dari jumlah tersebut terdapat gabungan kata yang berbentuk potongan klausa, seperti *ini merupakan, itu akan, ia tidak, ini adalah, dan ini sangat*, serta potongan frasa, seperti *informasi dan, komunikasi dan, barang dan, pihak yang, apa yang, hal yang, dan orang yang*. Bentuk-bentuk tersebut tidak dianalisis.

4. HASIL DAN PEMBAHASAN

Pada bagian ini majemuk diurai berdasarkan struktur, kategori, makna, dan fungsi. Berikut adalah uraian lengkapnya.

4.1 Struktur Majemuk

Beberapa ahli bahasa meragukan kata majemuk sebagai hasil dari proses morfologis. Mereka mencoba menjelaskan bahwa susunan kata majemuk itu merupakan hasil pembekuan konstruksi sintaktis. Fokker [20], salah satunya, berpendapat bahwa hubungan konstruksi seperti *memeras keringat* dan *membanting tulang* adalah hubungan sintaktis antara verba dan objeknya. Dengan kata lain, terjadi proses penurunan derajat (*downgrading*), yakni konstruksi predikatif yang diturunkan menjadi kata. Fokker juga mencatat sejumlah konstruksi yang biasa disebut kata majemuk sebagai frasa yang terdiri atas komponen-komponen yang tetap (sejenis kolokasi). Senada dengan Fokker, Moeliono dalam Masinambouw [21], dengan memperhatikan kumpulan data kata majemuk dari Kridalaksana yang memperlihatkan bahwa secara diakronis terjadi proses peralihan dari frasa ke bentuk majemuk, membagi bentuk majemuk menjadi dua, yaitu majemuk kata dan majemuk frasa. Majemuk kata memiliki ciri yang cenderung bersifat kata, sementara majemuk frasa cenderung memiliki ciri-ciri frasa [22]. Dalam hal ini, penulis bersetuju dengan pendapat mereka dengan didukung temuan data dalam penelitian ini.

Majemuk kata memiliki unsur-unsur penyusun yang berikatan erat sehingga tidak dapat disisipi kata lain dan letaknya tidak dapat dipertukarkan. Gabungan kata *tanah air*, *orang tua*, dan *budi daya* tidak mengizinkan penyisipan satuan dan pertukaran posisi atau urutan. Selain itu, Dari segi gramatika, majemuk kata biasanya dapat menempati satu slot dalam kalimat, yakni sebagai verba, nomina, adjektiva, dan seterusnya. Misalnya, *lalu lintas* memiliki unsur *lalu* dan *lintas* yang keduanya berkelas kata verba. Meskipun demikian, dalam konstruksi kalimat *lalu lintas* memiliki kedudukan sebagai nomina.

Sementara itu, majemuk frasa memiliki unsur-unsur penyusun yang berikatan longgar. Majemuk ini dapat disisipi kata lain, misalnya *laba rugi* dan *ilmu pengetahuan*. Bentuk tersebut dapat disisipi *atau* sehingga menjadi *laba atau rugi* serta disisipi *dan* sehingga menjadi *ilmu dan pengetahuan*. Meskipun demikian, majemuk ini tidak dapat diubah urutannya menjadi **rugi laba* atau **pengetahuan ilmu*. Secara gramatika, kedua majemuk frasa tersebut tersusun dari unsur penyusun berkategori nomina sehingga dalam konstruksi kalimat, majemuk frasa *laba rugi* dan *ilmu pengetahuan* berkedudukan sebagai nomina pula.

Secara morfologis, majemuk dapat dibentuk dari bentuk dasar bebas, bentuk berimbuhan, atau gabungan keduanya.

A. Bentuk dasar bebas + bentuk dasar bebas

Misalnya:

media massa	wajib pajak
media sosial	tenaga kerja
jangka waktu	kartu kredit

B. Bentuk dasar bebas + bentuk berimbuhan

Misalnya:

ilmu pengetahuan
tahun anggaran
surat pernyataan
surat berharga

C. Bentuk berimbuhan + bentuk dasar bebas

Misalnya:

peserta didik
lingkungan hidup
pelaku usaha
perguruan tinggi

D. Bentuk berimbuhan + bentuk berimbuhan

Misalnya:

kendaraan bermotor
pengambilan keputusan
perusahaan penjaminan
perusahaan penjaminan

4.2 Kategori Majemuk

Berdasarkan kategori unsurnya, majemuk dapat dikelompokkan menjadi majemuk verbal, majemuk nominal, majemuk adjektival, majemuk adverbial, majemuk preposisional, dan majemuk numeral. Keenam jenis majemuk tersebut dapat diuraikan sebagai berikut.

A. Majemuk Verbal

Yang dimaksud dengan majemuk verbal adalah majemuk yang pada satuan klausa berkategori verba. Perhatikan contoh berikut.

- Kami akan *bekerja sama* dengan pemerintah daerah dan dinas pertambangan untuk mengawasi mereka.

b. Kami siap *bertanggung jawab* atas kesalahan kami di masa lalu. Sebagai pengisi fungsi predikat kedua majemuk berintikan bekerja dan bertanggung yang berkategori verba. Komposisi majemuk verbal dapat dibentuk dari:

- i. Verba + verba, seperti *bertanggung jawab*
- ii. Verba + nomina, seperti *jatuh tempo*
- iii. Verba + adjektiva, seperti *bekerja sama*
- iv. Verba + preposisi, seperti *berhubungan dengan, bertentangan dengan, terdiri atas*
- v. Adverbial + verba, seperti *akan datang, dapat digunakan*

B. Majemuk Nominal

Majemuk nominal adalah majemuk yang dibentuk dari:

- i. Nomina + nomina, seperti *tenaga kerja, tanah air, masa depan, ilmu pengetahuan, surat kabar, wali kota, pemegang saham*
- ii. Nomina + verba, seperti *bahan bakar*
- iii. Nomina + adjektiva, seperti *jangka panjang, jangka pendek, orang tua, media sosial, bank sentral, bank umum, antara lain*
- iv. Nomina + pronominal, seperti *sementara itu, saat ini, saat itu, ketika itu, waktu itu, hal itu, hal ini*
- v. Numeralia + nomina, seperti *banyak hal, beberapa hal, per tahun, beberapa kali, dua tahun*

C. Majemuk Adjektival

Majemuk adjektival adalah majemuk yang dibentuk dari:

- i. Adjektiva + numeral, seperti *salah satu, salah seorang, sama sekali*
- ii. Adjektiva + preposisi, seperti *sama dengan, sesuai dengan, dekat dengan, jauh dari, khusus untuk, penting dalam*
- iii. Adjektiva + adverbial, seperti *tentu saja, baru saja*
- iv. Adverbial + adjektiva, seperti *sangat besar sangat baik, lebih mudah, lebih rendah*

D. Majemuk Adverbial

Majemuk adverbial adalah majemuk yang dibentuk dari:

- i. Adverbial + adverbial, seperti *akan lebih, akan terus, juga tidak, tidak lagi, tidak hanya, tak hanya, tidak terlalu*
- ii. Adverbial + modal, seperti *juga dapat, juga bisa, juga akan, hanya*

bisa, hanya dapat, tidak mampu, tidak akan, tak perlu, tak akan
iii. Pronomina + adverbia, seperti *begitu saja*

E. Majemuk Preposisional

Majemuk preposisional adalah majemuk yang dibentuk dari:

- i. Preposisi + nomina, seperti *oleh sebab, dengan tujuan, di antara, dengan cara, di samping*
- ii. Preposisi + verba, seperti *sebagaimana dimaksud, dengan menggunakan, untuk memenuhi*
- iii. Preposisi + pronominal, seperti *untuk itu, dengan demikian, seperti ini, seperti itu*
- iv. Preposisi + konjungsi, seperti *akan tetapi, oleh karena*

F. Majemuk Numeral

Majemuk numeral adalah majemuk yang dibentuk dari:

- i. Numeralia + nomina, seperti *dua puluh, dua kali*

4.1 Makna Majemuk

Ada majemuk yang memiliki makna idiomatis, baik berupa idiom penuh maupun idiom sebagian. Selain itu, ada juga majemuk yang memiliki makna non-idiomatis. Ketiga jenis majemuk tersebut diuraikan dalam subbagian berikut.

A. Idiomatis Penuh

Majemuk dengan idiom penuh memiliki makna yang tidak dapat ditelusuri dari unsur-unsur yang menyusunnya. Perhatikan contoh berikut.

- a. Tak sedikit *orang tua* yang menuntut anak-anaknya agar mendapat nilai-nilai akademik yang bagus.
- b. Dia mengucapkan *terima kasih* karena sudah belajar banyak di KPK.
- c. Di balik kesuksesannya, Risma memiliki latar belakang kehidupan yang penuh warna.
- d. Peluang itu sedang terhidang di halaman *Tanah Air* kita.
- e. Ini tentu bukanlah kecelakaan *lalu lintas* yang pertama di jalan raya.

Majemuk *orang tua, terima kasih, latar belakang, tanah air, dan lalu lintas* memiliki makna idiomatis. Dalam konteks

tersebut, *orang tua* bermakna ‘ayah-ibu’, *terima kasih* bermakna ‘rasa syukur’, *latar belakang* bermakna ‘pengalaman hidup’, *tanah air* bermakna ‘negeri tempat kelahiran’, dan *lalu lintas* bermakna ‘perihal perjalanan di jalan’

B. Idiomatis Sebagian

Majemuk dengan idiom sebagian memiliki makna yang dapat ditelusuri dari salah satu unsur penyusunnya. Perhatikan contoh berikut.

- Pemilu merupakan *salah satu* media bagi rakyat untuk terlibat langsung dalam proses bernegara.
- Kepemilikan perusahaan itu *atas nama* orang lain.
- Utang yang segera *jatuh tempo* tak mungkin dilunasi hanya dengan menarik angkot.
- Solusi daur ulang ini adalah bentuk *tanggung jawab* kami dalam menjaga lingkungan.
- Rendahnya nilai *mata uang* kita terhadap mata uang asing disebabkan kegagalan pemerintah menjaga tingkat inflasi.

Majemuk *salah satu*, *atas nama*, *jatuh temp*, *tanggung jawab*, dan *mata uang* merupakan majemuk diomatis Sebagian. Majemuk tersebut dapat ditelusuri dari kata *satu*, *nama*, *tempo*, *tanggung*, dan *uang*. Sementara itu, unsur kata lainnya memiliki makna yang berbeda dari makna literalnya. Majemuk *salah satu* bermakna ‘satu di antara yang ada’, *atas nama* bermakna ‘dengan nama’, *jatuh tempo* bermakna ‘batas waktu pembayaran/penerimaan’, *tanggung jawab* bermakna ‘keadaan wajib menanggung segala sesuatunya’, serta *mata uang* bermakna ‘satuan uang suatu negara’.

C. Non-idiomatis

Majemuk non-idiomatis adalah majemuk yang maknanya dapat ditelusuri dari unsur-unsur penyusunnya. Perhatikan contoh berikut.

- Karena itu, ia meminta para pengusaha *dalam negeri* meningkatkan daya saing.
- Akun beban penyusutan akan tampak dalam laporan *laba rugi*, sedangkan akun akumulasi penyusutan akan terlihat dalam neraca.
- Air mata* masih menggenang di kelopaknya.
- Dari *tempat tinggal* tersangka, polisi menemukan zat kimia potasium yang biasa digunakan untuk bahan peledak.

- e. Hanya enam pecatur yang akan mengikuti pelatnas *jangka panjang* pada setiap kelompok

Makna majemuk *dalam negeri*, yaitu *laba rugi*, *air mata*, *tempat tinggal*, dan *jangka panjang* dapat ditelusuri dari kedua unsur yang membentuknya. Dalam konteks tersebut, *dalam negeri* bermakna ‘lingkungan negeri sendiri’, *laba rugi* bermakna ‘untung rugi’, *air mata* bermakna ‘air yang meleleh dari mata’, *tempat tinggal* bermakna ‘rumah tempat orang tinggal’, serta *jangka panjang* bermakna ‘dalam waktu lama’. Majemuk yang berupa istilah (bidang ilmu) biasanya bersifat non-idiomatis, misalnya istilah *perusahaan asuransi*, *perusahaan pembiayaan*, dan *perusahaan penjaminan*.

4.3 Fungsi Majemuk

Fungsi majemuk dalam hal ini berkaitan dengan bagaimana bahasa bekerja untuk mencapai tujuan komunikatif tertentu. Penulis mengadopsi dan mengembangkan fungsi-fungsi yang telah digunakan oleh Biber [7] [9], Hyland [23] [24], dan Salazar [25]. Penelitian ini membagi fungsi-fungsi tersebut ke dalam tiga kelompok, yaitu fungsi referensial (*referential*), fungsi tekstual (*textual*), dan fungsi pendirian (*stance*). Berikut adalah penjabaran dari setiap fungsi tersebut.

A. Fungsi Referensial

Fungsi referensial merupakan fungsi untuk menyampaikan pesan isi (*content message*), yaitu yang merujuk pada objek, fenomena, atau fakta kehidupan nyata. Fungsi ini terdiri atas fungsi lokasi, prosedur, kuantifikasi, deskripsi, pengelompokan, dan topik. Paparan keenam fungsi tersebut adalah sebagai berikut.

1) Lokasi

Fungsi ini digunakan untuk menunjukkan waktu dan tempat/arah. untuk menunjukkan waktu, majemuk yang dapat digunakan adalah *pada masa*, *pada periode*, *pada saat*, *pada waktu*, *pada tahun*, *pada bulan*, *pada tanggal*, *pada usia*; *sejak tanggal*, *sejak tahun*, *ke depan*, *masa depan*, *masa lalu*, dan *akan datang*. Perhatikan contoh berikut.

- a. Menurut dia, *pada masa* pandemi, banyak orang mencari kegiatan agar imunitas tidak menurun.

- b. Pasal 29 KEK yang telah ditetapkan harus siap beroperasi paling lambat 3 (tiga) tahun *sejak tanggal* ditetapkan.
- c. Sulit membantah betapa pemilu yang *akan datang* itu akan menjadi pemilu yang lebih mahal dan juga lebih ketat tingkat kompetisinya dibandingkan dengan pemilu-pemilu yang lalu.
- d. Pembudidayaan ideologi kerakyatan merupakan modalitas paripurna guna menyongsong tantangan bangsa *ke depan*.

Sementara itu, untuk menunjukkan tempat, majemuk yang dapat digunakan adalah *di lingkungan, di kalangan, di sekolah, di rumah sakit, di luar negeri; bekerja di, berada pada, berada dalam; berasal dari, berdiri di, ada pada, datang dari, datang ke, duduk di, hidup di, dan terletak di*. Perhatikan contoh berikut.

- a. Banyak pengunjung asing ke Libya melukiskan, akademi itu *berada dalam* kompleks besar dengan pagar tembok yang kokoh.
- b. Ketika itu, planet Uranus *berada pada* jarak 4 miliar kilometer dari bumi.
- e. Danny Rayar, pemuda 17 tahun yang berada di lokasi kejadian, masih mengingat dengan jelas rentetan tembakan peluru tajam yang *datang dari* segala arah.
- f. Pemantauan juga dilakukan terhadap pemain-pemain muda Indonesia yang ada *di luar negeri*.

2) Prosedur

Fungsi ini digunakan untuk menunjukkan peristiwa, tindakan, dan metode. Majemuk yang digunakan dalam fungsi ini di antaranya adalah *diatur dalam, diatur dengan, diberikan kepada, diberikan oleh, digunakan dalam, digunakan oleh, digunakan sebagai, digunakan untuk, diikuti oleh, diisi dengan, dilaksanakan oleh, dilakukan dengan, dilakukan dalam, dilakukan melalui, dilengkapi dengan, bagi masyarakat, bagi perusahaan, untuk kepentingan, untuk mengetahui, untuk meningkatkan, adalah untuk, dengan cara; dengan menggunakan; secara langsung, dengan baik, dengan mudah, dengan cepat; dengan memperhatikan, sampai dengan, sejak awal, dalam waktu, dalam proses; dengan tujuan, untuk mencapai, untuk mengembangkan, untuk membantu, untuk mengatasi; kepada masyarakat, kepada pemerintah; bertujuan untuk; oleh pemerintah, oleh perusahaan, oleh masyarakat*. Berikut adalah beberapa contoh penggunaannya.

- a. Hal-hal mengenai warga negara dan penduduk *diatur dengan* undang-undang.
- b. Di sisi lain, upaya menjaga stabilitas juga *dilakukan melalui* pendalaman pasar keuangan domestik.
- c. Studi tersebut menawarkan cara ketiga untuk mengubah sel kulit menjadi sel induk dengan memanipulasi genom sel *secara langsung* menggunakan teknik peraturan gen CRISPR.
- d. Dengan kapabilitas rendah, tidak mudah bagi penduduk pedesaan *untuk mengembangkan* usaha, terutama di luar pertanian (*off farm*).

3) Kuantifikasi

Fungsi ini digunakan untuk menunjukkan ukuran, jumlah, proporsi dan perubahannya, termasuk pemerincian, misalnya *dengan jumlah, dengan harga; dalam berbagai, dalam beberapa, dalam setiap; di beberapa, di seluruh, di setiap; banyak hal, banyak orang, beberapa hal, beberapa hari, beberapa kali; dua hari, dua kali, dua orang, dua puluh, dua tahun, enam bulan, juta rupiah; sebagian besar; antara lain, dan di antaranya*. Perhatikan contoh berikut.

- a. Indonesia juga masuk jajaran 15 besar pemilik cadangan minyak di dunia *dengan jumlah* cadangan 11,6 miliar barel.
- b. Para pedagang bensin dadakan ini menjual Premium *dengan harga* hingga Rp 50.000 per liter atau lima kali lipat lebih mahal dari harga normal.
- c. Para petinju ini berlatih *dua kali* dalam sehari, pagi dan sore.
- d. Penyebaran tingkat kemiskinan kabupaten daerah tertinggal *sebagian besar* berada di atas garis tingkat kemiskinan nasional.
- e. Menurut dia, Pertamina akan mengeksekusi aset-aset yang dijaminan, *di antaranya* kilang dan tanah.

4) Deskripsi

Fungsi ini digunakan untuk menunjukkan keadaan, derajat, dan keberadaan, misalnya *berfungsi sebagai, berhubungan dengan, berkaitan dengan; berubah menjadi, dikenal sebagai, dikenal dengan; sangat baik, sangat besar, sangat penting, dan luar biasa*. Perhatikan contoh berikut.

- a. Di laut, terumbu karang ini *berfungsi sebagai* barisan penahan gempuran ombak dan pelindung pantai dari abrasi.

- b. Pasalnya, kasus tersebut *berkaitan dengan* oknum di bank, bukan sistem bank tersebut.
- c. Kedua kakinya dipakai untuk membantu membuat tas yang di Papua *dikenal dengan* nama noken.
- d. Temuan Komnas HAM seharusnya membuat kita semua tersadar bahwa ada masalah *sangat besar* yang mendesak diselesaikan oleh kita semua.

5) Pengelompokan

Fungsi ini digunakan untuk menunjukkan Kelompok, kategori, bagian, dan susunan, misalnya pada majemuk *sebagai bagian, sebagai bagian dari, bagian dari, menjadi bagian, terdiri atas, terdiri dari, sebagai salah satu, adalah salah satu, salah satu, satu dari, dan salah seorang*. Perhatikan contoh berikut.

- a. Segenap gambaran di atas menunjukkan lanskap luas pesantren *sebagai bagian dari* sub-kultur Islam di Indonesia.
- b. Bima Saksi *adalah salah satu* galaksi terbesar di jagat raya.
- c. Kesebelasan ini *salah satu dari* lima kesebelasan legenda nusantara yang lahir sejak era perserikatan 1930.
- d. Ia didata dan dijanjikan masuk sebagai *salah seorang* penerima bantuan bedah rumah.
- e. Lempengan *terdiri atas* beberapa segmen.

6) Topik

Fungsi ini berkaitan dengan hal yang sedang dibicarakan atau dibahas. Majemuk yang muncul pada fungsi ini pada umumnya berupa istilah yang terkait dengan bidang ilmu tertentu. Contohnya adalah sebagai berikut.

Istilah bidang ekonomi:

arus kas, badan hukum, badan usaha, badan usaha milik negara, bank sentral, bank umum, barang milik daerah, dana pension, jaminan sosial, jasa keuangan, kantor cabang, kantor pusat, kartu kredit, laporan keuangan, manajemen risiko, mata uang, modal sosial, modal ventura, nilai tukar, otoritas jasa keuangan, pemegang saham, pertumbuhan ekonomi, perusahaan asuransi, perusahaan pembiayaan, perusahaan penjaminan, rencana bisnis, suku bunga, sumber daya, sumber daya alam, sumber daya manusia, surat berharga, surat edaran, surat pernyataan, tahun anggaran, tenaga kerja, wajib pajak

Istilah bidang pemerintahan:

*anggota dewan, dasar negara, dewan pengawas, dewan perwakilan rakyat, direktorat jenderal, direktur jenderal, direktur utama, kepala daerah, kepala desa, ketua umum
lembaran negara, mahkamah agung, menteri keuangan, orde baru, partai politik, pegawai negeri, pemberantasan korupsi, pemerintah daerah, pemerintah pusat, pemerintahan daerah
pemilihan umum, pengambilan keputusan, peraturan daerah, peraturan Menteri, peraturan pemerintah, peraturan perundang-undangan, peraturan presiden, peraturan zonasi, reformasi birokrasi, tata cara, tata kelola, tata ruang, tindak pidana, unit kerja, wakil ketua, wakil presiden, wali kota, warga negara*

Selain berupa istilah, majemuk yang berfungsi sebagai topik dapat juga berupa kosakata umum dan nama diri. Berikut adalah majemuk berupa kosakata umum.

air mata, akhir tahun, bagi hasil, bahan bakar, bahan baku, bandar udara, budi daya, daya saing, hak asasi manusia, hari kerja, ilmu pengetahuan, jangka panjang, jangka pendek, jangka waktu, kata kunci, kendaraan bermotor, kereta api, kerja sama, laba rugi, lalu lintas, latar belakang, lingkungan hidup, luar negeri, masa depan, masa lalu, media massa, media sosial, orang tua, pelaku usaha, rumah sakit, rumah tangga, sepak bola, surat kabar, tanah air, tanggung jawab, tempat tinggal, terima kasih

Sementara itu, majemuk berupa nama diri adalah sebagai berikut.

*Allah Swt, Amerika Serikat, Bahasa Indonesia, Bahasa Inggris, Bangsa Indonesia,
Bank Indonesia, Bung Karno, DKI Jakarta, Hindia Belanda, Jawa Barat, Jawa Tengah,
Jawa Timur, Presiden Joko Widodo, Republik Indonesia*

B. Fungsi Tekstual

Fungsi tekstual merupakan fungsi yang digunakan untuk menyusun dan menata isi (informasi referensial) suatu teks. Fungsi ini terbagi atas aditif, konstatif, inferensial, kausatif, penstrukturan, pembingkai, generalisasi, dan komparatif.

1) Aditif

Fungsi ini digunakan untuk membuat hubungan penambahan antara suatu unsur dan unsur lain, misalnya *selain itu*, *di samping*, *dengan kata lain*; *setelah itu*; *pada akhirnya*; *ketika itu*, Perhatikan contoh berikut.

- a. Dalam aksinya, para perampok langsung menggasak berbagai macam perhiasan emas yang tersimpan pada laci-laci kecil dalam brankas dengan total nilai sekitar Rp 6 miliar. *Selain itu*, mereka juga mengambil uang tunai sebesar Rp 30,9 juta.
- b. Isu dalam sebuah pemilu faktor yang paling relevan biasanya adalah ekonomi. *Di samping* ekonomi, Jakarta punya segudang pekerjaan rumah, dari banjir, sampah, macet, hingga pelbagai masalah ekonomi serta sosial dan budaya.
- c. Mereka pun mendapatkan asimilasi ataupun pembebasan bersyarat. *Dengan kata lain*, hukuman yang sudah ringan itu menjadi semakin enteng gara-gara fasilitas ini.

2) Kontrastif

Fungsi ini digunakan untuk mengontraskan suatu unsur dengan unsur lain, misalnya *di sisi lain*, *akan tetapi*, *sementara itu*. Perhatikan contoh berikut.

- a. Warga biasa berkorban menerima pengurangan hak-hak pensiunan karena pengetatan anggaran negara. *Di sisi lain*, warga kaya berbuat curang dengan menghindari pajak.
- b. Pertumbuhan ekonomi selama ini cenderung meningkat dari tahun ke tahun. *Akan tetapi*, kualitas pertumbuhannya makin mengkhawatirkan.
- c. Dalam studi sebelumnya, respons gen pada pasien alzheimer diukur hanya di bagian jaringan otak. *Sementara itu*, studi ini tidak membedakan jenis sel.

3) Inferensial

Fungsi ini digunakan untuk membuat generalisasi dan simpulan yang ditarik berdasarkan data, misalnya *dengan demikian* dan *menunjukkan bahwa*. Perhatikan contoh berikut.

- a. Dengan sistem kandang kaca tersebut, ayam yang dibudidayakan tersebut lebih leluasa bergerak dan ini diharapkan dapat mengurangi tingkat stres mereka. *Dengan demikian*, produktivitas mereka diharapkan akan meningkat.

- b. Adanya batu sandaran itu *menunjukkan bahwa* lokasi itu pernah menjadi tempat untuk bermusyawarah.

4) Kausatif

Fungsi ini digunakan untuk menandakan adanya hubungan sebab dan akibat antara unsur satu dan unsur lain, misalnya *oleh karena itu*, *oleh karena*, *karena itu*, *oleh sebab*, dan *untuk itu*. Perhatikan contoh berikut.

- a. Perusahaan sesungguhnya memiliki tujuan yang mulia dan luhur apabila *setting*-nya didasarkan kepada perusahaan yang berketuhanan (*God corporate Governance*). *Oleh karena itu*, sudah saatnya cara pandang yang mendikotomikan antara agama dan bisnis itu untuk diakhiri.
- b. Mereka tidak bisa mengeluarkan iuran BPJS yang naik 100 persen. *Oleh sebab itu*, penurunan kelas menjadi salah satu pilihan terbaik.
- c. Inflasi yang tinggi jelas akan menambah beban hidup rakyat. *Untuk itu*, pemerintah telah merencanakan pemberian bantuan langsung sementara masyarakat (BLSM) Rp 17,08 triliun kepada 18,5 juta keluarga miskin (74 juta jiwa).

5) Penstruktur

Fungsi ini digunakan untuk menata teks atau yang mengarahkan pembaca pada teks tertentu, misalnya *sebagaimana dimaksud pada*, *sebagaimana dimaksud dalam*; *sebagai berikut*, *berbunyi sebagai berikut*, *di bawah ini*, *di atas*, *seperti ini*, *seperti itu*; *pada kalimat*, *pada tabel*, *pada gambar*, *pada bagian*, *pada angka*, *pada ayat*, Perhatikan contoh berikut.

- a. Penetapan ketua pengadilan *sebagaimana dimaksud pada* ayat (3) tidak dapat diajukan upaya hukum.
- b. Penerbitan izin *sebagaimana dimaksud dalam* Pasal 6 ayat (1) tidak dikenakan retribusi.
- c. Dalam menelusuri hal tersebut maka perlu didalami hal-hal *sebagai berikut*. Pertama, motivasi untuk melakukan penyiaran agama tetap menjadi hal yang intrinsik pada semua agama.
- d. Namun, seperti telah disebutkan *di atas*, sistem terpadu dan aturan yang jelas sangat terasa di Malaysia.
- e. Hubungan antara derajat curah hujan dan intensitas curah hujan dapat dilihat *pada Tabel 1*, sedangkan keadaan curah hujan *pada Tabel 2*.

6) Pembingkai

Fungsi ini digunakan untuk mengondisikan argumen dengan mengkhususkan keadaan yang terbatas, misalnya *dalam hal ini, dalam penelitian ini; dalam konteks, dalam rangka, dalam hal, dalam kondisi, dalam bidang, dari sisi, dari segi; sebagai salah satu, berdasarkan hasil, dan dilihat dari*. Perhatikan contoh berikut.

- a. Persamaan visi dan prioritas maritim antara Indonesia dan Norwegia ini tentunya membuka peluang kerja sama yang luas bagi kedua negara. *Dalam hal ini*, pemerintah dan para pelaku industri Norwegia siap berbagi pengalaman dan keahliannya untuk mendukung visi maritim Indonesia.
- b. *Dalam konteks* pengarusutamaan pembangunan pertanian ini, penulis melihat setidaknya upaya untuk mengembalikan ruh pertanian, perikanan, dan kelautan harus segera dilakukan melalui beberapa langkah strategis.
- c. *Dari segi* investasi, minat rumah tangga untuk berinvestasi di sektor properti masih merupakan preferensi yang tertinggi setelah tabungan/deposito.
- d. *Sebagai salah satu* bentuk komitmen dan konsistensi kepedulian dalam mendukung kemajuan serta peningkatan kualitas pendidikan anak bangsa, Bank Indonesia (BI) memberikan beasiswa kepada mahasiswa dari beberapa Perguruan Tinggi Negeri (PTN) di Indonesia.
- e. *Berdasarkan hasil* pantauan BMKG, masih ada pertumbuhan bibit badai atau tekanan rendah di Samudera Hindia dan Australia.

7) Penggeneral

Fungsi ini digunakan untuk menandakan fakta atau pernyataan yang secara umum sudah berterima, misalnya *pada umumnya, pada dasarnya, secara umum, secara keseluruhan, dan dianggap sebagai*. Perhatikan contoh berikut.

- a. *Pada umumnya*, karnivora memiliki tingkat enzim yang untuk memecah serat dan karbohidrat yang rendah.
- b. Semua anak *pada dasarnya* punya bakat yang nanti akan muncul secara alami.
- c. Kebiasaan tekanan kenaikan harga akan berlangsung sampai periode mudik-balik dan *dianggap sebagai* suatu kewajaran.
- d. *Secara keseluruhan*, pengeluaran konsumsi rumah tangga Indonesia

berdasarkan kemampuan daya beli di Asia-Pasifik mencapai 12.041,7 miliar dolar AS.

8) Komparatif

Fungsi ini digunakan untuk membandingkan suatu unsur dengan unsur lain, misalnya *dibandingkan dengan, lebih rendah, lebih mudah, lebih luas, dan lebih kecil*. Perhatikan contoh berikut.

- a. Di pasar darurat Pasar Pagi, misalnya, pedagang dan pembeli semakin ramai *dibandingkan dengan* 3-4 bulan lalu.
- b. Sementara itu, pemerintah memperkirakan realisasi impor daging beku tahun ini *lebih rendah* daripada kuota yang telah ditetapkan.
- c. Semua bukti ini menyatakan bahwa mutasi yang kami identifikasi *sama dengan* yang direkam van Gogh dalam lukisannya pada era 1800-an,

C. Fungsi Pendirian

Fungsi ini digunakan untuk menyampaikan sikap dan evaluasi penulis terhadap suatu proposisi. dan terbagi atas pendirian epistemik dan pendirian sikap.

1) Pendirian Epistemik

Pendirian epistemik berarti pendirian subjektif si pembicara terhadap kebenaran terhadap proposisi. Majemuk dengan fungsi ini digunakan sebagai sebagai pagar (*hedge*) dan menunjukkan tingkat probabilitas tertentu. Majemuk dengan fungsi itu di antaranya adalah *belum ada, berbeda dengan, bertentangan dengan, sesuai dengan, sama dengan, sama seperti, sudah ada, sudah menjadi, telah dilakukan, telah ditetapkan, telah menjadi, tidak memiliki, tidak tahu; belum pernah, bisa jadi, bisa saja, bukan hanya, hanya bisa, hanya dapat, hanya untuk, juga telah, juga tidak, sama sekali tidak, sering kali, sudah tidak, tidak pernah, tidak ada, tidak hanya, tidak pernah; tentu saja, cukup jelas, sangat baik, sangat besar, sangat penting, baik untuk, dekat dengan, jauh dari, jauh lebih, kurang dari, penting bagi, penting dalam, dan penting untuk*. Berikut beberapa contoh penggunaannya.

- a. Pengerjaan videotron dilaksanakan tidak *sesuai dengan* spesifikasi dan volume sehingga telah menimbulkan kerugian keuangan negara.
- b. Bila itu yang terjadi, *bisa jadi* negeri ini sedang dalam bahaya.
- c. Gagasan ini *tentu saja* sangat menarik dan penting untuk dipikirkan dan ditindaklanjuti.

- d. Meski suaranya pelan, *cukup jelas* kutangkap liriknya yang memecah konsentrasiku membaca.
- e. *Penting bagi* orangtua untuk mengawasi penggunaan gawai bagi anak-anaknya.

2) Pendirian Sikap

Fungsi ini digunakan untuk menyatakan keinginan, kewajiban, rencana/prediksi, dan kebolehan/kemampuan.

- i. Keinginan: *diharapkan dapat, tidak mau*

Misalnya:

- a. Data yang terkumpul *diharapkan dapat* mengungkap struktur bagian dalam Mars dan komposisinya.
- b. Kami *tidak mau* mengambil risiko.

- ii. Kewajiban: *harus dilakukan, harus memiliki, perlu dilakukan, juga harus, tak perlu, tidak perlu*

Misalnya:

- a. Agar mereka tertarik, saya *harus memiliki* tubuh atletis.
- b. Polisi dan jaksa pun *tidak perlu* repot memproses pengaduan kasus ini.
- c. Semua hal tersebut *harus dilakukan* secara masif dan sistemis agar semua komponen masyarakat menjadi tangguh bencana.

- iii. Rencana/Prediksi: *akan ada, akan terjadi, akan menjadi, akan membuat, akan dilakukan; akan terus, akan lebih, akan semakin, juga akan, tak akan, dan tidak akan*

Misalnya:

- a. Jika media tidak mempunyai sikap, ia hanya akan menjadi agen ‘penjualan’ berita.
- b. Bursa memang diperkirakan *akan terus* bergejolak lebih karena isu global, bukan karena *reshuffle*.
- c. Kemacetan di Ibu Kota *tidak akan* pernah selesai dengan memberikan fasilitas jalan bagi pengguna kendaraan.

- iv. Kebolehan/Kemampuan: *bisa dilakukan, bisa menjadi, dapat digunakan, dapat dilakukan, dapat dilihat, dapat melakukan, dapat memberikan, dapat meningkatkan, dapat menjadi, tak bisa, tidak dapat, tidak mampu, tidak boleh, juga bisa, dan juga dapat*.

Misalnya:

- a. Ibadah wajib hingga sunah *dapat dilakukan* setiap saat oleh para jamaah secara intensif selama 40 hari di Arab Saudi untuk haji dan sekitar 14-20 hari ibadah umrah.
- b. Makan makanan yang seimbang dan bergizi *dapat meningkatkan* kebugaran tubuh dan perasaan bahagia yang mampu melawan stres.
- c. Aturan yang ada sudah *tidak mampu* menjangkau kompleksitas persoalan yang terjadi di lapangan.

5. SIMPULAN DAN SARAN

Berdasarkan hasil pembahasan di atas, pendekatan distribusional yang berbasis frekuensi kemunculan dapat digunakan sebagai salah satu cara dalam menentukan gabungan kata yang berupa majemuk. Majemuk yang dihasilkan dapat berisifat kata dan bersifat frasa. Majemuk yang bersifat kata disebut majemuk kata, sedangkan majemuk yang bersifat frasa disebut majemuk frasa. Majemuk tersebut dapat terbentuk dari kata dasar bebas, kata berimbuhan, atau kombinasi keduanya. Berdasarkan kategori unsurnya, majemuk dapat dikelompokkan menjadi majemuk verbal, majemuk nominal, majemuk adjektival, majemuk adverbial, majemuk preposisional, dan majemuk numeral.

Dari analisis didapati bahwa ada majemuk yang memiliki makna idiomatis, baik berupa idiom penuh maupun idiom sebagian. Selain itu, ada juga majemuk yang memiliki makna non-idiomatis. Dalam hal fungsi, dalam hal ini berkaitan dengan bagaimana bahasa bekerja untuk mencapai tujuan komunikatif tertentu, majemuk memiliki fungsi referensial (*referential*), fungsi tekstual (*textual*), dan fungsi pendirian (*stance*). Fungsi referensial merupakan fungsi untuk menyampaikan pesan isi (*content message*), yaitu yang merujuk pada objek, fenomena, atau fakta kehidupan nyata. Fungsi ini terdiri atas fungsi lokasi, prosedur, kuantifikasi, deskripsi, pengelompokan, dan topik. Fungsi tekstual merupakan fungsi yang digunakan untuk menyusun dan menata isi (informasi referensial) suatu teks. Fungsi ini terbagi atas aditif, kontrastif, inferensial, kausatif, penstrukturan, pemingkaian, generalisasi, dan komparatif. Fungsi ini digunakan untuk menyampaikan sikap dan evaluasi penulis terhadap suatu proposisi. dan terbagi atas pendirian epistemik dan pendirian sikap.

Meskipun penggunaan pendekatan distribusional dapat membantu dalam menemukan dan menentukan majemuk dalam bahasa Indonesia, pendekatan ini masih menyisakan residu, yaitu potongan frasa dan potongan klausa.

Referensi

- [1] J. D. Parera, *Morfologi*, Jakarta: Gramedia, 2007.
- [2] S. Granger and M. Paquot, "Disentangling the phraseological web," in *Phraseology: An interdisciplinary perspective*, Amsterdam/Philadelphina, J. Benjamin, 2008, pp. 27-50.
- [3] A. P. Cowie, "The treatment of collocations and idioms in learners' dictionaries," *Applied Linguistic*, vol. 2, no. 3, pp. 223-235, 1981.
- [4] A. P. Cowie, *Phraseology: Theory, analysis, and applications*, Oxford: Oxford University Press, 1998.
- [5] J. Sinclair, *Corpus, concordance, collocation*, Oxford: Oxford Press University, 1991.
- [6] J. Sinclair, "The search for units of meaning," in *Trust the Text*, London, Routledge, 1984, pp. 34-58.
- [7] D. Biber, S. Johansson, G. Leech, S. Conrad and E. Finegan, *Longman grammar of spoken and written English*, London: Longman, 1999.
- [8] M. Stubbs, "On texts, corpora and models of language," in *Text, discourse and corpora: Theory and analysis*, London, Continuum, 2007.
- [9] D. Biber, *University language: A corpus-based study of spoken and written registers*, Amsterdam/Philadelphia: J. Benjamins, 2006.
- [10] M. Scott and C. Tribble, *Textual patterns: Key words and corpus analysis in language education*, Amsterdam: John Benjamins, 2006.
- [11] M. Stubbs, *Words and phrases: Corpus studies of lexical semantics*, Oxford: Blackwell, 2002.
- [12] S. D. Cock, *Recurrent sequences of words in native speaker and advanced learner spoken and written English: A corpus-driven approach*, Louvain: Université Catholique de Louvain, 2003.
- [13] B. Altenberg, "On the phraseology of spoken English: The evidence of reccurent word-combinations," in *Phraseology: Theory, analysis, and application*, Oxford, Clarendon Press, 1998, pp. 101-122.
- [14] H. Burger, *Phraseologie, Einfuhrung am Beispiek des Deutschen 5.*, neu bearbeitete Auflage, Berlin: Erich Schmidt Verlag, 2015.
- [15] H. Kridalaksana, *pembentukan kata dalam bahasa Indonesia*, Jakarta: Gramedia, 2007.

- [16] M. A. A. Ajie, "Phraseologisches Worterbuch Deutsch-Indonesisch am Beispiel der: Somatismen und anderer Phraseme," Dissertation, Fakultät für Philologie der Ruhr-Universität Bochum, Bochum, 2019.
- [17] T. Suhardijanto, R. Mahendra, Z. Nuriah and A. Budiwiyanto, "The framework of multiword expression in Indonesian language," in Proceedings of the 34th Pacific Asia Conference on language information and computation, Hanoi, 2020.
- [18] P. Rayson, "Computational tools and methods for corpus compilation and analysis," in The Cambridge handbook of English corpus linguistics, Cambridge, Cambridge University Press, 2015, pp. 32-49.
- [19] M. Scott, WordSmith tools version 8 (64 bit version), Stroud: Lexical Analysis Software, 2022.
- [20] A. A. Fokker, *Inleiding tot de studie van Indonesische Syntaxis*, Groningen-Jakarta: J.B. Wolters, 1951.
- [21] E. Masinambouw, *Kata majemuk: Beberapa sumbangan pikiran*, Jakarta: Fakultas Sastra Universitas Indonesia, 1980.
- [22] A. M. Moeliono, H. Lapoliwa, S. S. T. W. Sasangka and Sugiyono, *Tata Bahasa Baku Bahasa Indonesia Edisi Keempat*, Jakarta: Badan Pengembangan dan Pembinaan Bahasa, 2017.
- [23] K. Hyland, *Academic discourse: English in a global context*, London: Continuum, 2009.
- [24] K. Hyland, "As can be seen: Lexical bundles and disciplinary variation," *English for Specific Purposes*, vol. 27, no. 1, pp. 4-21, 2008.
- [25] D. Salazar, *Lexical Bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching*, Amsterdam/Philadelphia: J. Benjamins, 2014.
- [26] S. T. Gries, "Phraseology and linguistic theory: A brief survey," in *Phraseology: An interdisciplinary perspective*, Amsterdam/Philadelphia: J. Benjamin, 2008, pp. 4-26.
- [27] J. N. Sneddon, A. Adelaar, D. Jenar and M. C. Ewing, *Indonesian Reference Grammar (2nd Edition)*, New South Wales: Allen & Unwin, 2010.

PEMAJEMUKAN KATA DALAM BAHASA INDONESIA: ANCANGAN FRASEOLOGIS

¹M. Arie Andhiko Ajie dan ²Adi Budiwiyanto

¹Universitas Indonesia; ²Badan Pengembangan dan Pembinaan Bahasa

Abstrak

Kata majemuk dalam bahasa Indonesia dibahas mulai dari secara sederhana hingga yang bersifat teoretis. Ada hal-hal tertentu yang tidak konsisten dan masih menimbulkan pertanyaan, mulai dari definisi, makna, struktur hingga fungsi kata majemuk. Penelitian ini menelaah kata majemuk. Dengan berbasis korpus untuk memberikan bukti, penelitian ini mempermudah pembaca untuk memahami teori dengan contoh otentik. Korpus yang digunakan terdiri atas 29.987.513 kata (token), 445,500 kata unik (tipe), dan 17.277 teks yang berasal dari 12 ranah, yaitu koran, majalah, cerpen, novel, buku teks, jurnal, disertasi/ tesis/ skripsi, biografi, populer, perundang-undangan, surat resmi, dan laman resmi. Aplikasi yang digunakan untuk mengolah korpus adalah CQPweb v3.3.17 (<https://cqpweb.lancs.ac.uk/>). Analisis data dilakukan dengan pendekatan fraseologis. Berbeda dengan pendekatan distribusiional, pendekatan secara fraseologis divalidasi oleh korpus. Artinya, bukan frekuensi yang menjadi titik tolak pencarian kata majemuk, melainkan lebih ke kemampuan dan intuisi bahasa peneliti. Kemampuan dan intuisi bahasa ini divalidasi dengan data dari korpus. Identifikasi kata majemuk dilakukan dengan menggunakan tiga kriteria: polileksikalitas, kepaduan/keajekan dan keidiomatisan. Setelah teridentifikasi, kata majemuk diklasifikasi dan dianalisis berdasarkan struktur, makna dan fungsinya. Berdasarkan kombinasi morfem pembentuknya, kata majemuk bisa diklasifikasikan menjadi morfem bebas + morfem bebas, morfem bebas + morfem terikat, morfem terikat + morfem bebas, morfem terikat + morfem terikat dan juga kata majemuk yang terdiri atas lebih dari dua morfem. Berdasarkan maknanya, kata majemuk bisa diklasifikasikan menjadi idiom penuh, idiom sebagian dan kolokasi terbatas. Selain berdasarkan morfem pembentuknya, struktur kata majemuk bisa diklasifikasikan berdasarkan kelas kata unsur penyusunnya. Dengan klasifikasi ini, kata majemuk bisa dibagi menjadi majemuk verbal, majemuk nominal, majemuk adjektival dan majemuk preposisional. Klasifikasi kata majemuk juga bisa dibuat dengan berdasarkan kedudukan unsur pembentuknya. Unsur yang satu inti, unsur lainnya sebagai pelengkap. Klasifikasi semacam ini memunculkan istilah majemuk setara (koordinatif) dan taksetara (subkoordinatif). Berdasarkan hubungan makna antara unsur-

unsurnya, unsur-unsur kata majemuk memiliki hubungan kuantitatif, hubungan kualitatif, hubungan tujuan, hubungan perbandingan, hubungan kepemilikan dan hubungan bagian-keseluruhan. Sebagai penutup, penelitian ini meninjau fungsi kata majemuk. Beberapa fungsi kata majemuk, di antaranya adalah membentuk struktur kompleks, membentuk makna baru, membuat bentuk jamak dan menyangatkan maksud atau pernyataan.

Kata kunci: kata majemuk, fraseologi, frasem, konstruksi majemuk, bentuk majemuk

1. PENDAHULUAN

Pemajemukan merupakan salah satu proses pembentukan kata yang produktif. Dalam buku *Tata Bahasa Baku Bahasa Indonesia* edisi keempat (Moeliono dkk., 2017), pemajemukan dijelaskan berdasarkan kategori, yaitu dalam bab tentang verba, nomina, dan adjektiva. Pembahasan mengenai pemajemukan tidak dilakukan dalam satu bab khusus. Pembahasan secara terpisah itu mengakibatkan tidakimbang dan tidak konsistennya informasi yang diuraikan dalam tiap bab. Batasan umum tentang pemajemukan (juga disebut *pengompositan*) diuraikan dalam Bab IV. Di sana disebutkan bahwa pemajemukan adalah pembentukan leksem baru dengan menggabungkan dua leksem atau lebih (hlm. 117). Hasil dari pemajemukan adalah bentuk majemuk atau komposit. Kemudian, dijelaskan pula bahwa bentuk majemuk terbagi atas majemuk kata dan majemuk frasa, yang pembagiannya didasarkan atas erat-longgarnya ikatan antara unsur-unsur penyusun bentuk majemuk itu. Dalam subbagian tersebut tidak dijelaskan bagaimana mengidentifikasi sebuah bentuk majemuk/komposit.

Pengidentifikasian bentuk majemuk baru muncul pada bahasan mengenai verba majemuk: (1) makna yang dihasilkan oleh verba majemuk masih dapat dirunut dari unsur-unsurnya, (2) adanya kohesi yang kuat di antara unsur-unsurnya sehingga tidak dapat disisipi kata lain, dan (3) letak unsur-unsurnya tidak dapat dipertukarkan. Selain itu, pada akhir subbagian itu diuraikan bahwa verba majemuk berbeda dari idiom. Sementara itu, pada bagian pada pembahasan nomina majemuk (305–306), diuraikan kembali mengenai ciri bentuk majemuk dalam hal makna, yaitu masih dapat ditelusuri dari unsur-unsurnya. Namun, dalam pembahasan mengenai adjektiva majemuk, tidak diuraikan ciri-ciri adjektiva majemuk dan langsung menuju pada pembagian jenis adjektiva majemuk. Pada contoh-contoh yang diberikan, tampak bahwa

idiom merupakan bagian dari adjektiva majemuk (hlm. 224–225), misalnya *besar kepala, ringan tangan, dan tinggi hati*.

Ketakberimbangan dan ketakkonsistenan informasi tersebut mungkin saja tidak akan terjadi jika pembahasan mengenai pemajemukan dilakukan secara utuh. Oleh karena itu, Penelitian ini berusaha untuk membahas secara utuh pemajemukan dalam bahasa Indonesia yang mencakup batasan pemajemukan, ciri-ciri, struktur, makna, dan fungsi majemuk. Penelitian ini menggunakan ancangan fraseologis dan berbasis korpus. Ancangan fraseologis adalah ancangan yang melihat fraseologi sebagai satuan multikata berupa kontinum, satu sisi bersifat sangat kabur dan tetap dan di sisi lain bersifat sangat jelas dan bervariasi. dalam tradisi ini, satuan paling idiomatis, yang maknanya tidak dapat ditentukan dari unsur-unsur yang menyusun gabungan kata itu, disebut satuan paling inti dan dianggap sebagai prototipe satuan fraseologis (S. Th. Granger & M. Paquot: 2008, 27–38).

Salah satu manfaat korpus dalam mendeskripsikan suatu bahasa adalah bahwa korpus memberikan bukti dengan kualitas yang lebih baik karena korpus memungkinkan penggunaannya secara empiris memperlihatkan keteraturan pola berdasarkan kemunculan secara berulang melalui korpus (Sinclair, 2004). Korpus bukanlah sekadar kumpulan teks, melainkan kumpulan teks dari bahasa yang digunakan secara alami yang dipilih untuk menggambarkan keadaan atau variasi suatu bahasa. Korpus didesain dan disusun berdasarkan kriteria atau prinsip tertentu dan tersimpan dalam bentuk elektronik, serta dapat dibaca oleh mesin sehingga dapat dicari dengan menggunakan program komputer khusus. Hal itu penting untuk melakukan analisis kuantitatif yang berkaitan dengan linguistik korpus, seperti mengekstrak daftar frekuensi kata, konkordansi, daftar kata kunci, daftar kolokat, dan uji statistis. Selain itu, analisis kualitatif juga dapat dilakukan ketika peneliti mengakses konteks yang lebih banyak atau keseluruhan teks yang ada pada korpus yang di dalamnya terdapat beragam penggunaan bahasa.

2. TINJAUAN PUSTAKA

Pembahasan tentang kata majemuk dalam bahasa Indonesia merupakan pembahasan yang menarik bagi para linguis Indonesia, baik berkenaan dengan istilah yang digunakan maupun konsep kata majemuk itu sendiri. Dari segi istilah, kata majemuk telah lama digunakan dan hingga kini masih

digunakan oleh linguis Indonesia, seperti Mees (1954), Alisjahbana (1963), Slametmuljana (1969), Ramlan (2012), dan Simpen (2021). Parera (2007) dan Moeliono dkk. (2017) tidak menggunakan istilah kata majemuk, tetapi bentuk majemuk, karena proses pemajemukan itu melibatkan dua kata atau lebih sehingga secara struktur tidak tepat lagi menggunakan istilah *kata*. Sementara itu, Chaer (2015) menggunakan istilah komposisi, baik itu untuk proses dan hasilnya, sedangkan Kridalaksana (2007) menggunakan istilah kompositum atau paduan leksem yang menjadi calon kata majemuk.

Mees (1954: 70—71) mendefinisikan kata majemuk atau persenyawaan sebagai “dua buah perkataan atau tiga yang kalau diperhubungkan, menciptakan satu pengertian. Mees membagi kata majemuk menjadi dua, yaitu (1) persenyawaan kopulatif atau gabung (yang setara), misalnya. *ibu bapak, tanah air, dan hutan rimba belantara*; (2) persenyawaan determinatif (yang menentukan), yang terbagi ke dalam tiga kelompok: (a) yang mengandung perhubungan kasus, misalnya *ibu jari* dan *air mata*; (b) yang menjelaskan, misalnya *bunga mawar, ikan mas, dan sungai Musi*; dan (c) yang menyifatkan, misalnya *orang tua* dan *tanah lapang*; dan (3) persenyawaan posesif (yang mengandung arti kepunyaan), misalnya *mulut manis* dan *kepala botak*. Penulisan kata majemuk mengikuti ejaan Van Ophuysen, yaitu ditulis terpisah dan tanpa tanda hubung. Van Ophuysen sendiri dikenal sebagai orang yang pertama meneliti kata majemuk dalam bahasa Indonesia.

Mees juga menggunakan istilah aneksi untuk menyebutkan kata majemuk atau persenyawaan. khususnya yang memiliki hubungan yang sangat erat antar unsur-unsurnya sehingga memiliki makna baru, misalnya *saputangan* dan *keretaapi*. Aneksi adalah hubungan kata-kata yang tidak boleh dipisahkan atau dibalikkan susunannya, misalnya *lukisan Yusuf* ‘lukisan milik yusuf’ dan *nasihat ayah* ‘nasihat yang diberikan ayah’. Jadi, kata majemuk, menurut Mees, memiliki tiga ciri, yaitu (1) maknanya dapat bersifat komposisional atau nonkomposisional; (2) tidak dapat disisipi atau diperluas; dan (3) tidak dapat dibolak-balik.

Alisjahbana (1963) mendefinisikan kata majemuk sebagai kata gabungan dua kata yang maknanya tidak dapat ditelusuri dari kata penyusunnya, misalnya *rumahsakit, keretaapi, mesinterbang*. Kata majemuk ditulis serangkai dan jika direduklifikasi, keseluruhan kata mengalami reduklifikasi. Alisjahbana mengelompokkan kata majemuk menjadi tiga, yaitu (1) kata mejemuk berpola DM (unsur kedua menerangkan unsur pertama), misalnya *mataair* dan *rumahmakan, mabuklaut; keraskepala, panjangtangan, besar*

hati; (2) kata majemuk berpola DM yang dapat disisipi ‘dan’, yang terbagi atas (a). berlawanan, misalnya *lakibini* dan *siangmalam*; (b) mengeraskan, misalnya *hinadina* dan *sanaksaudara*; (3) kata majemuk berpola MD, misalnya *maharaja* dan *purbakala*.

Ramlan (2012) mendefinisikan kata majemuk sebagai kata yang terjadi dari dua kata sebagai unsurnya, misalnya *daya tahan*, *kamar tunggu*, *ruang baca*, *simpan pinjam*, dan *jual beli*. Ramlan mencirikan kata majemuk sebagai berikut. Pertama, salah satu atau semua unsurnya berupa pokok kata (satuan gramatik yang tidak dapat berdiri sendiri dalam tuturan biasa dan secara gramatik tidak memiliki sifat bebas), misalnya *kolam renang*, *pasukan tempur*, dan *medan tempur*. Kedua, unsur-unsurnya tidak dapat dipisahkan atau strukturnya tidak bisa diubah, misalnya *kaki tangan* tidak dapat menjadi **kaki dan tangan*. Ketiga, kata majemuk dengan unsur berupa morfem unik hanya bisa berpasangan dengan kata tertentu, misalnya, *gelap gulita*, *terang benderang*, dan *sunyi senyap*.

Sneddon (2010) mendefinisikan kata majemuk (*compound*) sebagai “a combination of two simple words which come together to form a complex word.” Gabungan dua kata simpleks yang membentuk sebuah kata kompleks tersebut ia contohkan dengan kata *tanda* dan kata *tangan* yang membentuk kata kompleks *tanda tangan* yang memiliki makna baru. Secara penulisan, Sneddon menyebutkan bahwa ada tiga cara penulisan, yaitu serangkai, dengan tanda hubung, dan terpisah. Kata majemuk dapat berupa nomina, adjektiva, dan verba. Secara makna, kata majemuk dapat bersifat non-komposisional maupun komposisional.

Sippen (2021) mendefinisikan pemajemukan sebagai proses pembentukan kata yang dilakukan dengan cara menggabungkan satu bentuk (bebas atau terikat) dengan satu bentuk (bebas atau terikat) yang lain sehingga menghasilkan kata majemuk. Makna setiap unsurnya lebur dan membentuk pengertian baru. Sippen membagi kata majemuk berdasarkan beberapa segi: (1) bentuk dasar yang membangunnya (morfem bebas+morfem bebas, morfem bebas+morfem terikat, morfem bebas+morfem unik, morfem terikat+morfem bebas); (2) kategori unsurnya (VV, VN, NN, NV, AA, NA, NNum); (3) hubungan antarunsurnya, yaitu setara dan taksetara; serta (4) susunannya, yaitu DM dan MD.

Kridalaksana (2007) mendefinisikan perpaduan/komposisi sebagai proses penggabungan dua leksem atau lebih yang membentuk kata. Keluaran proses perpaduan itu adalah paduan leksem/kompositum yang menjadi

calon kata majemuk. Kridalaksana menyebutkan tiga ciri paduan leksem/kata majemuk untuk dibedakan dari frasa: ketaktersisipan, ketakterluasan, dan ketakterbalikan. Namun, ia juga melihat bahwa konsep kompositum tidak sama benar dengan konsep kata majemuk. Hal itu dapat diuji dengan kompositum *satu padu* dan *bumi hangus*. Bentuk tersebut belum berstatus sebagai kata karena tidak dapat berdiri sendiri jika tidak mengalami proses afiksasi. Kridalaksana membagi kompositum ke dalam lima kelompok: (1) kompositum subordinatif substantif, (2) kompositum subordinatif atributif, (3) kompositum koordinatif, (4) kompositum berproleksem, dan (5) kompositum sintetis.

Chaer (2015) mendefinisikan komposisi sebagai “proses penggabungan dasar dengan dasar (biasanya berupa akar maupun bentuk berimbuhan) untuk mewadahi suatu konsep yang belum tertampung dalam sebuah kata.” Istilah *komposisi* yang digunakan Chaer sebagai proses ternyata digunakan juga sebagai hasil. Hal itu tampaknya dapat menimbulkan ketaksaan. Chaer membagi komposisi ke dalam lima kelompok berdasarkan aspek semantik: (1) komposisi koordinatif, misalnya *baca tulis* dan *tua muda*; (2) komposisi subordinatif, misalnya *sate ayam*; (3) komposisi penghasil istilah, misalnya *guru bantu* dan *buku ajar*, (4) komposisi pembentuk idiom, misalnya *memeras keringat* dan *meja hijau*; serta (5) komposisi penghasil nama, misalnya *Stasiun Gambir* dan *Selat Sunda*. Sementara itu, berdasarkan kategori, Chaer membagi komposisi ke dalam (1) komposisi nominal: bermakna gramatikal, bermakna idiomatis, bermakna metaforis, nama dan istilah, dan komposisi dengan adverbial; (2) komposisi verbal: bermakna gramatikal, bermakna idiomatis, komposisi dengan adverbial; dan (3) komposisi ajektival: bermakna gramatikal, bermakna idiomatis, komposisi dengan adverbial. dalam pembagian kategori di atas tampaknya ketidaksetaraan, khususnya di bagian ‘komposisi dengan adverbial’ karena subkategori yang lain berada pada tataran makna.

Sementara itu, Parera (2007) memberi batasan yang lebih sempit. Menurut Parera, bentuk majemuk itu adalah pasangan yang terikat, yaitu dua bentuk bahasa yang secara khusus, terbatas, dan tetap berpasangan dalam keseluruhan pelaksanaan bahasa. Ia tidak mempunyai kemungkinan untuk berpasangan dengan bentuk atau kata lain. Bentuk penghematan atau penghilangan partikel lain juga bukan bentuk majemuk.

3. METODE

Penelitian ini menggunakan metode kualitatif dengan ancangan berbasis korpus. Korpus yang digunakan terdiri atas 29.987.513 kata (token), 445,500 kata unik (tipe), dan 17.277 teks yang berasal dari 12 ranah, yaitu koran, majalah, cerpen, novel, buku teks, jurnal, disertasi/ tesis/ skripsi, biografi, populer, perundang-undangan, surat resmi, dan laman resmi. Aplikasi yang digunakan untuk mengolah korpus adalah CQPweb v3.3.17 (<https://cqpweb.lancs.ac.uk/>).

Pengumpulan data dilakukan dengan dua cara: pertama, dengan mengumpulkan kata majemuk potensial secara intuitif kemudian mengecek penggunaannya di dalam korpus CQPweb; kedua, dengan mengekstrak data n-gram dengan menggunakan WordSmith 8.0. Istilah n-gram biasa digunakan dalam bidang linguistik komputasional. Metode n-gram ini menggunakan program komputer untuk menghitung frasa atau rangkaian kata sinambung yang muncul dalam data korpus. beragam panjang n-gram yang dapat dihitung secara terpisah (Paul Rayson: 41). Dalam penelitian ini, n-gram diekstrak adalah 2-4 kata dengan menggunakan pengolah korpus WordSmith 8.0.

Analisis data dilakukan dengan pendekatan fraseologis. Identifikasi kata majemuk dilakukan dengan menggunakan tiga kriteria: polileksikalitas, kepaduan/keajekan, dan keidiomatisan. Setelah teridentifikasi, kata majemuk diklasifikasi dan dianalisis berdasarkan struktur, makna, dan fungsinya.

4. HASIL DAN PEMBAHASAN

Hingga saat ini, banyak tulisan yang membahas kata majemuk dari berbagai kriteria dan memberikan contohnya. Ketika ditelaah lebih dalam dengan berdasarkan pendapat berbagai ahli linguistik, kata majemuk tidak sesederhana yang dipaparkan di buku-buku pelajaran bahasa Indonesia di sekolah. Untuk membantu menelaah kata majemuk atau konstruksi majemuk, bidang fraseologi yang menelaah frasem memiliki potensi untuk menelaah kata atau konstruksi majemuk. Ini dikarenakan adanya kemiripan dan adanya sebagian kesamaan antara frasem dengan kata atau konstruksi majemuk.

Kata majemuk atau konstruksi majemuk memiliki kemiripan dengan frasem karena memenuhi dua kriteria, yakni polileksikalitas, kepaduan/keajekan dan sebagian memenuhi kriteria ke-tiga, yakni keidiomatisan

(Burger, 2015).

Frasem berbeda kelas katanya, struktur sintaksisnya, secara pragmatis pun penggunaannya terikat konteks komunikasinya. Frasem terdiri dari minimal dua kata (polileksikalitas). Kemunculannya sebagai rangkaian kata yang satu bukan hanya satu atau beberapa kali saja, kesatuan rangkaian kata ini digunakan oleh pengguna bahasa Indonesia layaknya sebuah kata atau kesatuan (memenuhi kriteria keajekan). Penutur jati bisa menggunakan frasem tanpa banyak berpikir ulang. Mereka tidak perlu merangkaikan lagi kata per kata untuk bisa menggunakan frasem itu. Mereka memahami frasem itu, menyimpannya di dalam ingatan dan menggunakannya seperti rangkaian yang satu, layaknya sebuah kata (keajekan psikolinguistik). Konteks komunikasi digunakannya frasem ini pun terbatas.

Berdasarkan penjabaran singkat di atas, ada kesamaan frasem dengan kata majemuk. Perbedaannya adalah, frasem lebih luas. Struktur frasem mencakup dari minimal dua kata hingga sebuah kalimat utuh, peribahasa misalnya. Peribahasa merupakan salah satu jenis frasem yang secara sintaksis berbentuk kalimat. Jika frasem ada yang bisa berbentuk kalimat, maka kata majemuk tetap harus berstatus kata. Dan hal inilah yang jadi keambiguan dengan penggunaan istilah "kata majemuk". Di pembahasan, didapatkan contoh kata majemuk yang terdiri atas lebih dari dua kata atau dua morfem bebas. Hal inilah yang membuat istilah "kata majemuk" beralih menjadi "konstruksi majemuk". *Anak semata wayang, di bawah tangan dan oleh karena itu* tentu akan terlihat aneh kalau masih disebut "kata majemuk".

Tulisan ini dibuat untuk memperkaya diskusi tentang kata atau konstruksi majemuk dengan mendeskripsikan jenisnya (berdasarkan morfem, berdasarkan makna, berdasarkan fungsi, berdasarkan kelas kata, berdasarkan relasi kedudukan unsur pembentuknya) dan fungsinya dengan menggunakan korpus. Dalam pemilihan contoh, kriteria fraseologis menjadi pertimbangan utama. Ancangan fraseologis ini berbeda dengan ancangan distribusional karena berangkat dari kemampuan dan intuisi bahasa peneliti. Langkah ini penting karena frekuensi kata majemuk atau frasem berbentuk kata majemuk dalam bahasa Indonesia terhitung rendah frekuensinya. Supaya penelitian lebih efektif, korpus digunakan untuk memvalidasi hipotesis peneliti. Hasil yang disampaikan di sini merupakan kata majemuk atau konstruksi majemuk yang berangkat dari hipotesis peneliti dan terbukti ada di dalam korpus.

4.1 Jenis Kata Majemuk

4.1.1 Berdasarkan berdasarkan morfem

Kata majemuk terbentuk dengan berbagai kombinasi morfem, baik morfem terikat, maupun morfem bebas. Berikut ini akan ditampilkan jenis-jenis kata majemuk berdasarkan morfem yang membentuknya dan urutannya dalam kata majemuk itu.

4.1.1.1 Morfem bebas + morfem bebas

Kombinasi komponen kata majemuk yang berupa gabungan dua morfem bebas adalah bentuk yang paling umum dijumpai. Contoh kombinasi kata majemuk yang berupa morfem bebas + morfem bebas adalah *tanggung jawab, kereta api, angkat kaki, rumah sakit, murah meriah, basah kuyup* dan sebagainya.

Kombinasi ini merupakan kombinasi yang paling umum. Hal ini bisa dikenali dengan membuka kamus-kamus ungkapan yang banyak mengandung idiom, baik idiom penuh, maupun idiom sebagian. Batasan antara kata majemuk dengan idiom pun kadang kurang tegas karena banyak idiom yang dikatakan merupakan kata majemuk. Kata majemuk adalah sebutan di tataran morfologis, sedangkan idiom di tataran semantis. Karena banyaknya kata majemuk yang bersifat idiomatis, maka banyak ada yang mendefinisikan bahwa kata majemuk memiliki makna baru, padahal tidak semuanya bersifat idiomatis, misalnya *salah sangka* dan *jumpa pers*.

Contoh idiom penuh yang memiliki konstruksi morfem bebas + morfem bebas:

cuci tangan, angkat tangan, tangan kanan, tutup mata, besar kepala, besar hati, angkat kaki, tutup mulut, naik daun, naik darah

Contoh idiom sebagian yang memiliki konstruksi morfem bebas + morfem bebas:

uang suap, uang pelicin, daftar hitam, penumpang gelap, kabar burung

Contoh kolokasi (karena tidak idiomatis atau rendah derajat keidiomatisannya) yang memiliki konstruksi morfem bebas + morfem bebas:

*balas budi, kereta api, kipas angin, wali murid, gagah berani,
setia kawan, pasar uang, pasar valuta asing, toko swalayan*

4.1.1.2 Morfem bebas + morfem terikat

Tidak hanya kombinasi morfem bebas + morfem bebas, kombinasi morfem bebas + morfem terikat juga didapati di korpus. Contoh kombinasi ini adalah *Sukuisme* (*suku* merupakan morfem bebas dan *-isme* merupakan morfem terikat).

Contoh kata majemuk yang komponen ke-duanya morfem terikat adalah:

daya juang, tanpa henti, modernisasi, santriwati, duniawi
daya (morfem bebas) + *juang* (morfem terikat)
tanpa (morfem bebas) + *henti* (morfem terikat)
modern (morfem bebas) + *-isasi* (morfem terikat)
santri (morfem bebas) + *-wati* (morfem terikat)
dunia (morfem bebas) + *-wi* (morfem terikat)

Selain menunjukkan contoh kata majemuk yang salah satu komponennya merupakan morfem terikat, bisa juga dilihat bahwa morfem terikat ada yang berbentuk kata, bukan hanya sufiks. Morfem terikat semacam ini contohnya *juang* dan *henti*. Meskipun berbentuk kata, *juang* dan *henti* tidak bisa berdiri sendiri tanpa adanya kombinasi dengan kata atau morfem lain untuk bisa dipergunakan. Berikut ini contoh penggunaannya: *pejuang, perjuangan, perjuangan, berjuang, diperjuangkan, berhenti, hentikan, perhentian* dan sebagainya.

4.1.1.3 Morfem terikat + morfem bebas

Kata majemuk yang komponen pembentuknya diawali dengan morfem terikat bisa ditemukan di dalam korpus. *Pascalahir* (*pasca-* merupakan morfem terikat, sedangkan *lahir* merupakan morfem bebas), *metafisika* (*meta-* merupakan morfem terikat, sedangkan *fisika* merupakan morfem bebas), *pramuwisata* (*pramu-* merupakan morfem terikat, sedangkan *wisata* merupakan morfem bebas).

Berikut ini adalah contoh kata majemuk lainnya yang merupakan gabungan morfem terikat dengan morfem bebas:

pasca- : *pascabayar, pascakrisis, pascaproduksi*
meta- : *metadata, metakognitif, metaanalisis*
swa- : *swalayan, swadaya, swafoto*
maha- : *mahasiswa, mahakuasa, mahatahu*
pra- : *prabayar, pramusim, praaksara*
mal- : *malpraktik, maladministrasi*
dwi- : *dwibahasa, dwitunggal*
tri- : *tridarma*
panca- : *pancasila, pancaindera, pancayadnya, pancaroba*
dasa- : *dasasila, dasawarsa*

4.1.1.4 Morfem terikat + morfem terikat

Konstruksi kata majemuk tidak memprasyaratkan adanya morfem bebas sebagai salah satu komponen. Kata majemuk bisa terbentuk dari gabungan morfem terikat semata. Contohnya adalah sebagai berikut:

teologi : *teo-* (morfem terikat) + *-logi* (morfem terikat)
teolog : *teo-* (morfem terikat) + *-log* (morfem terikat)
biologi : *bio-* (morfem terikat) + *-logi* (morfem terikat)
biografi : *bio-* (morfem terikat) + *-grafi* (morfem terikat)

Contoh lainnya adalah: *geologi, geolog, geografi, topologi, topografi*

Pada contoh di atas, *bio* merupakan morfem terikat yang dituliskan tersambung dengan morfem terikat lainnya. Di korpus, morfem *bio* bisa ditemukan berdiri sendiri tanpa ditempelkan dengan morfem lainnya, misalnya di kata *gemuk bio* dan *pelumas bio*. Meskipun ditulis terpisah, belum bisa dikatakan morfem *bio* bisa berdiri sendiri sehingga bisa disebut sebagai morfem bebas karena pemakaiannya terbatas dan frekuensinya sedikit.

4.1.1.5 Kata majemuk dengan lebih dari dua morfem bebas

Pada umumnya, kata majemuk yang sering dijadikan contoh hampir semuanya terdiri atas dua kata. Kedua kata ini biasanya dua morfem bebas. Namun, ternyata ada beberapa kata majemuk yang terdiri atas tiga kata atau tiga morfem bebas, yakni *dengan langkah gontai, penjahat kelas kakap, penjahat kelas teri, di bawah tangan, anak semata wayang, anak kemarin sore*

Contoh tersebut bisa dianggap sebagai kata majemuk dengan tiga morfem bebas. Kridalaksana et. al. (1985: 147) menyampaikan, kata majemuk, seperti istilahnya, haruslah tetap berstatus kata dan tidak sama dengan frase.

Kridalaksana (2007) merumuskan tiga kriteria yang bisa membedakan kata majemuk dengan bukan kata majemuk. Kata majemuk harus memenuhi kriteria ketidaktersipan, ketakterluasan dan ketakterbalikan. Yang dimaksudkan dengan ketidaktersipan adalah, kata majemuk tidak bisa disisipi apapun di antara komponen pembentuknya. Dari semua contoh tersebut, hanya *dengan langkah gontai* yang bisa disisipi dan sisipannya ditemukan di dalam korpus, yakni disisipi kata *agak* dan *sedikit* sehingga menjadi *dengan langkah agak gontai* dan *dengan langkah sedikit gontai*. Namun, di dalam korpus hanya ada sebuah contoh untuk masing-masing kata. Kriteria lainnya adalah ketakterluasan (komponen kata majemuk tidak dapat diafiksasi dan dimodifikasi, kecuali keseluruhan) dan ketakterbalikan (urutan komponen pembentuk kata majemuk tidak dapat dibalik).

4.1.2 Berdasarkan makna

Kata majemuk bisa diklasifikasikan secara semantis menjadi tiga, yakni kata majemuk idiomatis penuh, kata majemuk idiomatis sebagian dan kata majemuk kolokasi terbatas. Pembagian ini serupa dengan pembagian rangkaian kata ajek/ gugus kata/frasem yang secara semantis bisa dibagi menjadi tiga macam, yakni idiom penuh, idiom sebagian dan kolokasi. Di bagian ini akan dijelaskan mengenai tiga macam kata majemuk berdasarkan makna berikut dengan contoh yang diambil dari korpus.

4.1.2.1 Majemuk idiomatis penuh

Kata majemuk banyak yang memiliki makna idiomatis, bahkan makna idiomatis penuh. Hal ini membuat batasan antara kata majemuk dan idiom tidak mudah dikenali. Contoh-contoh yang sering diberikan untuk menjelaskan kata majemuk berupa idiom. Kridalaksana et. al. (1985: 147) menegaskan:

“Konsep kata majemuk harus dibedakan dari idiom dan semi-idiom. Kata majemuk adalah konsep sintaktis, sedangkan idiom adalah konsep semantik. Yang dimaksud dengan idiom ialah konstruksi yang maknanya tidak sama dengan gabungan makna anggota-anggotanya. Konstruksi semacam itu bisa berupa kata, seperti *pribumi*, frase seperti *kambing hitam*, klausa seperti *nona makan sirih*, kata berulang seperti *mata-mata*.”

Dari penjelasan ini bisa dipahami bahwa yang membedakan adalah konsepnya. Selain itu, idiom bisa terdiri atas lebih dari dua kata atau dua morfem bebas, sedangkan kata majemuk tidak lumrah memiliki konstruksi seperti itu. Berikut ini adalah contoh kata majemuk idiomatis penuh, yakni yang kesemua komponen pembentuknya tidak lagi mempertahankan makna leksikalnya:

<i>terus terang</i>	: mengatakan/menuliskan sesuatu dengan jujur
<i>besar kepala</i>	: sombong
<i>buah bibir</i>	: menjadi bahan pembicaraan
<i>buah tangan</i>	: oleh-oleh
<i>mata keranjang</i>	: sebutan untuk laki-laki yang suka memandangi perempuan lain yang bukan pasangannya
<i>naik daun</i>	: menjadi populer
<i>tangan kanan</i>	: orang kepercayaan
<i>kaki tangan</i>	: orang yang bekerja untuk orang lain dalam artian negatif
<i>patah hati</i>	: perasaan seseorang ketika cintanya tidak disambut dengan baik atau mengalami penolakan
<i>turun gunung</i>	: ketika seseorang yang dianggap memiliki ilmu atau kemampuan kembali ke tengah komunitas untuk mengurus sesuatu dengan kemampuannya itu
<i>di bawah tangan</i>	: secara tidak resmi
<i>kambing hitam</i>	: orang atau hal yang dipersalahkan atas suatu kejadian
<i>anak bawang</i>	: orang yang dianggap belum memiliki kemampuan matang sehingga tidak diberikan tanggung jawab penuh
<i>angkat tangan</i>	: menyerah
<i>turun tangan</i>	: ikut serta untuk melakukan atau menyelesaikan sesuatu
<i>angkat kaki</i>	: pergi meninggalkan suatu tempat, biasanya karena ada alasan yang tidak menyenangkan
<i>meja hijau</i>	: pengadilan
<i>buah hati</i>	: anak

Dari penjabaran kata majemuk dan makna idiomatisnya, tidak ada satupun contoh kata majemuk yang makna leksikal komponen pembentuknya mempertahankan makna leksikalnya. Oleh karena itu, kumpulan contoh kata majemuk di atas tergolong kata majemuk idiomatis penuh. Meskipun bersifat idiomatis penuh, makna kata majemuk tersebut mudah dipahami oleh penutur jati bahasa Indonesia karena frekuensi penggunaan kata majemuk tersebut

di berbagai media cukup sering. Ini adalah bukti bahwa kata majemuk disimpan di dalam otak manusia sebagai kesatuan. Pengguna bahasa tidak perlu lagi mengingat atau memikirkan makna leksikal komponen penyusun kata majemuk satu persatu. Ini berbeda bagi orang yang mempelajari bahasa Indonesia sebagai bahasa asing. Karena masih minimnya kontak dengan bahasa Indonesia, perbendaharaan idiom mereka masih sedikit, sehingga ada kecenderungan untuk memahami makna kata majemuk idiomatis penuh dengan tidak tepat karena memahami berdasarkan makna satu persatu komponennya.

4.1.2.2 Majemuk idiomatis sebagian

Berbeda dengan kata majemuk idiomatis penuh, kata majemuk idiomatis sebagian memiliki komponen yang masih mempertahankan makna leksikalnya. Secara semantis, kata majemuk idiomatis sebagian dikenal dengan idiom sebagian. *Rumah sakit* dan *penumpang gelap* merupakan contoh kata majemuk idiomatis sebagian. Kata *rumah* mempertahankan makna leksikalnya (meskipun rumah sakit besar tidak lagi berbentuk rumah, tetapi lebih seperti gedung), sedangkan kata *sakit* tidak menerangkan kata rumah (bandingkan misalnya dengan *orang sakit* dan *hewan sakit*). Hal yang sama juga bisa diamati pada contoh berikut ini:

- penumpang gelap* : **penumpang** yang tidak membayar atau ikut serta dengan cara yang benar
- daftar hitam* : daftar nama orang atau organisasi yang dianggap membahayakan keamanan atau daftar nama orang yang pernah dihukum karena melakukan kejahatan
- anak angkat* : seseorang yang diadopsi menjadi **anak**, meskipun sudah dewasa, statusnya akan tetap bisa disebut anak angkat
- kabar burung* : **kabar** yang belum jelas kebenarannya, **kabar** yang masih berupa isu
- lapangan hijau* : **lapangan** yang digunakan untuk permainan sepak bola. Meskipun lapangan itu hanya tanah merah tanpa rumput, tetap bisa disebut *lapangan hijau*. Sebaliknya, lapangan tenis yang dicat dengan warna hijau, tidak disebut *lapangan hijau*

masih hijau : sebutan yang ditunjukkan untuk seseorang yang dianggap **masih** muda atau belum banyak pengalaman
buah pikiran : hasil **pikiran**

4.1.2.3 Majemuk kolokasi terbatas

Kata majemuk yang memiliki komponen yang mempertahankan makna leksikalnya atau kata majemuk yang memiliki makna tidak idiomatis atau rendah derajat keidiomatisannya disebut kata majemuk kolokasi terbatas. Contohnya bisa dilihat pada kata-kata majemuk berikut ini:

ikut campur, ambil keputusan, jalan santai, gerak jalan, lari cepat, dengan langkah gontai

Karena tidak idiomatis atau rendahnya derajat keidiomatisan kata majemuk jenis ini, orang yang bukan penutur jati bahasa Indonesia bisa memahami maknanya dengan mengartikan makna kata majemuk kolokasi terbatas berdasarkan makna leksikal komponen pembentuknya. Dalam perspektif pengajaran bahasa asing, kata majemuk kolokasi terbatas memiliki tingkat kesulitan yang rendah untuk bisa dipahami dibandingkan kata majemuk idiomatis penuh dan kata majemuk idiomatis sebagian.

4.1.3 Struktur kata majemuk berdasarkan kelas kata

Kata majemuk dapat dikelompokkan berdasarkan kelas kata unsur penyusunnya: majemuk verbal, majemuk nominal, majemuk adjektival, dan majemuk preposisional. Berikut adalah pemaparan untuk setiap kategori.

4.1.3.1 Majemuk verbal

Majemuk verbal adalah kata majemuk yang berintikan verba, misalnya *gulung tikar, hancur lebur, hilang lenyap, terjun payung, jatuh bangun*. Kata *gulung, hancur, hilang, terjun*, dan *naik* merupakan inti dan unsur lainnya yang mengikuti dapat bersifat subordinatif (*tikar, tangan*), atributif (*lebur, lenyap*), dan koordinatif (*turun*). Perhatikan pemakaian majemuk verbal pada contoh berikut.

- Perusahaan itu *gulung tikar* akibat krisis ekonomi.
- Bangunan dua lantai itu *hancur lebur*.
- Rasa hampa yang kualami selama ini *hilang lenyap*, tergantikan dengan

rasa bahagia.

- d. Di Pekanbaru dia *terjun payung* sebelum berlatih terjun.
- e. Para pejalan kaki harus *naik turun* tangga penyeberangan.
Meskipun inti kata majemuk berupa verba, dalam pemakaian tertentu kata majemuk verbal dapat berfungsi sebagai nomina seperti pada contoh berikut.
- f. Nanti setelah pemeriksaan besok, saya akan *jumpa pers*.
- g. Rencananya mereka akan menggelar *jumpa pers* di kantor pusat untuk menyatakan sikap.

Pada kalimat (e) *jumpa pers* berfungsi sebagai verba, sedangkan pada kalimat (f) berfungsi sebagai nomina.

Sementara itu, terdapat majemuk verbal yang intinya berupa verba, tetapi majemuk tersebut berfungsi sebagai nomina. Majemuk verbal tersebut dapat berfungsi sebagai verb jika kata majemuk itu mendapatkan afiksasi. Perhatikan contoh berikut.

- h. Pengumpulan data dilakukan dengan wawancara telepon atau dengan *tatap muka*.
- i. Mereka cepat akrab meskipun baru pertama kali ini *bertatap muka*.
Dari segi bentuknya, majemuk verbal dapat dibagi menjadi tiga kelompok, yaitu majemuk verbal dasar, majemuk verbal berafiks, dan majemuk verbal berulang. Penjelasan ketiga bentuk tersebut akan dipaparkan berikut ini.

(1) Majemuk Verbal Dasar

Majemuk verbal dasar adalah kata majemuk yang unsur-unsur penyusunnya berupa verba yang tidak berafiks dan tidak mengandung komponen berulang. Perhatikan contoh berikut.

- a. Ia mengaku tak *ikut campur* dalam keputusan Presiden memilih penggantinya.
- b. Mereka hanya bilang bahwa saya *gegar otak*.
- c. Aku menatap Kadek, memastikan dia tidak sedang *mabuk laut*.
- d. Kiper itu harus *jatuh bangun* menyelamatkan gawangnya.
- e. Dia sendiri *pulang pergi* naik kendaraan umum.

(2) Majemuk Verbal Berafiks

Majemuk verbal berafiks adalah verba majemuk yang mengandung afiks tertentu. Afiks tersebut dapat melekat pada inti atau mengapit kata majemuk tersebut, seperti yang terdapat dalam kalimat berikut.

- a. Pasal itu menyebutkan KPK dapat *mengambil alih* kasus korupsi yang ditangani kejaksaan.
- b. Aku hanya bisa *berserah diri* kepada Allah.
- c. Saat itulah ia *tertangkap basah* sedang membuat sabu-sabu bersama pegawainya.
- d. Ia tak pernah mengeluh meski harus bekerja *membanting tulang* seperti kuda beban.
- e. Penelitian ini *mengikutsertakan* sebagian relawan masyarakat kampung.
- f. Mohon bantuan Bapak/Ibu untuk dapat *menyebarkan* informasi tersebut kepada mahasiswa.
- g. Kapal ini diharapkan dapat segera *diserahterimakan*.

(3) Majemuk Verbal Berulang

Majemuk verbal dalam bahasa Indonesia dapat direduksi, baik verbal dasar maupun verbal berimbuhan. Pengulangan dilakukan pada inti kata majemuk tersebut, bukan keseluruhan. Perhatikan contoh berikut.

- a. Saya *geleng-geleng kepala* ketika baca drafnya.
- b. Mereka *tidur-tidur ayam* di siang yang terik.
- c. Aku tak mau *berpanjang-panjang kata* soal itu.
- d. Mobil yang *berpindah-pindah tangan* juga bisa mempersingkat umur transmisi.
- e. Kami *berlari-lari kecil*, menyusuri jalan setapak yang sedikit licin.

4.1.3.2 Majemuk Nominal

Majemuk nominal adalah kata majemuk yang berintikan nomina, misalnya *uang muka*, *tingkah laku*, *adat istiadat*, *utang budi*, dan *ibu kota*. Kata *uang*, *tingkah*, *adat*, *utang*, dan *ibu* masing-masing menjadi inti pada kata majemuk tersebut. Sementara itu, unsur lainnya dapat diisi dengan subordinatif (*budi*, *kota*), atributif (*muka*), dan koordinatif (*laku*, *istiadat*). Berikut adalah contoh penggunaan majemuk nominal tersebut.

- a. Ketika mendapat harga murah, ia langsung membayar *uang muka*.
- b. Ia menyaksikan *tingkah laku* ayahnya sambil berusaha keras menahan tawa.
- c. Melalui antropologi ia belajar tentang *adat istiadat* dan pola pikir sebuah masyarakat.

- d. Politik etis dimaksudkan untuk membayar *utang budi* Belanda kepada Indonesia.
- e. Kita seperti kian tak berdaya dihadapkan pada kemacetan *ibu kota*.

Berdasarkan bentuk morfologisnya, majemuk nominal terbagi atas (1) majemuk nominal dasar, (2) majemuk nominal berafiks, (3) majemuk nominal dengan bentuk terikat, dan (4) majemuk nominal sintetis.

(1) **Majemuk Nominal Dasar**

Majemuk nomina dasar adalah kata majemuk yang unsur intinya berupa nomina dasar, misalnya *rumah tangga*, *kurun waktu*, *sanak saudara*, *kepala dingin*, dan *jalan raya*. Unsur yang melengkapi unsur inti dapat bersifat subordinatif (*tangga*), atributif (*dingin*, *raya*), dan koordinatif (*waktu*, *saudara*). Perhatikan contoh pemakaian berikut.

- a. Korban pemadaman ini bukan hanya *rumah tangga*, tetapi juga pelanggan industri.
- b. Berat badannya memang sempat turun, tetapi kemudian kembali naik dalam *kurun waktu* singkat.
- c. Ketika hijrah, mereka meninggalkan *sanak saudara*, harta benda, dan kampung halaman.
- d. Konflik tidak akan berlarut-larut andaikata urusan diselesaikan dengan *kepala dingin*.
- e. Sistem peringatan dini terbukti mampu memotong jumlah kecelakaan di *jalan raya*.

(2) **Majemuk Nominal Berafiks**

Majemuk nominal berimbuhan adalah kata majemuk yang berintikan nomina dan salah satu atau kedua unsur penyusunnya memiliki afiks, misalnya *penegak hukum*, *perguruan tinggi*, *mata pelajaran*, *surat berharga*, dan *perusahaan penjaminan*. Kata *penegak*, *perguruan*, *pelajaran*, *berharga*, serta perusahaan dan penjaminan merupakan unsur berafiks. Berikut adalah contoh pemakaian majemuk nominal berafiks.

- a. Lembaga *penegak hukum* semestinya semakin terpacu meningkatkan kinerja.
- b. Alih-alih duduk manis di bangku *perguruan tinggi*, ia malah mengasah kepalan di sasana tinju.
- c. Perubahan kurikulum pendidikan itu akan menyederhanakan sejumlah *mata pelajaran*.

- d. Tujuan pembelian *surat berharga* ini adalah memengaruhi likuiditas di pasar uang.
- e. Keberadaan *perusahaan penjaminan* ini akan membantu dan memudahkan UMKM dalam mengakses pembiayaan.

(3) Majemuk Nominal dengan Bentuk Terikat

Majemuk nominal dengan bentuk terikat ini terdiri atas dua unsur dan salah satunya adalah unsur terikat, yaitu unsur yang tidak dapat berdiri sendiri. Dalam penulisannya, majemuk nominal ditulis serangkai atau satu kata. Misalnya *swasembada* yang terbentuk dari bentuk terikat *swa* ‘sendiri’ dan *sembada* dan *pascasarjana* yang terbentuk dari bentuk terikat *pasca-* dan *dana*. Perhatikan contoh lainnya.

- a. Pertemuan itu dihadiri 40 utusan lembaga *swadaya* masyarakat dari 20 negara anggota OKI.
- b. Sang *mahasiswa* mendapat uang saku dan uang kuliah dari penelitian yang dijalankan.
- c. Saya akan menjadi *pramuwisata* Anda selama perjalanan transfer kita menuju hotel.
- d. Saat ini tim BMKG sedang memeriksa munculnya air dari tanah *pascagempa* ini.
- e. Kami mengharapkan semua pihak berpegang teguh pada asas *praduga* tak bersalah.

Contoh lain:

<i>serbaguna</i>	<i>pascabencana</i>
<i>metafungsi</i>	<i>prasayarat</i>
<i>swafoto</i>	<i>malapraktik</i>
<i>pramuniaga</i>	<i>caturwulan</i>
<i>mahakarya</i>	<i>dasalomba</i>

(4) Majemuk Nominal Sintetis

Majemuk nominal sintetis adalah kata majemuk yang terbentuk dari unsur-unsur yang terikat. Misalnya *biologi* yang terbentuk dari unsur terikat *bio-* dan *-logi*.

Contoh lain:

hidrologi dari *hidro-* dan *-logi*, polihedral dari *poli-* dan *-hedral*,
sosiologi dari *sosio-* dan *-logi*, geologi dari *geo-* dan *-logi*

Jenis majemuk ini jumlahnya terbatas dan umumnya terbatas pada bentuk serapan asing.

4.1.3.3 Majemuk adjektival

Majemuk adjektival adalah kata majemuk yang berintikan adjektiva, misalnya *cantik jelita*, *gagah berani*, *murah hati*, *gegap gempita*, dan *terang benderang*. Kata *gelap*, *besar*, *murah*, *panjang*, dan *terang* masing-masing menjadi inti pada kata majemuk tersebut. Perhatikan contoh pemakaian berikut.

- a. Sang raja hendak menikahkan putrinya yang telah tumbuh menjadi gadis *cantik jelita*.
- b. Dia adalah seorang abdi dan prajurit yang *gagah berani*.
- c. Sang kakek juga sangat *murah hati* karena sering memberi mereka hadiah.
- d. Namun, beberapa detik kemudian kudengar teriakan *gegap gempita* dari ribuan penonton.
- e. Publik berhak mendapat penjelasan yang *terang benderang* terkait informasi itu.

Dalam majemuk adjektival ini terdapat kata majemuk dengan kombinasi bentuk terikat, misalnya *a-* pada *asimetris*, *adi-* pada *adiluhung*, *in-* pada *independen*, *ekstra-* pada *ekstrakurikuler*, dan *hiper-* pada *hipersensitif*. Berikut contoh pemakaiannya.

- a. Kunci keberhasilan dalam perang *asimetris* adalah keunggulan teknologi informasi dan telekomunikasi.
- b. Masyarakat memberikan apresiasi yang tinggi terhadap karya *adiluhung* yang ditunjukkan dalam seni.
- c. Investigasi menyeluruh sepatutnya dilakukan oleh tim *independen* yang dipercaya publik.
- d. Berbagai manfaat dapat diperoleh melalui kegiatan *ekstrakurikuler* teater.
- e. Padahal, belum tentu anak yang banyak bergerak adalah penderita *hiperaktif*.

4.1.3.4 Majemuk Preposisional

Majemuk preposisional adalah kata majemuk yang berintikan preposisi, misalnya *ke sana kemari*, *ke sana-sini*, *ke kanan-kiri*, *di samping itu*, *di satu sisi*, dan *oleh karena itu*. Unsur *ke* (pada *ke sana kemari* dan *ke sana ke sini*), *di* (pada *di samping itu* dan *di satu sisi*), dan *oleh* (pada *oleh karena itu*) masing-masing merupakan preposisi dalam konstruksi tersebut. Perhatikan contoh berikut.

- a. Sudah melamar *ke sana kemari*, pekerjaan itu belum juga didapat.
- b. Saat dicari *ke sana-sini*, batu bersejarah itu tak kunjung tampak.
- c. Anak-anak yang lewat dekat hamparan kacang itu biasanya menoleh *ke kanan-kiri*, lalu menyambar segenggam kalau tidak ada yang mengawasi.
- d. Puasa mengajari itu semua. *Di samping itu*, puasa juga mengajarkan kejujuran.
- e. *Di satu sisi*, mereka tidak berhak menerima subsidi karena dianggap memiliki pekerjaan.
- f. Di sana diamanatkan pembicaraan mengenai koalisi akan dilakukan di rapimnas. *Oleh karena itu*, rapat (rapimnas) hari ini digelar sah sesuai ketentuan.

Berikut adalah contoh lainnya.

- g. ... dikuasai oleh negara sebagaimana amanah Pasal 33 UUD 1945. *Dengan demikian*, kesejahteraan rakyat Indonesia dapat segera tercapai.
- h. Kredibilitas adalah kunci utama. *Oleh sebab itu*, saya dan tim tidak pernah main saham untuk menjaga independensi.
- i. Mereka pun mendapatkan asimilasi ataupun pembebasan bersyarat. *Dengan kata lain*, hukuman yang sudah ringan itu menjadi semakin enteng gara-gara fasilitas ini.
- j. ... pemerintah pusat dan daerah menggunakan data untuk mengambil keputusan. *Dalam hal ini*, perlu dihubungkan antara sistem data dan desakan akuntabilitas.
- k. Di satu sisi PKS menjadi bagian kabinet, sedangkan di sisi lain berseberangan dengan kebijakan pimpinan kabinet.

4.2 Kedudukan Kata Majemuk

Unsur-unsur yang menyusun kata majemuk dapat berbeda-beda dalam hal kedudukannya. Unsur yang satu dapat bersifat sebagai inti, sedangkan

unsur lainnya bersifat pelengkap atau atributif. Oleh karena itu, berdasarkan kedudukannya, kata majemuk terbagi atas majemuk setara (koordinatif) dan taksetara (subkoordinatif).

4.2.1 Majemuk Setara

Majemuk setara adalah kata majemuk yang unsur-unsurnya memiliki kedudukan yang sama dalam hal kelas kata dan fungsinya yang tidak bersifat atributif, misalnya *asal usul*, *fakir miskin*, *ijab kabul*, *muda mudi*, dan *hati nurani*. Perhatikan contoh kalimat berikut.

- a. Salah satu cara untuk mengungkap keragaman ini adalah dengan menelusuri *asal usul* nenek moyang bangsa Indonesia.
- b. Bahkan, mereka yang menelantarkan anak yatim dan *fakir miskin*, dicap sebagai pendusta agama.
- c. Saksi pernikahan telah menunggu mereka melaksanakan *ijab kabul*.
- d. Ada suatu tempat yang nyaman di sana bagi *muda mudi* yang sedang bercengkerama.
- e. Mau tidak mau, harus ada mekanisme *bongkar pasang* yang lebih mudah.

Berdasarkan relasi unsur-unsurnya, kata majemuk setara dibagi atas empat kategori sebagai berikut.

(1) Majemuk Setara Bersinonim

Contoh:

- a. Badanku *basah kuyup* oleh peluh.
- b. Rupanya *jerih payah* Kartini tidak sia-sia.
- c. Teriakan dan *gelak tawa* dari tanah lapang terdengar cukup jelas.
- d. Sampai sekarang mereka mengawasi ketat segala *sepak terjang* Supernova.
- e. Saksi pernikahan telah menunggu mereka melaksanakan *ijab kabul*.

(2) Majemuk Setara Berdampingan

Contoh:

- a. Karies gigi sudah ditemukan pada *nenek moyang* kita ribuan tahun lalu.
- b. Masih banyak kasus konflik lahan di *tanah air* yang tak terekam media.
- c. Alih-alih hanya meminjam koran, keduanya jadi mengobrol *panjang lebar*.

- d. Anda tidak mempunyai *arah tujuan* yang jelas dan pasti.
- e. Mendikbud menambahkan bahwa literasi tidak melulu hanya *baca tulis*.

(3) Majemuk Setara Berlawanan

Contoh:

- a. Ia selalu menjalankan tugas dengan tuntas, tanpa mengeluh, tanpa menghitung *untung rugi*.
- b. Semua *utang piutang* dan segala milik VOC diambil alih oleh pemerintah.
- c. Saat ia pasrah dan ikhlas *jiwa raga* melakukan kehendak-Nya.
- d. Mengejar kecantikan *luar dalam* akan membuat perempuan menjadi lebih bernilai positif.
- e. Dalam bidang ini, *mau tak mau* harus diakui, agaknya kita masih ketinggalan.

(4) Majemuk Setara Berpilihan

Contoh:

- a. Jalanan kini sepi, hanya ada *satu dua* ekor kucing bermata kilau berkeliaran menikmati malam.
- b. Kukenali *dua tiga* wajah akrab tetangga.
- c. Beberapa kali aku bisa membayar *tiga empat* bulan uang indekos dari menang taruhan.

4.2.2 Majemuk Taksetara

Majemuk taksetara dibentuk dari unsur-unsur yang tidak setara atau tidak sama. Salah satu unsurnya memiliki kedudukan yang lebih tinggi daripada yang lain. Hal itu dapat terlihat dari fungsi unsur tersebut yang bersifat atributif dan memiliki kelas kata yang berbeda. Perhatikan contoh berikut.

- a. Bayi-bayi itu ditimang-timang sembari diajak bicara seolah *anak kandung* sendiri.
- b. Sering terjadi *salah paham* yang mengakibatkan saling “senggol” antara saudara.
- c. Mereka meminta *ganti rugi* atas lahan fasos/fasum yang mereka serahkan.
- d. Di saat seseorang *jatuh cinta*, semua hal terlihat begitu indah dan sesuai harapan.
- e. Rasanya tidak masuk akal melihat begitu banyak air di tempat paling

kering di dunia.

Apabila dicermati contoh-contoh kata majemuk di atas, tampak bahwa unsur inti terletak di bagian awal. Namun, ada kata majemuk yang terbentuk dari gabungan bentuk terikat memiliki urutan yang sebaliknya, yaitu unsur penjelas terlebih dahulu dan intinya terletak di belakang. Perhatikan contoh berikut.

- a. Salah satu perubahan sosial ini adalah hilangnya kekuasaan *maharaja* dan kelompok aristokrat (bangsawan) lainnya.
- b. Di sana disimpan berbagai jenis benda *purbakala* yang nilai historisnya sulit diukur dengan rupiah.
- c. Pelecehan yang dilaporkan korban dinyatakan hanya merupakan perbuatan *asusila* biasa dan tidak tergolong pelanggaran berat.
- d. PT Jasa Marga (Persero) Tbk menyebut tol ini sebagai *adikarya*.

4.3 Makna Kata Majemuk

Dilihat dari hubungan makna antara unsur-unsurnya, terdapat beberapa jenis hubungan dalam kata majemuk.

- (1) Hubungan kuantitatif, yaitu hubungan yang unsur pertama dan kedua berhubungan sebagai bagian keseluruhan.

Contoh:

- a. Waktu sudah hampir pukul tiga dini hari dan aku *setengah mati* menahan kantuk.
- b. Oleh sebab itu, memberantas korupsi harus pula dilakukan dengan keberanian *setengah gila*.
- c. Ia mengganti kostumnya seperti *orang kebanyakan*.
- (2) Hubungan kualitatif, yaitu hubungan unsur kedua menyatakan sifat, keadaan, atau jenis komponen pertama.

Contoh:

- a. Tertundanya penyelenggaraan jaminan sosial membuat *rakyat jelata*, tak mendapat perlindungan negara.
- b. Ia dikukuhkan sebagai *guru besar* dalam bidang ilmu manajemen sumber daya perairan.
- c. Kewajiban membayar tunjangan *hari raya* (THR) hendaknya dapat ditunaikan sesuai peraturan.
- (3) Hubungan tujuan, yaitu hubungan yang unsur kedua menyatakan tujuan.

Contoh:

- a. *Pesawat pengebom* ini dilengkapi dengan persenjataan konvensional.
 - b. Ia mempersoalkan kewenangan *dewan pengawas* untuk menilai kinerja direksi.
 - c. Bahasa asing dapat digunakan sebagai *bahasa pengantar* di perguruan tinggi.
- (4) Hubungan perbandingan, yaitu hubungan yang unsur pertama dibandingkan dengan unsur kedua.
Contoh:
- a. Pipinya yang *kuning langsung* bercumbu dengan air mata yang penuh dengan alasan.
 - b. Langit mulai bersemu *merah jambu* di cakrawala timur.
 - c. Saat ini di wajahnya yang *bulat telur* terpasang sebingkai kacamata minus empat.
- (5) Hubungan kepemilikan, yaitu hubungan yang unsur kedua menyatakan milik.
Contoh:
- a. Penurunan harga minyak mentah dunia membuat *kas negara* mengempis.
 - b. Hal inilah pekerjaan rumah yang perlu diselesaikan oleh segenap *anak negeri*.
 - c. Hasil penjualan *anak perusahaan* ini kemudian digunakan untuk menutup sebagian utang.
- (6) Hubungan bagian-keseluruhan, yaitu hubungan yang unsur pertama bagian dari unsur kedua.
Contoh:
- a. Kepalanya tiba-tiba mendongak dengan dua *biji mata* yang semakin berkaca-kaca.
 - b. Masyarakat Indonesia sangatlah plural baik dari segi agama, *suku bangsa*, bahasa dan sebagainya.
 - c. Kuambil busurku dan kupasang *anak panah*.

4.4 Fungsi Kata Majemuk

Ada beberapa fungsi kata majemuk, di antaranya adalah sebagai berikut.

- (1) Membentuk struktur kompleks

Contoh:

matahari dari *mata* dan *hari*

saptalomba dari *sapta-* dan *lomba*

darmabakti dari *darma* dan *bakti*

dukacita dari *duka* dan *cita*

purbakala dari *purba* dan *kala*

(2) Membentuk makna baru

Contoh:

meja hijau ‘pengadilan’
jago merah ‘api; kebakaran’
jantung hati ‘kesayangan’
tangan kanan ‘orang kepercayaan’
besar kepala ‘sombong’

Berikut adalah contoh penggunaannya.

- a. Tak sedikit konflik yang akhirnya harus berakhir di *meja hijau*.
- b. Dengan cepat si *jago merah* merambat ke dua lantai di atasnya.
- c. Meski anak pelatih, tak ada cerita ia jadi *anak emas* di sana.
- d. Ia juga menjadi *tangan kanan* Ferdi mencari gudang penyimpanan narkotika.
- e. Semua fasilitas tidak membuat kami manja dan *besar kepala*.

(3) Membuat bentuk jamak

Contoh:

harta benda ‘*harta* dan *benda*’
 daya upaya ‘*daya* dan *upaya*’
 bujuk rayu ‘*bujuk* dan *rayu*’
 bala tentara ‘*bala* dan *tentara*’
 belas kasihan ‘*belas* dan *kasihan*’

Berikut adalah contoh penggunaannya.

- a. Badai ini juga mampu menimbulkan korban jiwa dan kerusakan *harta benda*.
- b. Segala *daya upaya* telah dilakukan untuk mengatasinya.
- c. Dengan segala *bujuk rayu* dia diminta untuk tidak keluar dari perusahaan.
- d. Pasukan Yudistira sudah siap menunggu kedatangan *bala tentara* Kurawa.
- e. Mereka memandang kami seolah-olah kami hidup dari *belas kasihan* mereka.

(4) Menyangatkan maksud atau pernyataan

Contoh:

gelap gulita bermakna ‘sangat gelap’
susah payah bermakna ‘sangat susah’
penuh sesak bermakna ‘sangat penuh’
muda belia bermakna ‘sangat muda’
sunyi senyap bermakna ‘sangat sunyi’

Berikut adalah contoh penggunaannya.

- Kondisi *gelap gulita* selama gerhana berlangsung mengingatkan manusia akan suasana alam kubur kelak
- Meski dengan *susah payah*, peti itu berhasil diangkat ke atas dermaga.
- Dapat kurasakan bagaimana kapal yang *penuh sesak* itu terombang-ambing di atas lautan lepas.
- Ditemukan juga data bahwa pelaku kekerasan banyak yang usianya *muda belia*.
- Kalaupun tidak *sunyi senyap*, maka rumah besar itu penuh dengan intrik dan pertengkaran yang panas.

5. PENUTUP

Berdasarkan pemaparan berbasis korpus, bisa dilihat bahwa kata majemuk atau konstruksi majemuk bisa diklasifikasikan berdasarkan beberapa kriteria, baik berdasarkan kriteria semantis dan juga kriteria sintaksis. Dari contoh yang ditemukan, ternyata kata majemuk tidak harus terbentuk dari dua buah kata atau morfem bebas saja, tetapi bisa lebih. Namun, di sini masih memerlukan diskusi tentang batasan “kata majemuk”. Jika ingin mempertahankan istilah “kata”, maka janganal kalau menyebutkan contoh seperti *anak kemarin sore*, *penjahat kelas kakap*, *oleh karena itu* dan *dalam hal ini* sebagai sebuah kata. Oleh karena itu, pilihan “konstruksi majemuk” muncul untuk menggantikan “kata majemuk”. Hal ini tidak bisa dihindari ketika seorang pakar memutuskan untuk membahas kata majemuk atau konstruksi majemuk secara lebih luas dan mendalam. Berbagai klasifikasi itu dan diskusi singkatnya ditampilkan dalam tulisan ini.

Berdasarkan maknanya, kata majemuk bisa diklasifikasikan secara semantis menjadi tiga, yakni kata majemuk idiomatis penuh (contohnya: *naik daun* dan *tangan kanan*), kata majemuk idiomatis sebagian (contohnya: *penumpang gelap* dan *kabar burung*) dan kata majemuk kolokasi terbatas

(contohnya: *ambil keputusan* dan *jalan santai*).

Kata majemuk juga dapat dikelompokkan berdasarkan kelas kata unsur penyusunnya: majemuk verbal, majemuk nominal, majemuk adjektival, dan majemuk preposisional. Majemuk verbal adalah kata majemuk yang berintikan verba (contohnya: *gulung tikar*, dan *jatuh bangun*).

Dari segi bentuknya, majemuk verbal dapat dibagi menjadi tiga kelompok, yaitu majemuk verbal dasar (contohnya: *ikut campur* dan *mabuk laut*), majemuk verbal berafiks (contohnya: *berserah diri* dan *menyebarkan*) dan majemuk verbal berulang (contohnya: *geleng-geleng kepala* dan *berpindah-pindah tangan*).

Majemuk nominal adalah kata majemuk yang berintikan nomina. Berdasarkan bentuk morfologisnya, majemuk nominal terbagi atas (1) majemuk nominal dasar (contohnya: *rumah tangga* dan *kepala dingin*), (2) majemuk nominal berafiks (contohnya: *penegak hukum* dan surat berharga), (3) majemuk nominal dengan bentuk terikat, (contohnya *pascagempa* dan *praduga*) dan (4) majemuk nominal sintetis. (contohnya: *biologi* dan *sosiologi*).

Majemuk adjektival adalah kata majemuk yang berintikan adjektiva (contohnya: *cantik jelita*, *gagah berani*, dan *terang benderang*).

Majemuk preposisional adalah kata majemuk yang berintikan preposisi (contohnya: *di samping itu*, *di satu sisi*, *oleh karena itu*, *dengan kata lain* dan *dalam hal ini*). Majemuk preposisional merupakan konstruksi majemuk yang tidak dijadikan contoh prototipe atau contoh yang lazim digunakan sebagai contoh kata majemuk, meskipun frekuensinya sering dan tersebar di berbagai macam jenis teks.

Berdasarkan kedudukannya, kata majemuk terbagi atas majemuk setara (koordinatif) dan taksetara (subkoordinatif). Kata majemuk setara terdiri dari majemuk setara bersinonim (contohnya: *basah kuyup* dan *jerih payah*), majemuk setara berdampingan (contohnya: *nenek moyang* dan *tanah air*), majemuk setara berlawanan (contohnya: *untung rugi* dan *utang piutang*) dan majemuk setara berpilihan (contohnya: *satu dua (+nomina)*, *dua tiga (+nomina)*). Majemuk taksetara dibentuk dari unsur-unsur yang tidak setara atau tidak sama. Salah satu unsurnya memiliki kedudukan yang lebih tinggi daripada yang lain (contohnya: *anak kandung* dan *salah paham*).

Berdasarkan hubungan makna antara unsur-unsurnya, terdapat beberapa jenis hubungan dalam kata majemuk, yakni hubungan kuantitatif (contoh: *setengah mati* dan *orang kebanyakan*), hubungan kualitatif, (contoh: *rakyat*

jelata dan *guru besar*), hubungan tujuan (contoh: *pesawat pengebom* dan *dewan pengawas*), hubungan perbandingan (contoh: *kuning langsung* dan *bulat telur*), hubungan kepemilikan (contoh: *kas negara* dan *anak perusahaan*).

Ada beberapa fungsi kata majemuk, di antaranya adalah untuk membentuk struktur kompleks (contoh: *matahari* dari *mata* dan *hari* dan *purbakala* dari *purba* dan *kala*), untuk membentuk makna baru (contoh: *meja hijau* ‘pengadilan’ dan *jago merah* ‘api’), untuk membuat bentuk jamak (contoh: *harta benda* dan *bujuk rayu*), untuk menyangatkan maksud atau pernyataan (contoh: *gelap gulita* dan *susah payah*).

Melihat klasifikasi tersebut, jelas bahwa kata majemuk atau konstruksi majemuk merupakan bagian dari frasem. Artinya, bidang fraseologi yang memang bersifat interdisipliner bisa menelaah kata atau konstruksi majemuk karena memenuhi kriteria frasem, yakni polileksikalitas, kepaduan/keajekan dan adanya sebagian konstruksi itu yang bersifat idiomatis.

Referensi

- [1] Abdul Chaer, *Morfologi Bahasa Indonesia: Pendekatan Proses*, Jakarta: Rineka Cipta, 2015.
- [2] Anton M. Moeliono, H. Lapoliwa, H. Alwi, S.S.T Wisnu Sasangka, and Sugiyono, *Tata Bahasa Baku Bahasa Indonesia (4th Edition)*. Jakarta: Badan Pengembangan dan Pembinaan Bahasa, 2017.
- [3] C.A. Mees, *Tatabahasa Indonesia*. Djakarta-Groningen: J.B. Wolters, 1954.
- [4] Harald Burger, *Phraseologie. Eine Einführung am Beispiel des Deutschen*. 5., neu bearbeitete Auflage. Berlin: Erich Schmidt Verlag, 2015.
- [5] Harimurti Kridalaksana, *Pembentukan kata dalam bahasa Indonesia*, Jakarta: Gramedia, 2007.
- [6] I Wayan Simpen, *Morfologi: Kajian Proses Pembentukan kata*. Jakarta: Bumi Aksara 2021.
- [7] J N Sneddon, A Adelaar, D-N Djenar, and M-C Ewing, *Indonesian Reference Grammar (2nd Edition)*. New South Wales: Allen & Unwin, 2010.
- [8] Jos Daniel Parera, *Morfologi*. Jakarta: Gramedia, 2007.
- [9] M. Ramlan, *Morfologi: Suatu tinjauan deskriptif*. Yogyakarta: Karyono, 2012.

- [10] Mike Scott, *WordSmith Tools* version 8 (64 bit version) Stroud: Lexical Analysis Software, 2022.
- [11] Paul Rayson, “Computational tools and methods for corpus compilation and analysis,” in Douglas Biber and Randi Reppen (eds.), *The Cambridge handbook of English corpus linguistics*. Cambridge: Cambridge University Press, 2015, pp. 32–49.
- [12] S. Takdir Alisjahbana, *Tatabahasa Baru Bahasa Indonesia 2*. Djakarta: Pustaka Rakjat. 1963.
- [13] Slametmuljana, *Kaidah Bahasa Indonesia*. Flores: Nusa Indah, 1969.
- [14] Stefan Th. Gries, “Phraseology and linguistic theory: A brief survey,” in Sylviane Granger and Fanny Meunier (eds.), *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia: John Benjamin, 2008, pp. 4—26.
- [15] Sylviane Granger and Magali Paquot, “Disentangling the phraseological web,” in Sylviane Granger and Fanny Meunier (eds.), *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia: John Benjamin, 2008, pp. 27—50.

REDUPLIKASI SEBAGIAN, SALIN SUARA, DAN TRILINGGA DALAM BAHASA INDONESIA

David Moeljadi

Universitas Bahasa Asing Kanda, Jepang

Abstrak

Dari segi bentuknya, perulangan atau reduplikasi dalam bahasa Indonesia dapat dibedakan menjadi (1) reduplikasi utuh, (2) reduplikasi salin suara, (3) reduplikasi sebagian, dan (4) reduplikasi sinonim [1]. Artikel ini membahas bentuk dan makna reduplikasi sebagian atau dwipurwa, reduplikasi salin suara atau dwilingga salin suara, dan reduplikasi suku kata sebanyak tiga kali atau trilingga. Penulis mengumpulkan kata-kata ulang dwipurwa, dwilingga salin suara, dan trilingga dari Kamus Besar Bahasa Indonesia Edisi V (KBBI V), Malindo Morph [2] yang datanya bersumber dari korpus Leipzig [3], dan korpus TBIK. Hasil pencarian kata-kata ulang di korpus TBIK dianalisis berdasarkan frekuensi kemunculan, sumber data, bentuk, dan maknanya.

Kata kunci: kata ulang, reduplikasi, dwipurwa, dwilingga salin suara, trilingga

1. PENDAHULUAN

Bab ini membahas definisi jenis-jenis reduplikasi dan tujuan penulisan artikel. Ada tiga jenis reduplikasi yang dibahas: reduplikasi sebagian atau dwipurwa, reduplikasi salin suara atau dwilingga salin suara, dan trilingga. Istilah-istilah jenis reduplikasi ini dipungut dari bahasa Jawa yang menyerapnya dari bahasa Sanskerta: *dwi* berarti ‘dua’, *lingga* berarti ‘kata/bentuk dasar’, *purwa* berarti ‘awal’, dan *tri* berarti ‘tiga’ [4]. Dalam artikel ini selanjutnya digunakan istilah-istilah tersebut.

Definisi dwipurwa, dwilingga salin suara, dan trilingga di artikel ini mengacu pada definisi yang ada pada KBBI V. Dwipurwa adalah pengulangan sebagian atau seluruh suku awal sebuah kata atau leksem dengan pelemahan vokal (menjadi e pepet), misalnya *tamu* menjadi *tetamu*, *laki* menjadi *lelaki*. Bentuk dwipurwa ini tidak bertanda hubung. Dwilingga salin suara adalah pengulangan kata penuh dengan variasi vokal atau konsonan, misalnya *bolak-balik* dan *sayur-mayur*. Bentuk dwilingga salin suara bertanda hubung satu. Trilingga adalah pengulangan unsur (suku kata) sebanyak tiga kali dengan

variasi fonem, misalnya *dag-dig-dug*, *cas-cis-cus*, dan *ngak-ngik-ngok*. Bentuk trilingga bertanda hubung dua.

Artikel ini disusun dengan tujuan untuk mendeskripsikan jenis/subklasifikasi, bentuk/varian, dan fungsi/makna dwipurwa, dwilingga salin suara, dan trilingga. Selain itu, artikel ini membahas kecenderungan penggunaan (frekuensi) yang sering digunakan dan yang jarang digunakan, dirinci berdasarkan jenis teks dalam korpus bahasa Indonesia formal dan informal dalam sepuluh tahun terakhir (2011–2020). Dalam penjelasan artikel ini terdapat beberapa hal yang belum diatur dalam tata bahasa baku bahasa Indonesia [1].

2. TINJAUAN PUSTAKA

2.1 Dwipurwa

Kulsum [5] menulis bahwa dwipurwa merupakan reduplikasi yang tidak produktif dan mempunyai kecenderungan digunakan dalam ragam sastra (puisi atau lagu) dan media sosial karena beberapa media sosial sangat bergantung pada jumlah karakter. Dwipurwa dan dwilingga dengan kata dasar yang sama, misalnya *lelaki* dan *laki-laki*, dapat mempunyai makna yang sama atau berbeda. Dwipurwa mempunyai potensi untuk dikembangkan untuk memperkaya istilah dan kosakata, terutama dalam media sosial. Salah satu kendala dalam pengembangan dwipurwa adalah kata yang berhuruf awal vokal dapat dipastikan tidak dapat dibuat bentuk dwipurwanya.

Kridalaksana [6] menulis bahwa dwipurwa dapat berfungsi membentuk nomina yang bermakna (1) ‘jamak’, misalnya *dedaunan*, *pepohonan*, dan *rerumputan*; (2) sama dengan bentuk dasarnya, misalnya *lelaki* (sama dengan *laki*) dan *tetamu* (sama dengan *tamu*); (3) ‘segala macam’, misalnya *reramuan*, *sesajian*, dan *reruntuhan*; (4) ‘yang dianggap’, misalnya *leluhur*; (5) ‘kumpulan’, misalnya *dedaunan*, *reruntuhan*, dan *pepohonan*.

2.2 Dwilingga Salin Suara

Sariah [7] mengumpulkan data dwilingga salin suara dari Kamus Besar Bahasa Indonesia Edisi IV (KBBI IV). Dia mengkaji jenis, makna, dan penggunaan dwilingga salin suara dalam kalimat. Ada empat bentuk dwilingga salin suara, yaitu dwilingga salin suara dengan (1) perubahan fonem vokal, (2) perubahan fonem konsonan, (3) penambahan leksem, dan (4) penambahan

infiks. Dwilingga salin suara dengan perubahan fonem vokal didominasi bentuk *a-i*, *a-u*, dan *o a - a i*. Dwilingga salin suara dapat bermakna ‘banyak’ atau ‘bermacam-macam’, tiruan bunyi, kualitas, dan keadaan.

Denistia dan Baayen [8] menulis bahwa dwilingga salin suara dengan bentuk dasar adjektiva bermakna ‘intensif’ dan dwilingga salin suara dengan bentuk dasar nomina dan verba bermakna jamak. Dwilingga salin suara tidak produktif karena perubahan fonemnya (vokal atau konsonan) tidak dapat diprediksi. Biasanya unsur pertama dwilingga salin suara merupakan bentuk dasarnya, dengan kata lain unsur pertama adalah bentuk dasar yang merupakan morfem bebas dan unsur kedua adalah morfem terikat yang tidak dapat berdiri sendiri, misalnya *sayur-mayur* [9]. Selain itu, ada pula dwilingga salin suara yang bentuk dasarnya ada di posisi kedua, misalnya *corat-coret*, juga dwilingga salin suara yang unsur pertama dan keduanya merupakan morfem terikat, misalnya *mondar-mandir*.

Kridalaksana [6] menulis bahwa dwilingga salin suara dapat berfungsi membentuk adjektiva yang bermakna ‘sungguh-sungguh (intensif)’, misalnya *pontang-panting* dan *kucar-kacir*. Selain itu, juga dapat berfungsi membentuk nomina yang bermakna ‘bermacam-macam’, misalnya *sayur-mayur*, *warna-warni*, *serba-serbi*, dan *corat-coret*. Selain membentuk adjektiva dan nomina, dwilingga salin suara juga dapat membentuk verba, misalnya *mondar-mandir* dan *kelap-kelip*; dan dapat membentuk adverbia, misalnya *serta-merta*. Dalam bentuk *bolak-balik*, *gembar-gembor*, *kelap-kelip*, dan *corat-coret* unsur atau komponen kedua adalah morfem dasarnya, jadi proses reduplikasinya bersifat regresif. Di samping itu, terdapat reduplikasi resiprokal, yaitu *desas-desus*, *mondar-mandir*, *kocar-kacir*, dan *pontang-panting*.

2.3 Trilingga

Kridalaksana [6] hanya menyebutkan bahwa trilingga (salin suara) berfungsi membentuk verba, misalnya *cas-cis-cus* dan *dag-dig-dug*; dan membentuk nomina, misalnya *dar-der-dor* dan *ngak-ngeknok*. Tidak diuraikan lebih lanjut tentang makna bentuk trilingga.

3. METODE

Penulis menggunakan dua sumber utama dalam pengumpulan kata-kata ulang dwipurwa dan dwilingga salin suara, yaitu Kamus Besar Bahasa

Indonesia Edisi V (KBBI V) dan MALINDO Morph. KBBI V adalah kamus acuan/referensi bahasa Indonesia edisi termutakhir. MALINDO Morph adalah kamus morfologi daring bahasa Indonesia dan Melayu yang datanya berasal dari KBBI, Kamus Dewan, korpus Leipzig, dan penelitian lapangan. Kamus morfologi ini berisi informasi kata turunan, kata dasar, prefiks, sufiks, konfiks, dan tiga jenis reduplikasi yaitu dwilingga, dwipurwa, dan dwilingga salin suara. Pertama-tama, semua kata ulang tersebut dikumpulkan dari KBBI V dan diperoleh 47 dwipurwa dan 312 dwilingga salin suara. Setelah itu, kata ulang tersebut digabung dengan kata ulang di MALINDO Morph dan diperoleh 211 dwipurwa dan 347 dwilingga salin suara. Tidak ada informasi trilingga dalam MALINDO Morph.

Khusus untuk trilingga, penulis menggunakan sumber tambahan selain KBBI V, yaitu Tata Bahasa Baku Bahasa Indonesia Edisi Keempat (TBBBI IV) [1] dan korpus TBIK di CQPWeb [10]. Dari KBBI V dan TBBBI IV diperoleh lima trilingga. Data untuk trilingga dari korpus TBIK di CQPWeb diperoleh dengan menggunakan pencarian melalui *regular expression* (regex). Simbol regex yang digunakan adalah sebagai berikut.

$$[\text{word} = "[a-z]^+ - [a-z]^+ - [a-z]^+"]$$

Simbol itu digunakan untuk menjaring semua kata berhuruf a-z yang mengandung dua tanda hubung (-). Ada 723 tipe berbeda dari hasil pencarian tersebut. Akan tetapi, tidak semua tipe merupakan trilingga. Terdapat juga bentuk lain seperti kesalahan pemenggalan kata, kata asing, gabungan kata, imbuhan dengan kata dasar asing, dan onomatope yang bukan trilingga. Untuk itu, dilakukan penyeleksian data. Jumlah akhir trilingga di korpus setelah proses penyeleksian adalah tiga belas buah.

Dalam korpus terdapat juga bentuk dwilingga salin suara dan trilingga yang tidak ditulis dengan menggunakan tanda hubung. Namun, pencarian bentuk tersebut dalam CQPWeb tidak mudah. Oleh karena itu, data dibatasi hanya pada bentuk ulang yang penulisannya sesuai dengan kaidah ejaan bahasa Indonesia, yaitu menggunakan tanda hubung untuk dwilingga salin suara dan trilingga. Selain untuk pencarian kata, CQPWeb juga digunakan untuk penelitian kuantitatif (penghitungan frekuensi kemunculan dan sumber). Makna kata-kata ulang di dalam kalimat-kalimat hasil pencarian dianalisis secara kualitatif berdasarkan konteks pemakaian kata.

4. HASIL DAN PEMBAHASAN

4.1 Dwipurwa

Dari 211 dwipurwa yang berhasil dikumpulkan dari KBBI V dan MALINDO Morph, diperoleh sebanyak 74 dwipurwa yang ada di korpus TBIK. Sebagian besar dwipurwa berkelas kata nomina, seperti *lelaki*, *tetangga*, dan *wewenang*. Dwipurwa yang tidak berkelas kata nomina adalah sebagai berikut: *beberapa* (numeralia), *seseorang*, *sesuatu* (pronomina), *bebuyutan* (adjektiva), *sesekali* (adverbia), *tetirah*, *bebenah*, dan *sesumbar* (verba). Konjungsi *tetapi* bukanlah dwipurwa karena diserap dari bahasa Sanskerta *tathāpi* yang berarti ‘meskipun, namun, masih’ menurut informasi etimologis dalam KBBI V. Dwipurwa *bebenah* dan *sesumbar* memiliki label bahasa Melayu Jakarta dalam KBBI V, terutama digunakan dalam teks Melayu Jakarta. Sebagian besar dwipurwa terdiri atas tiga suku kata.

Dari segi bentuknya, 74 dwipurwa tersebut dapat diklasifikasikan sebagai berikut: (1) dwipurwa tanpa sufiks, misalnya *dedaun*; dan (2) dwipurwa dengan sufiks *-an*, misalnya *dedaunan*. Ada sebanyak 57 dwipurwa tanpa sufiks dan 17 dwipurwa dengan sufiks *-an*, sebagai berikut.

*pepohonan dedaunan bebatuan reruntuhan rerumputan wewangian
bebuyutan bebunyian sesembahan tetabuhan tetumbuhan geguritan
bebauan bebuahansesianganseserahan
sesengukan*

Hal yang menarik dari dwipurwa ini adalah adanya bentuk bersaing antara dwipurwa dengan sufiks *-an* dan dwipurwa tanpa sufiks *-an*. Ada lima pasang bentuk bersaing yang ditemukan di korpus, yaitu:

No.	Dwipurwa dengan <i>-an</i>	Jumlah Kemunculan	Dwipurwa tanpa <i>-an</i>	Jumlah Kemunculan
1	<i>pepohonan</i>	330	<i>pepohon</i>	6
2	<i>dedaunan</i>	211	<i>dedaun</i>	9
3	<i>bebatuan</i>	190	<i>bebatu</i>	2
4	<i>wewangian</i>	47	<i>wewangi</i>	1
5	<i>bebunyian</i>	20	<i>bebunyi</i>	3

Selain lima pasang bentuk bersaing di atas, ada juga kata *bebau* yang muncul dua kali dari sumber yang sama, yaitu biografi tahun 2015, tetapi kata ini bukanlah dwipurwa karena bermakna ‘pengurus desa’ yang tidak ada hubungannya dengan *bau*.

Dwipurwa *pepohon*, *dedaun*, *bebatu*, *wewangi*, dan *bebunyi* bermakna sama seperti bentuk bersaingnya dengan sufiks *-an*, yaitu bermakna jamak atau ‘segala macam’ atau ‘kumpulan’, seperti dalam contoh-contoh kalimat berikut ini. Dwipurwa *pepohon* dalam contoh (1) dan (2) dapat diganti dengan *pepohonan* atau *pohon-pohon*. Dwipurwa *dedaun* dalam contoh (3) dan (4) dapat diganti dengan *dedaunan* atau *daun-daun*. Demikian juga dengan dwipurwa *bebatu* dalam contoh (5) dan (6) yang dapat diganti dengan *bebatuan* atau *batu-batu*. Dwipurwa *bebatu* muncul dua kali dan keduanya berasal dari sumber (novel) yang sama. Dwipurwa *wewangi* hanya muncul sekali dalam contoh (7) dan bermakna sama dengan *wewangian*. Dwipurwa *bebunyi* berasal dari dua sumber, yaitu cerpen dan novel, seperti dalam contoh (8)--(10). Dwipurwa ini bermakna sama dengan *bebunyian*.

- (1) Di hadapannya, terhampar sebuah kebun yang tak begitu luas. *Pepohon* pisang, pepaya, dan mangga tumbuh di dalamnya. (Cerpen_C1A15070)
- (2) Kelelawar tinggal di sana. Ribu-ribu dalam kawanan, yang berpencaran ke *pepohon* dan rumah-rumah setiap malam. (Novel_D1A13003)
- (3) *Dedaun* kering yang basah oleh hujan subuh bergerisik saat beradu dengan sepatu-sepatu gereja. (Cerpen_C1A20070)
- (4) Wis melewati gapura besar bata merah, ia merasakan energi yang baik, seperti datang bersama zat asam yang dihembus-hembus *dedaun*. (Novel_D1A13003)
- (5) Dari puncak tebing *bebatu* itu kau bisa melihat Merapi di alas dan samudra di bawah: gunung dan laut yang keramat. (Novel_D1A13003)
- (6) Mereka duduk di *bebatu*, ia dan lelaki bertubuh kecil itu. (Novel_D1A13003)
- (7) Sebab bukan ia tak mau menulis omong kosong di kertas merah muda dengan *wewangi* dari bedak, tapi ia sungguh tak tahu apa mesti diperbincangkan, sebab tak ada yang menarik sepanjang hidupnya di rindangan pohon menanti orang yang gelisah sebab rambut telah mencolok mata. (Novel_D1A14002)
- (8) Satu-satunya sumber cahaya hanyalah bulan yang bulat penuh

menyerupai martabak manis kesukaanmu. Hanya ada *bebunyi* derik serangga malam dan goyangan daun-daun. (Cerpen_C1A18124)

- (9) Demi mendengar suara Maesa Dewi, nada tinggi dan penuh kecemasan, serta *bebunyi* selimut yang dihentakkan dan ranjang berderak serta kaki menjejak ke lantai, Margio kembali menyarangkan giginya ke rekahan merah gelap dan basah itu, ciuman kedua yang lebih mematikan dan dikuasai nafsu. (Novel_D1A14002)
- (10) Maesa memandanginya sejenak, sebelum berjinjit bergegas ke kamar mandi, hilang di sana memberi *bebunyi* deras kencing yang tumpah ke lubang toilet. (Novel_D1A14002)

Dari segi maknanya, dwipurwa menyatakan pluralitas, kesamaan atau kemiripan dengan bentuk dasarnya, dan ketidaktentuan. Makna-makna tersebut dapat diuraikan lebih lanjut dan digolongkan ke dalam tujuh kelompok sebagai berikut:

1. ‘kumpulan’ atau ‘segala/berbagai macam’. Dwipurwa yang memiliki makna ini sebagian besar berakhiran *-an*. Misalnya,

● <i>dedaunan, dedaun</i>	● <i>tetumbuhan</i>
● <i>bebatuan, bebatu</i>	● <i>rerumputan</i>
● <i>bebunyan, bebunyi</i>	● <i>reramuan</i>
● <i>wewangian, wewangi</i>	● <i>tetabuhan</i>
● <i>pepohonan, pepohon</i>	● <i>reruntuhan</i>
● <i>bebauan</i>	● <i>sesajian</i>
● <i>bebuahan</i>	● <i>sesajen</i>
2. ‘banyak’. Misalnya,
 - *tetamu*
 - *rerimbun*
 - *reranting*
 - *rerangka*
3. ‘yang dianggap’ atau ‘yang di-...-kan’. Misalnya,
 - *leluhur*
 - *tetua*
 - *sesepuh*
 - *sesembahan*

4. ‘sesuatu yang sama dengan bentuk dasarnya’. Misalnya, *lelaki*

- | | |
|-------------------|-------------------|
| ● <i>jejaka</i> | ● <i>gegara</i> |
| ● <i>bebenah</i> | ● <i>tetiba</i> |
| ● <i>sesumbar</i> | ● <i>wewenang</i> |
| ● <i>sesama</i> | ● <i>rerata</i> |

6. ‘sesuatu yang mirip dengan bentuk dasarnya’. Misalnya,

- *jejaring*
- *tetikus*
- *lelembut*

7. ‘tidak tentu’. Bentuk dasarnya berkelas kata numeralia seperti *suatu* dan *berapa*, mengandung numeralia seperti *seorang* (*satu orang*) dan *sekali* (*satu kali*), atau pronomina seperti *siapa*. Misalnya,

- *sesuatu*
- *seseorang*
- *sese kali*
- *sesiapa*
- *beberapa*

8. makna lainnya. Misalnya,

- *tetangga*
- *pepatah*
- *bebuyutan*
- *dedalu*
- *cecunguk*

Dwipurwa *tetangga*, *pepatah*, *dedalu*, dan *cecunguk* mungkin dapat dideskripsikan dengan menggunakan pendekatan etimologis atau asal-usul kata.

Tabel 1 memuat sepuluh dwipurwa terbanyak di korpus TBIK. Numeralia *beberapa* menduduki peringkat pertama dengan persentase 45,86%, hampir separuh dari keseluruhan frekuensi kemunculan semua dwipurwa yang ada di korpus. Setelah itu, pronomina *seseorang* (14,16%) dan *sesuatu* (12,65%) masing-masing menduduki peringkat kedua dan ketiga. Nomina *lelaki* menduduki peringkat keempat dengan persentase 9,48%. Dwipurwa yang lain

berfrekuensi kurang dari 3%.

Tabel 1 Sepuluh Dwipurwa Terbanyak

No.	Dwipurwa	Kelas Kata	Jumlah Kemunculan	Persentase
1	<i>beberapa</i>	Numeralia	29.750	45,86%
2	<i>seseorang</i>	Pronomina	9.186	14,16%
3	<i>sesuatu</i>	Pronomina	8.208	12,65%
4	<i>lelaki</i>	Nomina	6.149	9,48%
5	<i>tetangga</i>	Nomina	1.895	2,92%
6	<i>wewenang</i>	Nomina	1.846	2,85%
7	<i>sesama</i>	Nomina	1.690	2,61%
8	<i>sesekali</i>	Adverbia	1.405	2,17%
9	<i>lelembut</i>	Nomina	906	1,40%
10	<i>jejaring</i>	Nomina	764	1,18%

Bentuk *jejaring* dan *tetikus* telah memperkaya kosakata bahasa Indonesia dalam bidang teknologi informasi, seperti dalam contoh (11) dan (12). Ada pula bentuk *tetiba* (yang bermakna sama dengan *tiba-tiba*) dan *gegara* (yang bermakna sama dengan *gara-gara*) yang walaupun jumlah kemunculannya sedikit, berpotensi memperkaya bahasa Indonesia, seperti dalam contoh (13) dan (14). Dwipurwa *tetiba* muncul lima kali: dua kali dalam majalah (2013), satu kali dalam majalah (2019), dan dua kali dalam cerpen (2020), sedangkan dwipurwa *gegara* muncul empat kali: dua kali dalam biografi (2018), dan masing-masing satu kali dalam majalah (2019) dan cerpen (2020).

- (11) Bakrie Global Group dikabarkan telah membeli saham Path, perusahaan *jejaring* sosial asal Amerika Serikat. (Koran_A2DZ14001)
- (12) Komputer desktop dengan spesifikasi tinggi lengkap dengan headset, *tetikus*, dan keyboard menjadi “pegangan” standar selama bermain. (Majalah_B2H19029)
- (13) Saat buang air kecil, Dimas *tetiba* melihat sekelompok monyet yang sedang mengamati gerak-geriknya. (Majalah_B3A13006)
- (14) Di usia lima belas, Janitra mencakar pipi kekasihnya *gegara* mencuri ciuman pertamanya dengan kasar. (Cerpen_C1A20104)

Seperti yang ditulis oleh Kulsum [5], dwipurwa cenderung muncul dalam ragam sastra. Tabel 2 memuat frekuensi kemunculan dwipurwa berdasarkan sumber data. Cerpen dan novel merupakan dua sumber teratas, sedangkan

laman resmi, perundangan, dan surat resmi merupakan tiga sumber terbawah.

Tabel 2 Frekuensi Kemunculan Dwipurwa berdasarkan Kategori/Sumber

Kategori	Jumlah Kemunculan	Frekuensi Per Sejuta Kata
Cerpen	12.211	4.876,02
Novel	8.094	3.086,06
Populer	7.028	3.068,69
Buku teks	6.670	2.702,66
Disertasi, tesis, skripsi	6.565	2.428,23
Biografi	4.982	2.033,09
Majalah	4.900	2.008,51
Jurnal	4.814	1.997,42
Koran	4.162	1.540,90
Laman resmi	2.858	1.170,90
Perundangan	1.525	609,94
Surat resmi	1.022	418,75
Total	64.831	2.163,07

Tidak ada perbedaan frekuensi kemunculan dwipurwa berdasarkan tahun. Antara tahun 2011 hingga tahun 2020 dwipurwa muncul dalam data korpus sebanyak 6.000 hingga 7.000 kata per tahun.

4.2 Dwilingga Salin Suara

Dari 347 dwilingga salin suara yang berhasil dikumpulkan dari KBBI V dan MALINDO Morph, diperoleh sebanyak 85 dwilingga salin suara yang ada di korpus TBIK. Sebagian besar dwilingga salin suara berkelas kata nomina, seperti *basa-basi*, *asal-usul*, dan *warna-warni*. Dwilingga salin suara yang tidak berkelas kata nomina adalah sebagai berikut.

Verba: *bolak-balik*, *mondar-mandir*, *komat-kamit*, *kasak-kusuk*, *gonjang-ganjing*, *cengar-cengir*, *gonta-ganti*, *celingak-celinguk*, *kelap-kelip*, *gembur-gembor*, *otak-atik*, *wira-wiri*, *utak-atik*, *jungkat-jungkit*, *lontang-lantung*, *selang-seling*, *kebat-kebit*, *lenggak-lenggok*, *runtang-runtung*, *kempas-kempis*, dan *megal-*

megol.

Adjektiva: *kocar-kacir, compang-camping, karut-marut, morat-marit, remeh-temeh, ceplas-ceplos, ramah-tamah, zig-zag, petantang-petenteng, ketar-ketir, coreng-moreng, plin-plan, mencang-mencong, mencla-mencle, orak-arik, plintat-plintut, gerabak-gerubuk, hubar-habir, kopat-kapit, lekak-lekuk, plonga-plongo, dan ropak-rapik.*

Adverbia: *serta-merta, pontang-panting, dan serba-serbi.*

Beberapa dwilingga salin suara, seperti *ceplas-ceplos, mencla-mencle, plintat-plintut, runtang-runtung, kopat-kapit, plonga-plongo, unggah-ungguh* (Jawa), *petantang-petenteng* (Melayu Jakarta), dan *kebat-kebit* (Minangkabau) memiliki label bahasa daerah dalam KBBI V. Unsur-unsur pembentuk dwilingga salin suara *asal-usul*, yaitu *asal* dan *usul* masing-masing diserap dari bahasa Arab; dwilingga salin suara *zig-zag* diserap dari bahasa Inggris; dwilingga salin suara *tiki-taka* diserap dari bahasa Spanyol; dwilingga salin suara *wara-wiri, gono-gini, dan grassah-grusuh* diserap dari bahasa Jawa. Dapat dilihat bahwa kosakata serapan bahasa daerah, terutama bahasa Jawa, dan kosakata serapan bahasa asing telah memperkaya kosakata bahasa Indonesia melalui dwilingga salin suara. Sebagian besar unsur-unsur pembentuk dwilingga salin suara terdiri atas dua suku kata.

Dari segi bentuknya, 85 dwilingga salin suara tersebut dapat diklasifikasikan ke dalam dua kelompok, yaitu dwilingga salin suara dengan perubahan fonem vokal dan dwilingga salin suara dengan perubahan fonem konsonan. Kata-kata ulang dengan penambahan leksem seperti *lalu-lalang* dan kata-kata ulang dengan penambahan infiks seperti *ayun-temayun* tidak dibahas dalam artikel ini karena berdasarkan definisi pada bab pendahuluan, dwilingga salin suara adalah pengulangan kata penuh dengan variasi vokal atau konsonan. Dalam definisi tersebut tidak disebutkan adanya pengulangan dengan penambahan leksem atau infiks.

Dwilingga salin suara dengan perubahan fonem vokal dapat dikelompokkan lebih lanjut ke dalam dua subkelompok, yaitu (1) dwilingga salin suara dengan perubahan satu fonem vokal dan (2) dwilingga salin suara dengan perubahan dua fonem vokal. Masing-masing subkelompok terdiri dari lima pola perubahan fonem vokal sebagai berikut.

4.2.1 (I) Dwilingga salin suara dengan perubahan satu fonem vokal

(I-1) Perubahan fonem vokal *a-i*. Dwilingga salin suara yang termasuk dalam kelompok ini sangat banyak digunakan. Misalnya, *basa-basi*

- | | |
|------------------------|--------------------------|
| ● <i>basa-basi</i> | ● <i>jungkat-jungkit</i> |
| ● <i>warna-warni</i> | ● <i>selang-seling</i> |
| ● <i>gerak-gerik</i> | ● <i>serba-serbi</i> |
| ● <i>teka-teki</i> | ● <i>kebat-kebit</i> |
| ● <i>pernak-pernik</i> | ● <i>cipika-cipiki</i> |
| ● <i>ketar-ketir</i> | ● <i>dansa-dansi</i> |
| ● <i>cengar-cengir</i> | ● <i>derak-derik</i> |
| ● <i>kelap-kelip</i> | ● <i>kempas-kempis</i> |
| ● <i>wira-wiri</i> | ● <i>pringas-pringis</i> |

Unsur-unsur pembentuk dwilingga salin suara *cipika-cipiki* terdiri atas tiga suku kata dan merupakan gabungan akronim *cipika* (*cium pipi kanan*) dan *cipiki* (*cium pipi kiri*).

(I-2) Perubahan fonem vokal *a-u*. Misalnya,

- *desas-desus*
- *celingak-celinguk*
- *unggah-ungguh*
- *plintat-plintut*
- *runtang-runtung*
- *gedebak-gedebuk*
- *lekak-lekuk*

(I-3) Perubahan fonem vokal *a-o*. Misalnya,

- *ceplas-ceplos*
- *gambar-gembor*
- *lenggak-lenggok*
- *mencang-mencong*
- *megal-megol*
- *plonga-plong*

(I-4) Perubahan fonem vokal *a-e*. Misalnya,

- *corat-coret*
- *mencla-mencle*

(I-5) Perubahan fonem vokal *i-a*. Misalnya,

- *plin-plan*
- *zig-zag*

4.2.2 (II) Dwilingga salin suara dengan perubahan dua fonem vokal

(II-1) Perubahan fonem vokal *o a - a i*. Misalnya,

- | | |
|--------------------------|------------------------|
| ● <i>bolak-balik</i> | ● <i>gonta-ganti</i> |
| ● <i>mondar-mandir</i> | ● <i>otak-atik</i> |
| ● <i>pontang-panting</i> | ● <i>kolang-kaling</i> |
| ● <i>komat-kamit</i> | ● <i>obrak-abrik</i> |
| ● <i>kocar-kacir</i> | ● <i>orak-arik</i> |
| ● <i>compang-camping</i> | ● <i>kopat-kapit</i> |
| ● <i>morat-marit</i> | ● <i>orang-aring</i> |
| ● <i>gonjang-ganjing</i> | ● <i>ropak-rapik</i> |

(II-2) Perubahan fonem vokal *a a - u u*. Misalnya,

- *asal-usul*
- *kasak-kusuk*
- *grasah-grusuh*
- *gerabak-gerubuk*

(II-3) Perubahan fonem vokal *u a - a i*. Misalnya,

- *utak-atik*
- *hubar-habir*

(II-4) Perubahan fonem vokal *u a - a u*. Misalnya,

- *luntang-lantung*

(II-5) Perubahan fonem vokal beraturan lainnya. Misalnya,

- *huru-hara*
- *tiki-taka*
- *petantang-petenteng*
- *lontang-lantung*
- *wara-wiri*
- *gono-gini*
- *cekakak-cekikik*

Dari daftar pengelompokan di atas, dapat dilihat bahwa dwilingga salin suara dengan perubahan fonem vokal *a-i* dan *o a - a i* termasuk produktif.

Dwilingga salin suara dengan perubahan fonem konsonan dapat dikelompokkan ke dalam dua subkelompok, yaitu (1) dwilingga salin suara dengan unsur kedua berawalan konsonan *m* dan (2) dwilingga salin suara dengan unsur kedua berawalan konsonan selain *m*, sebagai berikut.

(2-1) Unsur kedua berawalan konsonan *m*. Misalnya,

- *serta-merta*
- *sayur-mayur*
- *karut-marut*
- *carut-marut*
- *coreng-moreng*
- *kawin-mawin*

(2-2) Unsur kedua berawalan konsonan selain *m*. Misalnya,

- *seluk-beluk*
- *lauk-pauk*
- *remeh-temeh*
- *ramah-tamah*

Dapat dilihat bahwa unsur kedua dwilingga salin suara dengan perubahan fonem konsonan kebanyakan berawalan konsonan *m*. Dibandingkan dengan bentukan yang memiliki perubahan fonem vokal, dwilingga salin suara dengan perubahan fonem konsonan dapat dikatakan tidak produktif.

Dari segi maknanya, dwilingga salin suara menyatakan pluralitas (baik nomina dalam hal kuantitas maupun verba dalam hal repetisi), tingkat intensitas, dan tiruan bunyi yang selanjutnya dapat digolongkan ke dalam enam kelompok sebagai berikut.

1. ‘banyak dan bermacam-macam’. Misalnya,

- | | |
|----------------------|------------------------|
| ● <i>sayur-mayur</i> | ● <i>remeh-temeh</i> |
| ● <i>warna-warni</i> | ● <i>gono-gini</i> |
| ● <i>serba-serbi</i> | ● <i>gerak-gerik</i> |
| ● <i>carut-marut</i> | ● <i>pernak-pernik</i> |
| ● <i>kawin-mawin</i> | ● <i>kelap-kelip</i> |
| ● <i>seluk-beluk</i> | ● <i>dansa-dansi</i> |
| ● <i>lauk-pauk</i> | |

2. ‘banyak dan tidak beraturan’. Misalnya,

- *mencla-mencle*
- *corat-coret*
- *karut-marut*
- *pontang-panting*
- *kucar-kacir*
- *kocar-kacir*
- *carut-marut*
- *coreng-moreng*
- *hubar-habir*
- *pontang-panting*
- *compang-camping*
- *morat-marit*
- *obrak-abrik*
- *orak-arik*
- *ropak-rapik*
- *plin-plan*
- *plintat-plintut*

3. berbagai tiruan bunyi. Misalnya,

- *cekakak-cekikik*
- *gedebak-gedebuk*
- *desas-desus*
- *kasak-kusuk*
- *gerabak-gerubuk*

4. ‘sungguh-sungguh (intensif)’. Misalnya,

- *ramah-tamah*
- *asal-usul*
- *gonjang-ganjing*
- *serta-merta*

5. repetisi, berulang-ulang. Misalnya,

- *celingak-celinguk*
- *zig-zag*
- *tiki-taka*
- *lekak-lekuk*
- *selang-seling*
- *jungkat-jungkit*
- *wara-wiri*
- *utak-atik*
- *otak-atik*
- *bolak-balik*
- *mondar-mandir*
- *komat-kamit*
- *gonta-ganti*
- *kolang-kaling*
- *kopat-kapit*
- *megal-megol*
- *cipika-cipiki*
- *kempas-kempis*
- *kebat-kebit*

6. makna lainnya. Misalnya,

- *kolang-kaling*
- *orang-arang*
- *ungguh-ungguh*
- *basa-basi*
- *teka-teki*

Tabel 3 memuat sepuluh dwilingga salin suara terbanyak di korpus TBIK. Tiga dwilingga salin suara terbanyak adalah *basa-basi* (8,87%), *bolak-balik* (8,53%), dan *asal-usul* (8,47%).

Tabel 3 Sepuluh Dwilingga Salin Suara Terbanyak

No.	Dwilingga Salin Suara	Kelas Kata	Jumlah Kemunculan	Persentase
1	<i>basa-basi</i>	Nomina	290	8,87%
2	<i>bolak-balik</i>	Verba	279	8,53%
3	<i>asal-usul</i>	Nomina	277	8,47%
4	<i>serta-merta</i>	Adverbia	249	7,62%
5	<i>warna-warni</i>	Nomina	167	5,11%
6	<i>gerak-gerak</i>	Nomina	131	4,01%
7	<i>desas-desus</i>	Nomina	111	3,40%
8	<i>teka-teki</i>	Nomina	108	3,30%
9	<i>mondar-mandir</i>	Verba	107	3,27%
10	<i>seluk-beluk</i>	Nomina	93	2,84%

Dwilingga salin suara cenderung muncul dalam ragam sastra dan cerita populer, bukan dalam ragam formal/resmi. Tabel 4 memuat frekuensi kemunculan dwilingga salin suara berdasarkan kategori/sumber. Lima kategori/sumber teratas adalah novel, cerpen, biografi, majalah, dan koran, sedangkan lima kategori/sumber terbawah adalah surat resmi, perundangan, laman resmi, jurnal, dan buku teks.

Tabel 4 Frekuensi Kemunculan Dwilingga Salin Suara berdasarkan Kategori/Sumber

Kategori	Jumlah Kemunculan	Frekuensi Per Sejuta Kata
Novel	844	321,80
Cerpen	709	283,11
Biografi	455	185,68
Majalah	357	146,33
Koran	297	109,96
Populer	192	83,83
Disertasi, tesis, skripsi	161	59,55
Buku teks	124	50,24
Jurnal	64	26,55
Laman resmi	35	14,34
Perundangan	14	5,60
Surat resmi	11	4,51
Total	3.263	108,87

Tidak ada perbedaan frekuensi kemunculan dwilingga salin suara berdasarkan tahun. Antara tahun 2011 hingga tahun 2020 dwilingga salin suara muncul dalam data korpus sebanyak 300 hingga 400 kata per tahun.

4.3 Trilingga

Penulis mengumpulkan 13 trilingga dari KBBI V, TBBBI, dan korpus TBIK. Semua trilingga tersebut berkelas kata nomina. Selain berkelas kata nomina, trilingga *dag-dig-dug* dan *pat-pet-pot* juga berkelas kata adjektiva. Dari segi bentuknya, 13 trilingga tersebut dapat dikelompokkan berdasarkan perubahan fonem vokalnya sebagai berikut.

- (1) Perubahan fonem vokal *a-i-u*. Misalnya,
 - *ba-bi-bu*
 - *cas-cis-cus*
 - *dag-dig-dug*
 - *tak-tik-tuk*
- (2) *tang-ting-tung* Perubahan fonem vokal *a-e-o*. Misalnya,
 - *dar-der-dor*
 - *pat-pet-pot*
- (3) Perubahan fonem vokal *a-i-o*. Misalnya,
 - *ngak-ngik-ngok*

- (4) Perubahan fonem vokal *a-o-e*. Misalnya,
 - *ceplak-ceplok-ceplek*
- (5) Perubahan fonem vokal *a-o-o*. Misalnya,
 - *grak-grok-grok*
- (6) Perubahan fonem vokal *o-a-o*. Misalnya,
 - *grok-grak-grok*
- (7) Perubahan fonem vokal *i-o-i*. Misalnya,
 - *ping-pong-ping*
- (8) Perubahan fonem vokal *o-i-o*. Misalnya,
 - *pong-ping-pong*

Dapat dilihat bahwa ada kecenderungan penggunaan fonem vokal *a* pada unsur pertama, fonem vokal *i* atau *e* pada unsur kedua, dan fonem vokal *u* atau *o* pada unsur ketiga.

Dari segi maknanya, sebagian besar bermakna tiruan bunyi atau sesuatu yang berhubungan dengan bunyi, seperti *dag-dig-dug* yang bermakna ‘tiruan bunyi debaran jantung’, *dar-der-dor* yang bermakna ‘tiruan bunyi tembakan’, dan *ngak-ngik-ngok* yang bermakna ‘ingar bingar (tentang musik)’. Trilingga *dag-dig-dug* mempunyai perluasan makna ‘berdebar-debar (tentang jantung) karena sangat takut, cemas, bersemangat, dan sebagainya’. Beberapa trilingga memiliki makna yang berhubungan dengan bicara atau cakap, seperti *ba-bi-bu* yang bermakna ‘bicara sesuatu’ dan *cas-cis-cus* yang bermakna ‘banyak cakap (omong)’ atau ‘lancar bercakap dalam bahasa asing’. KBBI V memberi label ragam cakapan pada *ngak-ngik-ngok*, *ba-bi-bu*, dan *cas-cis-cus*.

Dari 13 trilingga yang berhasil dikumpulkan, diperoleh sebanyak 10 trilingga yang ada di korpus TBIK. Tabel 5 memuat frekuensi dan jumlah kemunculan trilingga tersebut. Tiga trilingga terbanyak adalah *ba-bi-bu*, *dag-dig-dug*, dan *ping-pong-ping*. Trilingga *ba-bi-bu* dan *dag-dig-dug* masing-masing muncul sebanyak 6 kali, sedangkan *ping-pong-ping* hanya dua kali.

Tabel 5 Sepuluh Trilingga Terbanyak

No.	Trilingga	Kelas Kata	Jumlah Kemunculan	Persentase
1	<i>ba-bi-bu</i>	Nomina	6	30,00%
2	<i>dag-dig-dug</i>	Nomina	6	30,00%
3	<i>ping-pong-ping</i>	Nomina	2	10,00%
4	<i>pong-ping-pong</i>	Nomina	1	5,00%
5	<i>ceplak-ceplok-ceplek</i>	Nomina	1	5,00%

6	<i>dar-der-dor</i>	Nomina	1	5,00%
7	<i>grak-grok-grok</i>	Nomina	1	5,00%
8	<i>grok-grak-grok</i>	Nomina	1	5,00%
9	<i>pat-pet-pot</i>	Adjektiva	1	5,00%
10	<i>tang-ting-tung</i>	Nomina	1	5,00%

Trilingga *ba-bi-bu* selalu didahului oleh *tanpa*, seperti dalam contoh berikut ini.

- (15) Entah ikut tegang dengan cerita ini, si Latino tiba-tiba terdiam. Lalu tanpa *ba-bi-bu*, dia menutup telepon. (Novel_D1A13001)

Trilingga *dag-dig-dug* yang bermakna tiruan bunyi hanya muncul satu kali, seperti dalam contoh (16).

- (16) Mendadak, degup di jantung Bi Mar terasa tak normal. Ada *dag-dig-dug* yang tak biasa. (Cerpen_C1A11017)

Makna *dag-dig-dug* yang sering muncul adalah ‘berdebar-debar’, seperti dalam contoh (17).

- (17) Kuliah MENJELANG Dunia dalam Berita, hati Luru makin *dag-dig-dug*. (Cerpen_C1A15117)

Trilingga *dar-der-dor* hanya muncul satu kali dengan makna ‘menggunakan pistol’ atau ‘menembak’, bukan tiruan bunyi tembakan, seperti dalam contoh (18).

- (18) Opa meminjamiku pistol tua, memberikan penutup telinga, lantas kami sibuk *dar-der-dor* di halaman kanan rumah yang disulap jadi tempat latihan tembak dadakan. (Novel_D1A12002)

Trilingga lainnya bermakna tiruan bunyi, misalnya tiruan bunyi bola ping-pong yang membentur meja (19), tiruan bunyi sol sepatu lars (20), tiruan suara babi (21), tiruan bunyi orang yang memainkan alat musik (22), dan tiruan bunyi bel kereta (23).

- (19) Soekram kadang-kadang masih mendengar suara bola ping-pong yang membentur meja kalau ia berangkat tidur malam-malam, padahal anak-anak sudah tidak main lagi.
Ping-pong-ping pong-ping-pong ping-pong-ping. (Novel_D1A15004)
- (20) Polisi Baplang berjalan mondar-mandir di hadapan para waria, langkahnya sengaja pelan, dibuat berat biar terkesan mengancam. Sol sepatu larsnya berdecit-decit, menimbulkan bunyi *ceplak-ceplak-ceplek* saat ia hilir mudik di tanah yang becek. (Novel_D1A19003)
- (21) Babi Lumpur menoleh ke belakang. Ia mengeluarkan suara *grok-grak-grak*, Babi Muda menjawab dengan *grak-grok-grok*. (Novel_D1A19003)
- (22) Waktu aku masih belasan tahun, aku sering datang ke sini, berkemah, memancing, berburu, mengebut dengan speedboat, atau sekadar bengong duduk di beranda dermaga, menatap senja bersama Opa yang *pat-pet-pot* memainkan alat musik. (Novel_D1A12002)
- (23) Bunyi *tang-ting-tung* dan suara seseorang di pengeras suara menandakan keretaku segera datang. (Cerpen_C1A16090)

Trilingga hanya muncul pada kategori majalah, cerpen, dan novel, seperti yang dicantumkan dalam Tabel 6 yang memuat frekuensi kemunculan trilingga berdasarkan kategori/sumber.

Tabel 6 Frekuensi Kemunculan Trilingga berdasarkan Kategori/Sumber

Kategori	Jumlah Kemunculan	Frekuensi Per Sejuta Kata
Novel	11	4,19
Cerpen	6	2,40
Majalah	3	1,23
Total	20	2,64

Dari segi frekuensi kemunculan berdasarkan tahun, trilingga paling banyak muncul pada tahun 2011 sebanyak lima kali dalam majalah, cerpen, dan novel. Selebihnya, trilingga hanya muncul sebanyak satu hingga empat kali, kecuali pada tahun 2014 dan 2020 trilingga tidak muncul sama sekali.

5. SIMPULAN

Dibandingkan dengan kata ulang utuh atau dwilingga, baik dwipurwa, dwilingga salin suara, maupun trilingga tidak produktif. Meskipun demikian, dwipurwa lebih banyak digunakan daripada dwilingga salin suara dan trilingga. Selain itu, dibandingkan dengan dwilingga salin suara dan trilingga, ada kreasi baru dalam dwipurwa seperti *gegara*, *jejaring*, dan *tetikus*. Dari segi kesamaan maknanya, baik dwipurwa, dwilingga salin suara, maupun trilingga sama-sama menyatakan pluralitas. Dwipurwa, dwilingga salin suara, dan trilingga cenderung digunakan dalam ragam sastra (novel dan cerpen) dan takresmi/nonformal (majalah dan bacaan populer).

Referensi

- [1] Moeliono, A. M., Lapoliwa, H., Alwi, H., & Sasangka, S. S. T. W. (2017). *Tata Bahasa Baku Bahasa Indonesia Edisi Keempat*. Jakarta: Badan Pengembangan dan Pembinaan Bahasa.
- [2] Nomoto, Hiroki, Hannah Choi, David Moeljadi and Francis Bond. (2018). MALINDO Morph: Morphological dictionary and analyser for Malay/Indonesian. Kiyooki Shirai (ed.) *Proceedings of the LREC 2018 Workshop "The 13th Workshop on Asian Language Resources"*, 36-43.
- [3] Goldhahn, Dirk, Thomas Eckart and Uwe Quasthoff (2012): Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 2012.
- [4] Robson S. and Wibisono S. (2002). *Javanese-English Dictionary*.
- [5] Kulsum, Umi (2015). Dwipurwa dan Potensinya dalam bahasa Indonesia. *Sawerigading*, 21(3), 391-403.
- [6] Kridalaksana, H. (2009). *Pembentukan Kata dalam Bahasa Indonesia*. Jakarta: Gramedia Pustaka Utama.
- [7] Sariah. (2019). Dwilingga Salin Suara dalam bahasa Indonesia. *Widyaparwa*, 46(2), 126-144.
- [8] Denistia, K., & Baayen, R. H. (2022). The morphology of Indonesian: Data and quantitative modeling. In Shei, C., and Li, S. (Eds.), *The Routledge Handbook of Asian Linguistics* (pp. 605-634). London: Routledge.

- [9] Sumarsih, N. (2013). Tipe-Tipe Reduplikasi Semantis Bahasa Indonesia: Kajian Bentuk dan Makna. *Widyaparwa*, 41(01), 81-90.
- [10] Hardie, A. (2012). CQPweb—combining power, flexibility and usability in a corpus analysis tool. *International journal of corpus linguistics*, 17(3), 380-409.

AKRONIM DALAM BAHASA INDONESIA

Nazarudin

Universitas Indonesia

Abstrak

Akronim merupakan salah satu dari beberapa proses pembentukan kata yang cukup produktif dalam Bahasa Indonesia. Di dalam buku-buku tentang tata bahasa Indonesia, ternyata tidak semua mencantumkan proses pemendekan kata atau akronim dan abreviasi. Menurut beberapa peneliti (Soenjono, 1979, Smith-Hefner, 2010), di dalam Bahasa Indonesia sendiri, penutur Bahasa Indonesia bertendensi untuk menyamakan antara akronim dan abreviasi. Artikel ini bertujuan untuk mendeskripsikan bagaimana pola pembentukan singkatan atau akronim dalam Bahasa Indonesia. Pola pembentukan akronim tersebut diperoleh berdasarkan data-data dari korpus Tata Bahasa Indonesia Kontemporer (TBIK) yang termasuk ke dalam korpus umum.

Kata kunci: akronim, abreviasi, bahasa Indonesia, linguistik korpus

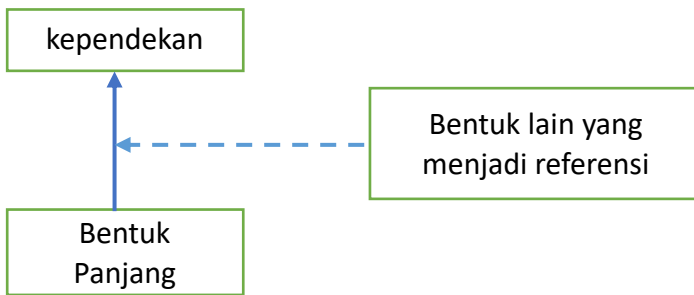
1. PENDAHULUAN

Proses akronim¹ dapat ditemukan dalam banyak Bahasa di dunia. Untuk Bahasa Indonesia, khususnya, akronim memiliki fungsi yang unik dan penting dalam peta perkembangan Bahasa Indonesia. Meskipun demikian, akronim dalam Bahasa Indonesia baru populer dan produktif setelah tahun 60-an. Dalam akademi militer di Indonesia, misalnya, para taruna diajarkan bagaimana menggunakan akronim-akronim yang dipakai dalam dunia militer sebagai bagian dari proses pembelajaran mereka (Departemen Angkatan Darat, 1968). Setelah itu, penggunaan akronim pun semakin meluas karena masyarakat mulai banyak yang mengakronimkan kata-kata dengan bebas. Bahkan, menurut Soenjono (1979), dalam Bahasa Jawa pada saat itu muncul akronim *pentil kecakot* sebagai akronim dari “*penilik tilpun kecamatan kota*”. Hal ini menjadi semacam bentuk kreativitas penutur dan juga permainan Bahasa oleh penutur karena akronim setiap kata dari “pentil kecakot” mengandung makna

1 Dalam artikel ini, penulis menggunakan istilah “akronim”, “singkatan”, dan “abreviasi” secara bergantian secara bersinonim dengan pertimbangan bahwa jenis-jenis pemendekan kata lainnya dianggap termasuk ke dalam ketiga istilah tersebut.

tersendiri dalam Bahasa Jawa, yaitu ‘pentil tergigit’.

Kridalaksana (2010) menyatakan bahwa kependekan yang berupa gabungan huruf atau suku kata atau bagian kata yang ditulis dan dilafalkan sebagai kata sesuai dengan kaidah fonotaktik sebuah bahasa disebut akronim. Di sisi lain, Bauer (1988) mendefinisikan akronim dengan ‘kata yang diciptakan dari huruf awal nama kata, nama judul, atau frasa karena kata ciptaan itu sebenarnya dilafalkan seperti kata baru’. Kridalaksana (2010) dalam hal ini memasukkan akronim ke dalam salah satu bagian dari abreviasi. Selain itu, dalam bukunya, Kridalaksana juga menekankan bahwa sering kali kependekan itu muncul bukan hanya sebatas proses morfologis, melainkan juga ada aspek asosiasi makna yang terlibat di dalamnya. Misalnya, Proses pembentukan tersebut dalam dilihat melalui bagan berikut ini.



Akronim merupakan salah satu dari beberapa proses pembentukan kata yang cukup produktif dalam Bahasa Indonesia. Bahkan, Sutan Takdir Alisjahbana pernah menyatakan kekhawatirannya seperti yang ditulisnya dalam harian Suara Karya, Senin, 9 Desember 1985. Dalam artikel tersebut, Beliau menuliskan, “Banjir akronim dan kependekan sekarang ini yang membuat kita sukar membaca surat kabar harus dibendung. Bahasa Indonesia oleh karenanya akan menjadi amat sulit, sebab tiap kependekan merupakan bentuk yang baru, sedangkan tidak ada isi baru di bawahnya, ingatan kita terlampau dibenahinya.”

Singkatan merupakan salah satu bentuk dari proses abreviasi yang berupa huruf atau gabungan huruf, misalnya *KKN* (Kuliah Kerja Nyata), *UNM* (Universitas Negeri Malang), *OSIS* (organisasi Siswa Intra Sekolah), *dll* (dan lain-lain), dan *dst* (dan seterusnya). Penggalan merupakan salah satu bentuk dari proses abreviasi yang berupa pengekalan atau pemenggalan sebagian unsur dalam kata, misalnya *Prof* (Profesor), *Bu* (Ibu), *Pak* (Bapak), *perpus*

(perpustakaan). Unsur bahasa dalam kata yang dipenggal dapat berupa fonem atau suku kata.

Artikel ini bertujuan untuk mendeskripsikan bagaimana pola pembentukan singkatan atau akronim dalam Bahasa Indonesia. Untuk mencapai tujuan tersebut, artikel ini disampaikan dalam beberapa bagian dan dimulai dengan bagian pendahuluan yang menjelaskan tentang tujuan penelitian ini. Kemudian, pada bagian 2 dilanjutkan dengan memaparkan semacam sintesis dari beberapa penelitian terdahulu yang sudah ada. Pada bagian 3 dijelaskan tentang metodologi penelitian yang digunakan di dalam artikel ini. Sementara itu, bagian 4 berisi tentang kategorisasi akronim menurut beberapa ahli dan ditambah dengan penjelasan kategorisasi mana yang akan digunakan dalam artikel ini. Bagian 5 berisi tentang analisis pola akronim dalam Bahasa Indonesia yang kemudian juga di dalamnya disajikan contoh-conton yang relevan yang didapat dari korpus Tata Bahasa Indonesia Kontemporer (TBIK). Bagian 6 dalam tulisan ini menunjukkan penggunaan akronim berdasarkan ranah penggunaan yang ada dalam korpus data TBIK di CQPWeb. Artikel ini ditutup dengan diskusi singkat dan kesimpulan yang disampaikan pada bagian 7.

2. PENELITIAN TERDAHULU

Di dalam buku-buku tentang tata bahasa Indonesia, ternyata tidak semua mencantumkan proses pemendekan kata atau akronim dan abreviasi. Contohnya, dalam buku Tata Bahasa Baku Bahasa Indonesia, tidak tercantum sub-bab tentang akronim dalam Bahasa Indonesia. Dalam Tata Bahasa Baku Bahasa Indonesia (2017), tidak terdapat penjabaran yang komprehensif tentang singkatan atau akronim dan abreviasi. Selain itu, dalam buku Tata Bahasa Praktis Bahasa Indonesia yang ditulis oleh Abdul Chaer (2000), misalnya, juga tidak disebutkan tentang akronim.

Acuan-acuan yang mencantumkan sedikit tentang akronim atau pemendekan kata di antaranya adalah Kridalaksana (2010). Harimurti Kridalaksana (2010) mengklasifikasikan gejala bahasa hasil pemendekan itu ke dalam lima kategori, yakni (a) singkatan, (b) penggalan, (c) akronim, (d) konstruksi, dan (e) lambang huruf. Sementara itu, Notosusanto (1979) dan JS Badudu (1983) membaginya menjadi dua, yakni singkatan dan akronim.

Menurut beberapa peneliti (Soenjono, 1979, Smith-Hefner, 2010), di

dalam Bahasa Indonesia sendiri, penutur Bahasa Indonesia bertendensi untuk menyamakan antara akronim dan abreviasi. Dalam artikelnya, Abidin (2012), misalnya, juga mengkategorikan SVM, yang merupakan kependekan dari *Support Vector Machine* dalam bidang ilmu komputer, sebagai sebuah akronim, meskipun dilafalkan sesuai dengan fonem-fonem dari huruf pembentuknya. Selain itu, dalam artikelnya, Ujian Nasional yang disingkat menjadi UN juga disebut sebagai sebuah akronim, terlepas dari apakah singkatan tersebut dilafalkan huruf per huruf dan tidak dilafalkan sebagaimana melafalkan kata secara umum.

Sebenarnya perdebatan tentang istilah-istilah ini sudah ada sejak lama. Kridalaksana (2010), misalnya, menyatakan bahwa akronim merupakan bagian dari abreviasi. Isa (2006) dalam artikelnya membahas bahwa bentukan leksem atau ciptaan kata baru yang berbeda-beda ditentukan oleh jenis abreviasinya. Abreviasi, menurut Bauer (1988: 39), dapat dibagi menurut jenisnya menjadi dua bagian, yaitu paduan (*blends*) dan akronim (*acronyms*). Akan tetapi, Kridalaksana (2010) memasukkan akronim, kontraksi (*contractions*), lambang huruf, pemenggalan (*clippings*), dan penyingkatan sebagai abreviasi. Selain itu, Kridalaksana (2010) juga memerikan jenis abreviasi lain, seperti paduan dan paduan pinjam (*loan blends*). Sementara itu, abreviasi, menurut Matthews (1997: 1), dibagi menjadi akronim, paduan, dan pemenggalan. Di sisi lain, Haspelmath (2002: 25) hampir sependapat dengan Kridalaksana (2010) dalam pembagian abreviasi, yaitu akronim, alfabetisme, pemenggalan, dan paduan. Sementara itu, Beard (1998: 56) membagi jenis abreviasi itu menjadi pemenggalan, paduan, akronimisasi, dan formasi analogis. Beard (1998: 56) juga menganggap bahwa keempat proses pembentukan kata itu sebagai ekspansi stok leksikal (*lexical stock expansion*), yang semuanya sesuai dengan deskripsi derivasi balik di dalam cara yang signifikan.

3. METODOLOGI

Terlepas dari pelbagai perdebatan yang disampaikan pada bagian sebelum ini, artikel ini bertujuan untuk mendeskripsikan bagaimana pola pembentukan singkatan atau akronim dalam Bahasa Indonesia. Pola pembentukan akronim tersebut diperoleh berdasarkan data-data dari korpus Tata Bahasa Indonesia Kontemporer (TBIK) yang termasuk ke dalam korpus umum. Korpus ini bukanlah korpus spesialis yang datanya hanya diperoleh dari satu ranah

saja, namun data korpus TBIK diperoleh dari berbagai domain seperti koran, cerpen, perundangan, dan sebagainya yang akan kita bahas pada bagian selanjutnya. Korpus TBIK, seperti juga Longman Corpus, juga digunakan sebagai sumber data buku TBIK ini. Namun, perlu juga digarisbawahi dalam bagian pendahuluan ini, akronim atau singkatan bukannya satu topik yang mudah untuk diteliti melalui pendekatan linguistik korpus. Linguistik korpus dalam penelitian ini hanya digunakan untuk mencari frekuensi kemunculan contoh akronim dalam sebuah konteks.

Artikel ini menggunakan gabungan dari pendekatan kualitatif dan kuantitatif. Pendekatan kualitatif dalam artikel ini digunakan untuk menganalisis pola pembentukan akronim yang ada dalam data. Namun, sebelumnya, penulis mensintesisakan kategorisasi yang sudah disusun oleh beberapa peneliti, seperti Soenjono (1979), Kridalaksana (2010), dan juga Isa (2006). Setelah itu, penulis mencari contoh-contoh akronim yang ada dalam korpus TBIK untuk mengecek frekuensi kemunculan dan juga melihat bagaimana penggunaannya dalam konteks (KWIC—Key Words in Context). Dengan demikian, kita dapat melihat deskripsi penggunaan akronim dalam Bahasa Indonesia berdasarkan kemunculannya pada data kontemporer.

Korpus ini dapat diakses melalui CQPWeb. CQP sendiri berasal dari corpus query processor yang berbasiskan web yang dikelola oleh Lancaster University di Inggris. Seluruh data korpus yang berhasil dihimpun oleh Badan Bahasa kemudian diunggah ke dalam web tersebut agar dapat diakses dengan lebih mudah. Pada data ini juga telah dilakukan POS Tag sederhana untuk mengenali kelas kata dari kosakata yang ada dalam data korpus.

Dalam korpus ini juga dapat dilihat pesebaran data per tahun sejak tahun 2011—2020. Pesebaran data selama lebih kurang satu dekade ini dianggap dapat merepresentasikan penggunaan bahasa tulis di masyarakat dalam 10 tahun terakhir. Data yang digunakan dalam korpus juga bervariasi, mulai dari data jurnalistik, tulisan ilmiah, hingga data terbitan populer yang ada dalam sepuluh tahun terakhir.

4. PEMBAHASAN

4.1 Jenis Akronim dalam Bahasa Indonesia

Ada beberapa perdebatan di kalangan linguis mengenai bagaimana orang Indonesia menyamakan istilah abreviasi dengan hanya sebutan akronim saja. Isa (2006), misalnya, menganggap bahwa kekeliruan informasi dan definisi

yang disajikan selama ini. Dirinya juga memperlihatkan hal itu melalui penelitiannya yang dilakukan dengan menggunakan data abreviasi yang didapat dari dua koran besar di Indonesia, yaitu KOMPAS dan Jakarta Post (Isa, 2006). Selain itu, dirinya juga mendukung definisi Bauer's yang menyatakan bahwa "*acronyms are words made up of initial letters of words, title, or phrase which are then pronounced as if they are new words*" (Bauer, 1988).

Isa juga menjelaskan untuk singkatan yang tidak mencakup semua huruf awal dari semua frasa panjang seperti BSc, RAPBN, TOEFL, dan U-RAISE. Dua contoh terakhir memang dapat diucapkan sebagai kata biasa dan dia mendefinisikannya sebagai jenis singkatan akronim. Untuk dua huruf pertama yang membutuhkan setiap huruf untuk dibunyikan, mereka adalah jenis singkatan 'alphabetism', yang kemungkinan besar termasuk ke dalam 'initialism' dalam kategorisasi Isa. Mungkin, sekilas tampak mengesankan bahwa Isa berhasil mengklasifikasikan proses kreatif pembentukan kata-kata baru dengan singkatan yang digunakan orang Indonesia menjadi 16 jenis, setidaknya satu di antaranya tampak campur aduk atau tidak perlu dikategorikan sebagai jenis yang berbeda. Misalnya saja, penjelasan untuk 'sinkronim' hampir identik dengan 'campuran'.

Selain itu, penggunaan berbagai istilah untuk menyebut singkatan, abreviasi, dan akronim ini juga seringkali masih dianggap sama oleh pengguna Bahasa Indonesia. Dengan pemikiran ini, jelas bahwa orang Indonesia mengambil kebebasan dalam pembentukan kata-kata. Ini adalah praktik yang sangat viral di media tentang politik atau bahkan hanya untuk tujuan komedi. Sangat umum bagi orang Indonesia untuk membuat akronim menggunakan huruf apa pun yang tersedia dari nama-nama kandidat pemerintah untuk meningkatkan keakraban dan viralitas, seperti nama Syahrul Yasin Limpo dan Agus Arifin Nu'mang disingkat menjadi Sayang, yaitu sebenarnya merupakan kata biasa yang sudah mapan yang dimaksudkan sebagai istilah *sayang* (Rijal, 2015).

Orang Indonesia juga sering sengaja membuat versi kata yang lebih panjang dengan menggunakan kata biasa yang sudah ada sehingga menjadi singkatan. Bupati adalah sebutan untuk kepala daerah dalam bahasa Indonesia, tetapi kemudian dalam konteks humor kata tersebut menjadi akronim yang bernada nyinyir untuk *buka paha tinggi-tinggi*, yang mengacu pada orang yang menunjukkan bagian paha yang terbuka. Selanjutnya, orang Indonesia akan memberikan definisi baru untuk akronim yang sudah dikenal seperti Golput (Golongan Putih) menjadi *Gerakan orang kesal dan putus asa*, atau Orba (Orde Baru) menjadi *Orang-orang banyak akal*.

Dengan demikian, kita dapat melihat alasan mengapa orang Indonesia hanya menggabungkan semua jenis singkatan sebagai akronim. Kemungkinan juga karena mereka secara teratur berlatih membuat banyak kata menjadi kata-kata baru yang mudah diucapkan dan diingat. Karena praktik yang sangat umum ini, orang awam tidak akan memperhatikan diri mereka sendiri untuk memastikan bahwa mereka menggunakan istilah spesifik yang benar untuk “struktur internal dari potensi kompleks kata-kata suatu bahasa” (Aronoff & Anshen, 2017). Oleh karena itu, bagi orang Indonesia, ‘akronim’ mungkin lebih merupakan istilah yang lebih akrab daripada ‘abreviasi’, meskipun mereka yang mempelajarinya memahami bahwa ‘abreviasi’ merupakan istilah umum untuk berbagai jenis pembentukan kata.

Dalam subbab ini, penulis menjelaskan beberapa kategorisasi yang sudah disampaikan oleh para peneliti Bahasa. Kategorisasi pertama diambil dari kajian yang sudah dilakukan oleh Kridalaksana (2010). Dalam kajiannya, Kridalaksana (2010: 159), memberikan definisi bahwa abreviasi adalah proses penanggalan beberapa bagian leksem atau kombinasi leksem sehingga jadilah bentuk baru yang berstatus kata. Dalam proses ini, leksem atau gabungan leksem menjadi kata kompleks atau akronim atau singkatan dengan pelbagai abreviasi, yaitu dengan pemenggalan, kontraksi, akronimi, dan penyingkatan.

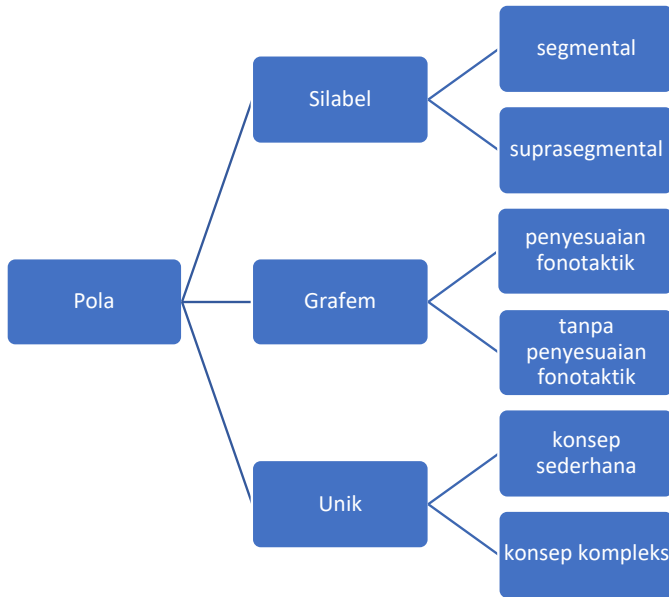
- a. *Singkatan*, merupakan salah satu hasil proses pemendekan yang berupa huruf atau gabungan huruf, baik yang cara membacanya dieja huruf demi huruf maupun yang tidak. Contoh: *Kuliah Kerja Nyata (KKN)*, *Sekolah Dasar (SD)*.
- b. *Penggalan*, yaitu proses pemendekan yang mengekalkan salah satu bagian dari leksem. Contoh: *Profesor (Prof)*, *Ibu (Bu)*.
- c. *Akronim*, merupakan proses pemendekan yang menggabungkan huruf atau suku kata atau bagian lain yang ditulis dan dilafalkan sebagai sebuah kata yang sedikit banyak memenuhi kaidah fonotaktik bahasa Indonesia seperti, *SIM (Surat Izin Mengemudi)*, *IKIP (Institut Keguruan dan Ilmu Pendidikan)*.
- d. *Kontraksi*, yaitu proses pemendekan yang meringkaskan leksem dasar atau gabungan leksem seperti, *takkan (tidak akan)*, *rudal (peluru kendali)*, *sendratari (seni drama tari)*.
- e. *Lambang huruf*, yaitu proses pemendekan yang menghasilkan satu huruf atau lebih yang menggambarkan konsep dasar kuantitas, satuan atau unsur, seperti *cm (centimeter)*, *kg (kilogram)*.

Berikutnya, Isa (2006) juga meneliti tentang abreviasi. Dalam risetnya tersebut, meskipun Isa (2006) mengangkat tentang abreviasi dalam Bahasa Inggris, namun argumennya tentang abreviasi dapat kita kutip pula di sini untuk dijadikan sebagai salah satu kajian terdahulu yang cukup penting. Kesimpulan dari hasil risetnya terkait pada proses pembentukan kata dalam level morfologi dan dalam risetnya dinyatakan bahwa abreviasi dapat terdiri dari beberapa jenis berikut ini.

- a. Alfabetisme/Inisialisme: sebuah tipe abreviasi yang menggunakan huruf awal dari seperangkat kata untuk membentuk sebuah kata yang baru yang tidak dapat dilafalkan sebagaimana kata umumnya. Contohnya adalah *Thank God It's Friday* ◇ TGIF.
- b. Akronim: sebuah tipe abreviasi yang menggunakan huruf awal dari seperangkat kata untuk membentuk kata yang baru yang dapat dilafalkan sebagaimana melafalkan kata biasa. Contohnya adalah National Aeronautics and Space Administration yang disingkat menjadi NASA.
- c. Akronim sebagian: sebuah tipe abreviasi yang menggunakan huruf awal dari seperangkat kata dan juga huruf awal dari preposisi dan penghubungnya untuk membentuk kata yang baru yang dapat dilafalkan sebagaimana melafalkan kata biasa. Contohnya adalah *search and rescue* yang disingkat menjadi SAR.
- d. Koinisasi Akronim: sebuah tipe abreviasi yang menggunakan huruf awal untuk menciptakan sebuah kata sebagai sebuah judul/nama (biasanya sebuah produk). Contohnya adalah *central processing unit* yang disingkat menjadi CPU.
- e. Sinkronim: merupakan sebuah tipe abreviasi yang memakai silabel dan akronim, menargetkan kemiripan huruf tertentu yang dimiliki oleh kedua kata (syllabic juncture) sehingga kata tersebut dapat lebih mudah dilafalkan. Contohnya, kata *motor* dan *hotel* yang kemudian disingkat menjadi *motel*.
- f. Campuran: sebuah tipe abreviasi menggunakan bagian kata yang bersifat *non-morphemic* (*not morpheme-based*). Contohnya adalah gabungan kata *information* dan *entertainment* yang kemudian disingkat menjadi *info + tainment* ◇ *infotainment*.
- g. Percampuran Leksikal: sejenis abreviasi yang merupakan penggabungan dari kependekan bagian beberapa kata, seperti *breakfast + lunch* ◇ br- + -unch ◇ *brunch*.

- h. Percampuran bentuk pinjaman: tipe abreviasi yang menggunakan gabungan dari morfem pinjaman dan kata asli. Contoh dalam Bahasa Inggris adalah sebagai berikut. [French] aer(o) + [British] plane ◇ aeroplane.
- i. Percampuran terpisah: tipe abreviasi yang merupakan hasil percampuran dari satu atau beberapa huruf yang merupakan bagian dari kata tersebut. Terkadang bentuk ini terlihat seperti secara sengaja menggunakan huruf-huruf yang jika digabungkan dapat dilafalkan seperti kata. Contohnya adalah Association Southeast Asian Nations ◇ ASEAN, United States of America ◇ USA, Inspektorat Jenderal ◇ Irjen.
- j. Amalgam: suatu tipe abreviasi yang menggunakan dua huruf pertama dari kata pertama dan seluruh huruf dari kata kedua. Contohnya adalah glass + asphalt ◇ gl- + asphalt ◇ glasphalt.
- k. Bentuk Analogikal: suatu tipe abreviasi yang menggunakan derivasi atas bentuk kata secara penuh, seperti work + -aholic ◇ workaholic dan Instagram + -able ◇ instagramable.
- l. Kontraksi: sebuah abreviasi yang diperoleh dari proses pemendekan leksem dasar. Contohnya bisa dilihat dari do not ◇ don't dan tidak ada ◇ tiada.
- m. Simbol huruf: sebuah tipe abreviasi yang menghasilkan satu atau dua huruf untuk mendeskripsikan konsep dalam konteks kuantitas atau elemen. Contohnya adalah gram ◇ g, calcium ◇ Ca.

Di samping pendapat tentang kategorisasi oleh Kridalaksana (2010) dan Isa (2006), perlu juga dilihat kategorisasi lain dari Soenjono (1979). Kategorisasi pola akronim berdasarkan Soenjono (1979) berikut ini dapat dianggap sebagai awal pemetaan pola yang berfokus pada pola pembentukan kata secara fonologis dan morfofonologis (berdasarkan pada pola struktur silabelnya). Pola ini cukup berbeda dengan kategorisasi dari peneliti lain. Ada tiga kategorisasi besar untuk akronim sesuai dengan yang dinyatakan oleh Soenjono (1979). Ketiga kategori tersebut terbagi berdasarkan pola pembentukan dari akronim dalam Bahasa Indonesia.



Sumber: Soenjono (1979)

Gambar 1 Pola Akronim dalam Bahasa Indonesia

Dalam artikelnya, Soenjono (1979) memberikan contoh-contoh dari tiap pola pembentukan akronim tersebut. Misalnya saja, contoh untuk pola pembentukan akronim berpola segmental silabel (suku kata) dapat terbagi lagi menjadi beberapa pola sebagai berikut.

- Suku kata pertama dari bentuk asal, seperti orba (orde baru) dan muker (musyawarah kerja)
- Suku kata terakhir dari bentuk asal, contohnya adalah dandim (komandan kodim)
- KV pertama + K final, misalnya ditjen yang berasal dari direktorat jenderal
- K pertama + V akhir, misalnya kasad dan letda
- K akhir yang mewakili keseluruhan kata, seperti di dalam singkatan kopasgat (komando pasukan gerak cepat)
- Penghilangan sebagian kata, misalnya dalam menlu (menteri luar negeri) dan banser (barisan ansor serbaguna)

Selain ini, disebutkan juga tentang pola yang terbentuk dari unsur suprasegmental. Namun, bentuk seperti ini tidak lazim ditemukan dalam Bahasa Indonesia. Dalam artikelnya, Soenjono (1979) mencontohkan bentuk /han/ yang diambil dari akronim *hankam* merupakan akronim yang muncul karena unsur suprasegmental, mengingat /han/ dalam kata “pertahanan” merupakan suku kata yang mendapatkan stress atau penekanan dalam pelafalannya.

Meskipun Soenjono (1979) sudah menyampaikan tentang pola pembentukan akronim, namun belum ada penjelasan yang cukup lengkap untuk mendeskripsikan pola-pola tersebut. Namun, dalam artikel tersebut sudah diungkit tentang faktor kognitif di balik produktifnya pembentukan akronim di antara penutur Bahasa Indonesia ini. Penciptaan akronim tampaknya hampir secara eksklusif didasarkan pada norma-norma yang secara inheren ada dalam bahasa dan, oleh karena itu, dimiliki oleh anggota komunitas tutur. Kita dapat menganggap ini disebabkan oleh fakta bahwa sebagian besar struktur kata akronim dan jumlah suku kata dalam akronim sangat sesuai dengan bahasa Indonesia. Ini juga merupakan faktor yang membuat penutur asli mengatakan “kedengaran enak di telinga” ketika ditanya mengapa kata tertentu diakronimkan dengan cara tertentu. Singkatan seperti bimas (bimbingan massal), turba, pemilu, terutama dll tentunya juga didasarkan pada prinsip ini.

Dalam beberapa kasus, akronim dibuat sedemikian rupa sehingga juga merupakan kata-kata asli Indonesia—tentu saja, dengan arti yang berbeda. Pemilihan *pe li ta* (pembangunan lima tahun) ‘pembangunan lima tahun’, *Jaya* (Jakarta Raya) ‘Jabodetabek’, *KAMI* (Kesatuan Aksi Mahasiswa Indonesia) ‘Persatuan Pelajar Indonesia’, pasti didasarkan pada kenyataan bahwa *pelita*, *jaya*, dan *kami* memang juga merupakan kata-kata bahasa Indonesia yang masing-masing berarti, ‘lampu’, ‘kemenangan’, dan ‘kita’.

Masih dalam beberapa kasus lain bentuk akronim mungkin telah ditentukan tidak hanya oleh keberadaan kata-kata asli Indonesia tetapi juga oleh nilai-nilai budaya yang berlaku di masyarakat. Kereta mewah berwarna biru yang beroperasi pada malam hari dari Jakarta ke Surabaya, *Biru Malam*, bisa disebut **Bilam*, **Rulam*, atau **Ruma*, yang semuanya mengikuti aturan fonotaktik bahasa. Namun nama resminya adalah *Bima* ‘(harfiah) Biru di Malam Hari’, karena, saya kira, kata ini kebetulan adalah nama pahlawan yang paling kuat secara fisik dalam Mahabharata versi Jawa.

Dapat kita anggap bahwa apa yang oleh penutur asli non-linguis disebut

“enak di telinga” sebenarnya adalah prinsip linguistik yang sangat mendasar yang juga disebut sebagai kompetensi (*competence*). Kompetensi inilah yang memungkinkan orang Indonesia menghasilkan akronim yang terdengar/terlihat bagus. Ketika “enak didengar” dianggap sebagai dasar yang cukup untuk akronimisasi, tentu saja konsekuensinya ada banyak inkonsistensi di mana bentuk lengkapnya disingkat menjadi beberapa akronim yang berbeda. Misalnya saja, kata *militer* menjadi *mil*, *m*, dan *mi* masing-masing seperti di *koramil*, *kodam*, dan *mahmilub*. Sementara pemilihan *mi* alih-alih *mil* dalam *mahmilub* mungkin dipengaruhi oleh kehadiran ganda /l/ yang tidak sepenuhnya asing tetapi sangat jarang. Dengan demikian, tidak ada alasan mengapa militer di *koramil* dan *kodam* harus mengambil dua bentuk yang berbeda, terutama ketika istilah-istilah ini ditemukan oleh sumber yang sama--kantor militer di Jakarta. Akronimnya bisa saja menjadi *koram* (meski ada juga akronim *korem*!) dan *kodam*, atau *koramil* dan *kodamil* –yang kesemuanya mengikuti aturan fonotaktik bahasa, dan sama enaknyanya dengan akronim yang ada. Namun, pada kenyataannya tidak seperti itu.

Setelah memahami berbagai pendekatan dan sudut pandang dari beberapa ahli tersebut, dalam artikel ini akan digunakan gabungan dari kategorisasi Kridalaksana (2010) dan Soenjono (1979). Dalam artikel ini, kategorisasi yang akan digunakan hanya empat kategori karena “singkatan” dan “akronim” digabungkan ke dalam satu kategori saja dalam artikel ini. Dengan demikian, kategorisasi yang akan digunakan dalam penelitian ini adalah sebagai berikut.

- a. *Singkatan/akronim*, merupakan proses pemendekan yang menggabungkan huruf atau suku kata atau bagian lain, baik yang ditulis dan dilafalkan sebagai sebuah kata atau pun baik yang cara membacanya dieja huruf demi huruf yang sedikit banyak memenuhi kaidah fonotaktik bahasa Indonesia seperti, *SIM* (*Surat Izin Mengemudi*), *IKIP* (*Institut Keguruan dan Ilmu Pendidikan*), *Kuliah Kerja Nyata (KKN)*, *Sekolah Dasar (SD)*.
- b. *Penggalan*, yaitu proses pemendekan yang mengekalkan salah satu bagian dari leksem. Contoh: *Profesor (Prof)*, *Ibu (Bu)*.
- c. *Kontraksi*, yaitu proses pemendekan yang meringkaskan leksem dasar atau gabungan leksem seperti, *takkan* (*tidak akan*), *rudal* (*peluru kendali*), *sendratari* (*seni drama tari*).
- d. *Lambang huruf*, yaitu proses pemendekan yang menghasilkan satu huruf atau lebih yang menggambarkan konsep dasar kuantitas, satuan atau unsur, seperti *cm* (*centimeter*), *kg* (*kilogram*).

Berdasarkan beberapa pendapat tersebut di atas, dapat disimpulkan bahwa pembentukan akronim Bahasa Indonesia paling tidak melewati tiga tahapan seperti pada gambar berikut ini.



4.2 Pola Akronim dalam Bahasa Indonesia

Pada subbab ini, penulis akan memperlihatkan pola yang cukup umum ditemukan dalam akronim di Indonesia. Selain itu, penulis juga akan menunjukkan bagaimana akronim digunakan dalam korpus TBIK. Dengan demikian, dengan melihat penggunaan akronim dalam korpus, kita jadi bisa mendapat gambaran mengenai frekuensi kemunculan, kolokasi, dan produktivitas akronim tersebut. Namun, sebelum melihat penggunaan akronim, ada beberapa generalisasi yang dapat kita buat mengenai pola bentuk kata-kata akronim. Pertama, bentuk-bentuk kanonik bahasa Indonesia yang relatif sederhana tetapi inheren pasti memberikan tekanan struktural pada bentuk kata-kata akronim. Sebagian besar akronim dalam bahasa Indonesia saat ini dihasilkan dari kombinasi dua ini: KV, KVK, VK dan V. Beberapa contoh akronim yang berasal dari dua kata dapat dilihat sebagai berikut.

- a. KV + KVK : *muker* dari *musyawarah kerja*, *dubes* dari *duta besar*, *caper* dari *calon perwira*, dan juga *caper* dari *cari perhatian*, serta *baper* dari *bawa perasaan*.

Jika kita melihat ke penggunaan akronim *caper*, dalam korpus TBIK yang digunakan ternyata tidak muncul bentuk akronim *caper* dari *calon perwira*.

Seluruh hasil yang muncul adalah akronim caper yang berasal dari cari perhatian. Hal itu terlihat dari tabel berikut ini yang menunjukkan ada 20 penggunaan akronim caper dalam korpus.

No.	Query result	No. of occurrences	Percent
1	caper_NN	19	95.00%
2	caper_VB	1	5.00%

No.	Query result	No. of occurrences	Percent
1	dubes_NNP	139	76.80%
2	Dubes_NN	41	22.65%
3	dubesnya_NN	1	0.55%

Your query "parpol*" returned 458 matches in 88 different texts (in 29,087,513 words (17,277 texts); frequency: 15.373 instances per million words), reduced to results where query node matches word-tag combinations: parpol_NN (440 hits)

No.	Text	Solution 1 to 50	Page 1 / 9
1	A12211017	politik yang kuat sebagai elemen...terutama kalangan pemerintah...parlemen...parpol	dan lain sebagainya...Ka renu itu...berdasarkan KKR
2	A12211061	Sabri (1971) ia juga meminta partai politik (parpol) lain untuk berani menanggung jika me...	lain untuk berani menanggung jika me...yang mereka dirugikan...Dengan
3	A12211079	dapat dia duk di kansi DPRD karena " me rumpes " lak parpol	lain sebagai akibat biko- nya partai beranggapan dari anbang ba tau...
4	A18211010	Penulis Legislatif Papua Barat mulai berubah...Maka di tingkat provinsi pengusaan parpol	nasionalis masih tinggi...di tingkat kota/kabupaten parpol kecil dan parpol Islam
5	A18211010	di tingkat provinsi pengusaan parpol nasionalis masih tinggi...di tingkat kota/kabupaten	kecil dan parpol Islam mulai bangkit meraih dukungan...Keterkaitan sosiologi-ideologi
6	A18211010	pengusaan parpol nasionalis masih tinggi...di tingkat kota/kabupaten parpol kecil dan	Islam mulai bangkit meraih dukungan...Keterkaitan sosiologi-ideologi pada parpol yang
7	A18211010	dan parpol Islam mulai bangkit meraih dukungan...Keterkaitan sosiologi-ideologi pada	yang belum terlalu kuat sebagai tempat untuk partai untuk memperhebat pilihan politik
8	A18211010	PDK) dan Partai Persatuan Pembangunan (PPP)...Suaran pilihan parpol	lebih kritis di posisi kedua...yang biasanya didominasi PDI-P...maka
9	A18211010	ada 22 partai yang menjadi anggota parlemen provinsi...Karakter terbuka Bureaynya parpol	yang mendapatkan kursi di parlemen menjadi peran besarnya peluang partai menang di
10	A18211010	wilayah kapita barang ini yang cenderung terbuka dan tak eksklusif pada satu parpol	terutama meski partai ini dominan...Politikus Golkar...John Fathie
11	A12211093	parpol politik... "Terbuka dan profesional di kabinet seanya tidak ada kepentingan parpol politik... "	yang masuk... " oleh Franky...Kritik Unsur Asosiasi Peneliti dan Budhanta
12	A12212091	Kalimat yang rumpes mengaitkan bebas 285 kilogram...Paran Ktrial Parpol	Keberadaan par tai politik di wa sa ini nyaris seperti sebuah paradoks
13	A12212091	di mana 51,4 persen responden menyatakan tidak puas de agan kinerja parpol	Meski tetap terbuka peluang untuk tidak menggaribarkan si tua si seangapahnya
14	A12213091	hasil sur vai tersebut dan berbagai survei lainnya mengungkapkan makin mudahnya citra parpol	di mana politik... Situasi ini berbanding terbalik di awal reformasi
15	A12213091	sebuah media pengabdian sehingga nyaris tak terduga...Beberapa manfaat Perilaku peron parpol	yang dapat dirasakan langsung dengan konsep partai itu sendiri dalam mengedipkan deyat

Query "caper*" returned 650 matches in 96 different texts (in 29,987,513 words (17,277 texts); frequency: 21.676 instances per million words)

Showing frequency breakdown of both words and annotation in this query, at the query node; there are 19 different types and 650 tokens at this concordance position. (in 4.12 seconds)

No.	Query result	No. of occurrences	Percent
1	caper_NN	595	91.54%
2	caperwaj_NN	22	3.38%
3	caperis_NN	11	1.69%
4	caper-caper_NN	3	0.46%
5	caper-caperwaj_NN	2	0.30%
6	caper-caperis_NN	2	0.31%
7	caperis_NN	2	0.31%
8	Caperis_NN	1	0.15%
9	Caper-Mawa_NN	1	0.15%
10	Caper-Budi_NN	1	0.15%
11	Caper-Sia_NN	1	0.15%
12	Caper-Jaka_NN	1	0.15%
13	Caper-Mawa_NN	1	0.15%
14	Caper-Wicak_NN	1	0.15%
15	Caper-Hadi_NN	1	0.15%
16	Caper-Sarkas_NN	1	0.15%
17	caperis_NN	1	0.15%
18	caper-Mawadati_NN	1	0.15%
19	caper-Madi_NN	1	0.15%

Sementara itu, akronim dubes cukup banyak ditemukan dalam korpus sebagaimana yang ditunjukkan dalam table berikut ini.

No.	Query result	No. of occurrences	Percent
1	dubes_NNP	139	76.80%
2	Dubes_NN	41	22.65%
3	dubesnya_NN	1	0.55%

- b. KV + KV : bima dari Biru Malam ‘nama kereta api Jakarta-Surabaya’, dan pati dari perwira tinggi. Bentuk akronim tipe ini tidak banyak muncul lagi dalam korpus TBIK.
- c. KVK + KVK : parpol dari partai politik, cerpen dari cerita pendek, golput dari golongan putih ‘sebutan bagi orang-orang yang memutuskan tidak memilih ketika pemilu’, dan bandes untuk bantuan desa.

Your query "parpol*" returned 458 matches in 88 different texts (in 29,087,513 words [17,277 texts]; frequency: 15.273 instances per million words), reduced to results where query node matches word-tag combination: **parpol_NN** (440 hits) (0.047 seconds - retrieved from cache)

No	Text	Solution 1 to 50	Page 1 / 9
1	A1Z211002	politik yang kian sepiang elemen , terutama kalangan pemerintah , parpol dan lain sebagainya . Ka rnu ita , kehadiran KKR	
2	A1Z211002	Suhu (167) . Ia juga meminta partai politik (parpol) lain untuk benar membongkar jika memang memang diragukan . Dengan	
3	A1Z211076	dapat du duk di kari DPRD karena me rupakan 'hak parpol lain sebagai wakil idola nya parti beranggotakan dari berbagai ba tas	
4	A38211010	Perole Legislatif Papua Barat mulai berubah . Meski di tingkat provinsi penguasaan parpol nasionalis masih tinggi , di tingkat kota/kabupaten parpol kecil dan parpol Islam	
5	A38211010	di tingkat provinsi penguasaan parpol nasionalis masih tinggi , di tingkat kota/kabupaten parpol kecil dan parpol Islam mulai bangkit merah dukungan . Keterkaitan sosiologi-ideologi	
6	A38211010	penguasaan parpol nasionalis masih tinggi , di tingkat kota/kabupaten parpol kecil dan parpol Islam mulai bangkit merah dukungan . Keterkaitan sosiologi-ideologi pada parpol yang	
7	A38211010	dan parpol Islam mulai bangkit merah dukungan . Keterkaitan sosiologi-ideologi pada parpol yang belum terula kan adanya menjadi sebab relatif makin menyebarkan pilihan publik	
8	A38211010	PEK , dan Partai Perseutan Pembangunan (PPP) . Seluruh pilihan parpol lebih karena di posisi kedua , yang biasanya dikendalikan PDIP , mada	

Your query "parpol*" returned 458 matches in 88 different texts (in 29,087,513 words [17,277 texts]; frequency: 15.273 instances per million words), reduced to results where query node matches word-tag combination: **parpol_NN** (440 hits) (0.047 seconds - retrieved from cache)

No	Text	Solution 1 to 50	Page 1 / 9
1	A1Z211002	politik yang kian sepiang elemen , terutama kalangan pemerintah , parpol dan lain sebagainya . Ka rnu ita , kehadiran KKR	
2	A1Z211002	Suhu (167) . Ia juga meminta partai politik (parpol) lain untuk benar membongkar jika memang memang diragukan . Dengan	
3	A1Z211076	dapat du duk di kari DPRD karena me rupakan 'hak parpol lain sebagai wakil idola nya parti beranggotakan dari berbagai ba tas	
4	A38211010	Perole Legislatif Papua Barat mulai berubah . Meski di tingkat provinsi penguasaan parpol nasionalis masih tinggi , di tingkat kota/kabupaten parpol kecil dan parpol Islam	
5	A38211010	di tingkat provinsi penguasaan parpol nasionalis masih tinggi , di tingkat kota/kabupaten parpol kecil dan parpol Islam mulai bangkit merah dukungan . Keterkaitan sosiologi-ideologi	
6	A38211010	penguasaan parpol nasionalis masih tinggi , di tingkat kota/kabupaten parpol kecil dan parpol Islam mulai bangkit merah dukungan . Keterkaitan sosiologi-ideologi pada parpol yang	
7	A38211010	dan parpol Islam mulai bangkit merah dukungan . Keterkaitan sosiologi-ideologi pada parpol yang belum terula kan adanya menjadi sebab relatif makin menyebarkan pilihan publik	
8	A38211010	PEK , dan Partai Perseutan Pembangunan (PPP) . Seluruh pilihan parpol lebih karena di posisi kedua , yang biasanya dikendalikan PDIP , mada	
9	A38211010	ada 22 partai yang menjadi anggota parlemen provinsi . Karakter terbuka Banyuwangi parpol yang mendapatkan kursi di parlemen menjadi petir besarnya peluang partai menang di	
10	A38211010	wilayah kepala daerah ini yang cenderung terbuka dan hai dihalal pada satu teritori meski partai ini dominan . Pilkada Golkar , John Fahie ,	
11	A3C211002	parlai politik , "Tantangan kalangan profesional di kabinet upaya tidak ada kepentingan parpol yang masuk , "saja Franky . Ketua Union Asosiasi Pendidikan Indonesia	
12	A1Z212001	Kalvin) yang mampu mengangkat beban 285 kilogram . Pemas Kinal parpol Kedebaras per sui politik du wa su ini nyaris seperti sebuah paradoks	
13	A1Z212001	di mana 51,4 persen responden menyatakan tidak puas di ngin kinerja parpol Meski tidak terbagi peluang untuk tidak menggariskan si tas si sesungguhnya	
14	A1Z212001	hasil sui vai sebarbar dan berbagai variasi lainnya menghidkan makin meluasnya citra parpol di mata publik . Situasi ini berdampak terakut di awal reformasi	
15	A1Z212001	sebuah media penghiburan sehingga nyaris tak terdengar . Beberapa manfaat Peltusan peme parpol yang dapat dirasakan langsung dengan komposisi partai ini sendiri dalam mendukung dnyat	

Query "cerpen*" returned 650 matches in 94 different texts (in 29,987,513 words [17,277 texts]; frequency: 21.676 instances per million words)

Showing frequency breakdown of both words and annotation in this query, at the query node: there are 19 different types and 650 tokens at this concordance position. (0.513 seconds)

No.	Query result	No. of occurrences	Percent
1	cerpen_NN	595	91.54%
2	cerpenias_NN	22	3.38%
3	cerpenas_NN	11	1.69%
4	cerpen-cerpen_NN	3	0.46%
5	cerpen-cerpenias_NN	3	0.46%
6	cerpenkolong_NN	2	0.31%
7	cerpenku_NN	2	0.31%
8	Cerpenia_NN	1	0.15%
9	Cerpen Alimad_NN	1	0.15%
10	Cerpen Asat_NN	1	0.15%
11	Cerpen Gak_NN	1	0.15%
12	Cerpen Jaki_NN	1	0.15%
13	Cerpen Maksa_NN	1	0.15%
14	Cerpen Wulak_NN	1	0.15%
15	Cerpen Wulu_NN	1	0.15%
16	cerpen-cerpenias_NN	1	0.15%
17	cerpenias_NN	1	0.15%
18	cerpen-kolong_NN	1	0.15%
19	cerpen Maksa_NN	1	0.15%

- d. KVK + KV : serda dari sersan dua, deplu untuk departemen luar negeri, dan letda untuk letnan dua,

Query "deplu" returned 70 matches in 7 different texts (in 29,987,513 words [17,277 texts]; frequency: 2.334 instances per million words)				
Showing frequency breakdown of both words and annotation in this query, at the query node there is 1 different type and 70 tokens at this concordance position. (0.332 seconds)				
 < << >>		Breakdown position: node B	Show: Frequency breakdown of words and annotation B	Download frequency breakdown table (for both words and annotation)
No.	Query result	No. of occurrences		Percent
1	deplu, PNP	70		100.00%

Your query "duplu"" returned 70 matches in 2 different texts (in 29.987.513 words [17,277 texts]; frequency: 2.334 instances per million words), reduced to results where query node matches word-tag combination: Duplu_NNP (70 hits): (1,637 results – retrieved from cache)					
[< << >> >]		Show Page: 1	1/100 hits	Display results order	Choose action:
No	Text	Solution 1 to 50	Page 1 / 2		
1	ECU-1952	Departemen Dalam Negeri (Duplgi) s. : Departemen Luar Negeri (Dupla - / : Departemen Kehutanan, d. Departemen Keuangan - n. Departemen		
2	ECU-1952	ECU number, "Sajarah Diplomat RI dan Masa ke Masa	Dupla 2009/1 Yeliduk Klaten dan Masa ke masa merupakan hasil dari		
3	ECU-1952	Pada Peristiwa Sumpah Departemen Luar Negeri ... 1971. Jakarta	Dupla 1971, 484,259,120. Tanpa jangkai pada dan jangkai panjang sila		
4	ECU-1952	ECU Number, "Sajarah Diplomat RI dan Masa ke Masa	Dupla 2009/1 dan 259,120,120. Tanpa jangkai pada dan jangkai panjang sila		
5	ECU-1952	Go idn. Justru kita yang sebenarnya punya banyak peluang	Dupla 2009/1 dan 259,120,120. Tanpa jangkai pada dan jangkai panjang sila		
6	ECU-1952	Indonesia memiliki pendidikan bahwa penerapan Indonesia harus berlaku pada hasil	Dupla 1997,89,100. Buku yang dikeluarkan oleh Departemen Luar		
7	ECU-1952	kebutuhan itu; tugas atau depresi atau dari para petugas untuk	Dupla 1997,18,130.0. Seperti yang akan dilakukan di atas.		
8	ECU-1952	Masing-masing pemerintah beritikad untuk mengupayakan penghapusan perang dan kemiskinan pemerintah	Dupla 2004,982.1. Departemen perajajaran beritahu penghapusan perijaraan Roon-Roon. Pada		
9	ECU-1952	3. Konferensi Meja Bundar disusupi diadukan di Den Haag	Dupla 1,2004,982.1. Perjanjian antara Me. Roon dengan Dr.		
10	ECU-1952	peristiwa tersebut bereslah setelah diungkap dan dibawa ke Inggris oleh Belanda	Dupla 2004,982.1. Kemungkinan pemerintah Indonesia ke Yogyakarta merupakan pengaduan		
11	ECU-1952	ini hanya diadukan oleh wali-wali Republik Indonesia dan Perjanjian Mayawahan Federal	Dupla 12004,921.1. Pada pertemuan dan hasil Konferensi Antar-Indonesia akan menyojiki pemerintahan		
12	ECU-1952	dan memana pemerintah-Indonesia-departemen dan pemerintah Indonesia (beutan Belanda)	Dupla 2004,982.1. 3. Konferensi Meja Bundar dan Belanda Republik		
13	ECU-1952	Meja Bundar di Den Haag pada 22 Januari 1949. Antara lain	Dupla 2004,927,920.1.1. Tajam Konferensi. Seperti tersebut dalam		
14	ECU-1952	Dipenuhi. Menurut dan Negeri, 1960/172. Dengan pokok-pokok	Dupla 12004,952,950.1. A. Negeri. Negeri-Negeri B. Pagan. Pagan Republik C. Status		
15	ECU-1952	Sekitar dan Belanda beresnya yang penting Belanda beres dan Amerika	Dupla 12004,952.1. Upacara diadukan Belanda oleh Belanda dan Menteri Belanda W.		
16	ECU-1952	ditutup dengan mengundikan bahwa Indonesia Ray dan juga keahungan Belanda Wilhelm	Dupla 1,2004,950.1. Sedangkan di Jakarta dengan waktu yang bersamaan		
17	ECU-1952	Negara sebagai belian. Republik Indonesia diadukan oleh Negeri Pagan Belanda	Dupla 19,997,48.1. Dengan penandatanganan sebagai Pagan-Pagan Belanda oleh Belanda kepada		
18	ECU-1952	menanggung beban pembayaran atau sebagai nilai Belanda yang akan dan tidak	Dupla 19,997,47,48.1. Menurut Hatta yang akan diadukan menjadi Wakil Presiden Republik		
19	ECU-1952	tapi juga akan menanggung sebagai Yogyakarta di Jakarta.	Dupla 19,997,49.1. Selain Hatta, Program Kerja Kabinet Republik Indonesia seperti		

- e. VK + KVK : ormas dari organisasi massa, akpol dari akademi polisi, dan aspal dari asli tetapi palsu.

Your query "aksep*" returned 9 matches in 7 different texts (in 20,987,513 words [17,277 texts]; frequency: 0.300 instances per million words), reduced to results where query node matches word-tag combination: Aksep_NN (9 hits) (0.027 seconds)			
 < << >> >		Word-tag: 1 2	Use view Show 8 synchro nodes Choose filters
No	Text	Solution 1 to 9	Page 1 / 1
1	ASB1094	Petri dan menyandang pangkat jenderal bintang tiga. Budi Wacno menjabat sebagai	Aksep
2	ASD1014	yang juga waku kivil apa sebagai, serta sudah sudah jadi	Aksep
3	BC15012	(802), Kepala Badan Narkotika Nasional Komandan Jenderal Adnan	Aksep
4	BC15012	serta Kepala Badan Penanggulangan dan Keamanan Komandan Jenderal Pital	Aksep
5	BC15012	Kernis, sudah lahir ini, dioret dari pencemaran. Alasannya, adalah	Aksep
6	HJA1002	yang sebelumnya memang sudah Terima AAN, AAN, AAN, dan Terima	Aksep
7	K1A115025	Terasa AAN 90 yang, Terasa AAN 88 yang, dan Terasa	Aksep
8	K1A10017	Muslami akan akan program tinggi operasi Akademi Militer 1 juta	Aksep
9	L1C1H011	Tinggi Militer Kopolitase - 79 926 Kopolitase 64001 Akademi Militer 1 juta	Aksep

Your query "ormas*" returned 329 matches in 128 different texts (in 29,987,513 words [17,777 texts]; frequency: 10,971 instances per million words), reduced to results where query mode matches word-tag combination: ormas_nns_330 (363 hits) [0.48 seconds]					
K	<<	>>	P	Show page	Like text
				Show in statistics area	Choose action...
No	Text	Solution 1 to 50	Page 1 / 7		
1	A12201017	(dialog dengan pola yang paling baik) . Kita berharap	Idara di Indonesia setidaknya satu- staff dalam menyepaki beragam persoalan yang berpeten-		
2	A12201028	publik bahwa beberapa tanggapan berbagai ka hingga tentamen para siswa akan	Idara terkait Dunia Muka-muk Komitin (MK) . Maple yang		
3	A12201048	seperti Kay Per y diperbolehkan i juga mungkin bahwa pada suatu kejadian	seperti dari Pemberi lisan (PL) dan beberapa kali		
4	A12201050	atau malah akan. Dalam situasi yang berubah pelai , masyarakat juga	o-pu ber , semakin meningkat dengan peningkatan yang		
5	A12201055	aka Ma homanly dalam hal ini dan kadarnya sudah terbeli	in baru- besar akan diperlihatkan ma bu kat baru tentang situasi . Selain		
6	A12201087	jumlah mereka memang tidak yang telah banyak , juga ma	Idara . Khususnya Mahasiswa (Mahasiswa) , memonition diri sebagai ka katan sari		
7	A12201087	kebermatan teknologi terakibat di sini to bu ran mungkin bagian dari itu	Idara agar kemasangan dan eksn di tingkat masyarakat . Dalam konteks		
8	A12201097	yang datang . Media sosial kini menjadi sarana alternatif di tengah adanya	terhadap isi itu kemasangan yang terakibat di tingkat . Ormas Idara		
9	A12201097	tergah adanya ormas terdapat iku in kemasangan yang terakibat di tingkat . Ormas	Idara (Mahasiswa) lebih subyektif persoalan pelai . administrasi organisasi		
10	A12201096	yang ada . Khusus bagi masyarakat Idara yang menjadi bagian maha di sini	teratur te , biasanya mereka akan lebih megak liki pendapat maha di sini		
11	A12201086	Mengingat kini pemerintah lebih serta merta mengontrolisasi atau dengan metode mading mading	tidak memonition atau memonition kemasn . Yang juga , sama proses		
12	A12201096	terjadi dari ini . Dalam 10 tahun terakhir ini sukukuk akan mading mading	in lebih yang mengontrol kemasn ataupun kemasn ataupun kemasn yang dengan jemaah		
13	A12201097	kini sukukuk akan sukukuk ormas , baik yang mengontrol kemasn ataupun kemasn	kebahasaan yang dengan jemaah madinglik pelai besar bu ter ataupun di bu		
14	A12201097	yang telah terse- main dalam payung baw Pancasila dapat kemasn sukukuk sukukuk	yang kio- ng pahan terapan fahala kemasn dan mading kemasnlik Ahfahd		
15	A12201097	tes- ini to menjadi agenda kita . Bahkan ka . (5	di PNSU) untuk melakukan ini ke Merko Pribadi . Ini ka- du		
16	A12201097	Pemerintah dan atau , kudu to kudu kemasnlik atau kemasnlik mading mading	tersebut . "Ma mering , diukapan pemerintah dan pemerintah juga ormas yang		
17	A12201097	ma ormas terakibat . "Ma mering , diukapan pemerintah dan pemerintah juga ormas	yang terapan terakibat kemasn . Diukapan terakibat terapan terapan terapan		
18	A12201014	Kerlap , representasi kepemimpinan . Praktek , anti bu pital Idara mading	Idara akan panya dengan kemasnlik yang dapat mengontrolisasi atau Idara		
19	A12201005	ormas juga kemasnlik . atau Praktek 2014 . atau Praktek 2014 . atau Praktek 2014	dan juga sistem di kemasnlik . atau Praktek 2014 . atau Praktek 2014 . atau Praktek 2014		

f. KVK + VK : balar dari balai arkeologi, dan dirut dari direktur utama,

Your query "dirut" returned 145 matches in 79 different texts (in 20,987,513 words [17,277 texts]; frequency: 4.835 instances per million words), reduced to results where query node matches word-tag combinations: Dirut_RNP (111 hits) : 112 percent

No.	Text	Solution 1 to 50	Page 1 / 3
1	AJ81412 pegawai Kementerian, "Pecutan" empat kanvas berisi uang itu berasal dari	Dirut PT Asa Neta Indonesia, Said Pasha David Niswan atau Paul Niswan	
2	A2213308 angkutan Lebaran tahun ini, sementara bagi kendaraan buspel kuning. Eks	Dirut PT Kenta Api Indonesia ini mengaku sempat gresir dengan apa yang dimilikinya	
3	A2213302 Dengan intensi, jika, dan penekanan yang ditekankan, "Wakil	Dirut Menghadir dan Perwakilan JAKARTA, Wakil Direktur Utama PT (Persero)	
4	A3221904 untuk memperoleh cukup baik walik yang dapat digunakan sebagai bar. Momen	Dirut Pirel Indi Terungkap JAKARTA, Komisi Pembeantasan Korupsi (PKP)	
5	A3819010 itu. Adapun dalam kasus ini, Kenyabang, yang merupakan mantan	Dirut PT Garuda Indonesia, ditugaskan KPK sebagai temannya karena diduga menerima suap	
6	A3820002 Rahu (1/1/2002). Salah satu saksi yang diperiksa adalah	Dirut PT Transcar Fardj Istomana Dewanto Achom, Dhirubhai, PT TPI termasuk	
7	B1612000 Caterpillar. Beberapa CEO dari perusahaan-perusahaan Indonesia terlihat hadir, antara lain	Dirut BNI Gusto Mediatore, Dirut Bank Mandiri Zulkifli Zaini, dan Dirut	
8	B1612000 perusahaan-perusahaan Indonesia terlihat hadir, antara lain Dirut BNI Gusto Mediatore,	Dirut Bank Mandiri Zulkifli Zaini, dan Dirut BRI Selyan Basir, Acara	
9	B1612000 Dirut BNI Gusto Mediatore, Dirut Bank Mandiri Zulkifli Zaini, dan	Dirut BRI Selyan Basir, Acara Indonesia Investment Day ini dikecandikan oleh Bank	
10	B2110020 nilai tagihan bukan berarti piutang perusahaan dititikan menumpuk, "Arisian Pak	Dirut (Dwi Setijono) jalan, semua piutang harus baik. Skarang	
11	B220011 ketubuhan informasi tersebut. Solokhaya, kata dia, "Ako caru hantu	Dirut "Vice President of Corporate Communication PT Pertamina (Persero)	
12	B1113002 Jembatan Merak Plaza (JMP) Kota Surabaya mulai 13-30 Oktober 2012.	Dirut PT Lami Citra Tbk selaku pengelola JNP, Priyo Setio Budi,	
13	B1113002 untuk itu, saya meminta Pak Syafi dan Pak Benay Wicaksono	Dirut Bank Mega Syariah untuk melakukan pendekatan kepada Menteri Agama agar Bank	
14	B1113002 dan menobatkan tembak tebel bingkisan yang menghiasi langkahnya. Ketika dipercaya menjadi	Dirut PLN dan kemudian dipercaya Presiden SBY menjadi menteri negara BUMN, nama	
15	B1113002 dan jenderal. Ketika diangkat Presiden SBY ke istana dan diwarai jabatan	Dirut PLN, sebenarnya Dahlan sudah menah. Berkali-kali dia katakan kepada presiden	
16	B1113002 "saya Dahlan... kemudian penjelasan presiden. Dahlan menerima amanat sebagai	Dirut PLN (di besar, Rp50 juta, "saya Dahlan, dan	
17	B1113002 apa pun dari perusahaan negara tersebut. "Palatol, gaji sebagai	Dirut PLN " "Tidak... 5 saya memang sudah (benar)	
18	B1113002 Jali Pak Dahlan tidak mengucil perubahan pendapatan dengan bekerja sama sebagai	Dirut PLN dan tak sampai dua tahun diwarai jabatan lebih tinggi sebagai	
19	B1113002 Tahun dan Mulu Rijkman: Rencan saja, begini dan dilantik sebagai		

Pola ini menjadi semakin kompleks jika kita mengambil akronim yang berasal dari kependekan gabungan kata lebih dari dua kata. Ada yang secara konsisten hanya mengambil huruf awal dari gabungan kata pada bentuk kepanjangannya, seperti TNI dari Tentara Nasional Indonesia dan BIN dari Badan Intelijen Negara. Contoh lainnya juga dapat dilihat dari penamaan fakultas di universitas di Indonesia, seperti FMIPA dari Fakultas Matematika dan Ilmu Pengetahuan Alam dan FKIP dari Fakultas Keguruan dan Ilmu Pendidikan.

Jika kita melihat dalam korpus TBIK, misalnya untuk mencari frekuensi kemunculan kata TNI, kita akan dapat melihat hasil yang cukup menarik sebagai berikut.

Query "TNI*" returned 4,810 matches in 755 different texts (in 29,987,513 words [17,277 texts]; frequency: 160.460 instances per million words)

Showing frequency breakdown of both words and annotation in this query, at the query node; there are 45 different types and 4,810 tokens at this concordance position. (1,645 separate)

No.	Query result	No. of occurrences	Percent
1	TNI_RNP	4034	96.34%
2	TNI-RNP_RNP	37	1.34%
3	TNI-RNP_RNP	28	0.58%
4	TNI-AG_RNP	12	0.25%
5	TNI-AL_RNP	9	0.19%
6	TNI-Angkatan_RNP	9	0.19%
7	TNI-AG_RNP	8	0.17%
8	TNI-AG_RNP	6	0.12%
9	TNI_RNP	4	0.08%
10	TNI-POLITIKAS_RNP	3	0.06%
11	TNI_RNP	3	0.06%
12	TNI-Prinsip_RNP	2	0.04%
13	TNI-Pemimpin_RNP	2	0.04%
14	TNI-memilih-dalam-pengambilan-keputusan-lengkap-lengkap_RNP	2	0.04%
15	TNI-Pemimpin-pemimpin_RNP	1	0.02%
16	TNI-AG_RNP	1	0.02%
17	TNI-AG_RNP	1	0.02%
18	TNI-AG_RNP	1	0.02%
19	TNI-AG_RNP	1	0.02%
20	TNI-AG_RNP	1	0.02%
21	TNI-AG_RNP	1	0.02%
22	TNI-Angkatan-Pemimpin_RNP	1	0.02%
23	TNI-Pemimpin_RNP	1	0.02%
24	TNI-Pemimpin_RNP	1	0.02%
25	TNI-Pemimpin-Pemimpin_RNP	1	0.02%
26	TNI-Pemimpin_RNP	1	0.02%
27	TNI-Pemimpin_RNP	1	0.02%
28	TNI-Pemimpin_RNP	1	0.02%
29	TNI_RNP	1	0.02%
30	TNI-Angkatan_RNP	1	0.02%
31	TNI-Pemimpin_RNP	1	0.02%
32	TNI_RNP	1	0.02%

Berdasarkan temuan di atas, kita dapat melihat bahwa akronim TNI dalam korpus ini sangat banyak jumlahnya, yaitu lebih dari 4600 kali kemunculan. Hal lain yang dapat kita jabarkan dari table di atas adalah bahwa kata TNI selalu berkolokasi dengan akronim lain yang seranah. Dari akronim-akronim tersebut, kata TNI lebih sering muncul dan berkolokasi dengan akronim Polri (Polisi Republik Indonesia). Kemudian, disusul dengan AU (Angkatan Udara), AL (Angkatan Laut), dan AD (Angkatan Darat). Hal ini berbeda dengan kata ABRI. Kata ABRI memiliki kemunculan cukup tinggi, sekitar 537 kali kemunculan dalam korpus. Namun, kata ini tidak ditemukan berkolokasi dengan kosakata yang tadi disebutkan berkolokasi dengan TNI.

Beberapa singkatan begitu sering digunakan, banyak orang tidak menyadari sifatnya sebagai singkatan, seperti radar (radio detection and ranging), scuba (self-contained underwater breathing apparatus), laser (light amplification by stimulated emission of radiation), dan juga rudal (peluru kendali). Ini menunjukkan bahwa karena frekuensinya yang tinggi, singkatan-singkatan ini telah mengaburkan identitasnya sebagai singkatan, khususnya akronim, dan sekarang diperlakukan sebagai kata-kata yang sebenarnya. Frekuensi dan kemunculannya di dalam konteks pada korpus TBIK dapat dilihat sebagai berikut.

No.	Word	Frequency
1	radar	255
2	radarnya	14
3	radar-radar	3
4	radar"-nya	1
5	radar bogor	1
6	radar grammetry	1
7	radar ku	1

No.	Query result	No. of occurrences	Percent
1	rudal_NN	223	94.49%
2	rudal-rudal_NN	3	1.27%
3	rudalnya_JJ	3	1.27%
4	Rudal/KCR_NN	2	0.85%
5	rudalnya_NN	2	0.85%
6	rudal-..._NN	1	0.42%
7	rudal-..._NN	1	0.42%
8	rudal-..._NN	1	0.42%

Berdasarkan frekuensi yang muncul dalam korpus tersebut, dapat dilihat bahwa akronim yang sudah lazim dan diperlakukan seperti kata ini memiliki kemunculan yang cukup tinggi. Hanya kata *scuba* yang kemunculannya dapat dianggap sedikit, namun ini tidak menjadi masalah mengingat bahwa korpus ini masih terlalu umum dan tidak spesifik untuk bidang olahraga saja.

4.3 Penggunaan Akronim Berdasarkan Ranah Penggunaan Bahasa

Berdasarkan ranah penggunaan bahasa, Anam, Ahmad Khoiril, dkk. (2021) pernah melakukan riset yang menunjukkan bahwa pada data jurnalistik (mereka menggunakan koran *Pos Kota* sebagai sumber datanya) dapat dilihat frekuensi penggunaan akronim sebagai berikut.

Kategori Pembentukan Akronim Sesuai Bidanganya	Jumlah Temuan Akronim dalam Angka	Jumlah Temuan Akronim dalam Persen
Bidang Forensik	74	33,48 %
Bidang Politik	52	23,53 %
Bidang Pemerintahan dan Ketatanegaraan	25	11,31 %
Bidang Khas Geografis	20	9,05 %
Bidang Olahraga	16	7,24 %
Bidang Organisasi	12	5,43 %
Bidang Hiburan	7	3,17 %
Bidang Nama Diri	6	2,72 %
Bidang Teknologi	5	2,26 %
Bidang Hak Cipta dan Badan Hukum	3	1,36 %
Bidang Pendidikan	1	0,5 %
Jumlah	221 kata	100 %

Sementara itu, pada data yang lebih beragam variasi penggunaan bahasanya seperti korpus TBIK di CQPWeb, hasilnya cukup berbeda. Hal ini dipengaruhi oleh akronim itu sendiri, misalnya akronim *baper* hany muncul dalam surat kabar dan tulisan populer saja, misalnya majalah, novel, dan tulisan populer, seperti yang ditemukan pada gambar berikut ini.

Distribution of hits for query "baper", restricted to texts meeting criteria "sumber: A_koran or B_majalah or C_cerpen or D_novel or E_buku_teks or F_jurnal or G_disertasi_tesis_skripsi or H_biografi or I_populer or J_perundangan or K_laman_resmi or L_surat_resmi", returned 5 matches in 4 different texts
Currently displaying distribution across text classifications.

Display distribution of: Show as:

Cross-tabulating against: Other actions:

Based on classification: sumber				
Category <input type="text" value="1"/>	Words queried within category	Hits in category	Dispersion (no. texts with 1+ hits)	Frequency <input type="text" value="1"/> per million words queried within category
B_majalah	2,439,621	1	1 out of 2,835	0.41
D_novel	2,622,758	3	2 out of 38	1.14
I_populer	2,290,231	1	1 out of 66	0.44
Total:	7,352,610	5	4 out of 2,939	0.68

Based on classification: tahun				
Category <input type="text" value="1"/>	Words queried within category	Hits in category	Dispersion (no. texts with 1+ hits)	Frequency <input type="text" value="1"/> per million words queried within category
2016	2,950,811	2	2 out of 1,465	0.68
2018	3,205,346	2	1 out of 1,835	0.62
2019	3,070,041	1	1 out of 1,825	0.33
Total:	9,226,198	5	4 out of 5,125	0.54

Sementara itu, di sisi lain, kosakata akronim yang identik dengan ranah politik, seperti *caleg*, memiliki kecenderungan penggunaan yang berbeda. Misalnya, pesebaran penggunaan kata *caleg* dan *parpol* pada data korpus

TBIK di CQPWeb menunjukkan bahwa kedua akronim tersebut lebih sering digunakan di surat kabar, tulisan biografi, dan tulisan ilmiah. Hal ini memperlihatkan bahwa kedua kata ini memang ranah penggunaannya tidak terlalu luas. Bahkan, jika melihat data berdasarkan persebaran penggunaan per tahun, penggunaan kedua kata ini pada tahun 2014 menunjukkan angka yang tertinggi. Hal ini tentu juga berhubungan dengan pemilu yang diadakan pada tahun 2014 tersebut. Penggunaan kata *caleg* pada data dapat dilihat melalui gambar berikut ini.

Distribution of hits for query "caleg", restricted to texts meeting criteria "sumber: A_koran or B_majalah or G_disertas_tesis_skripsi or H_biografi or I_populer or J_perundangan or K_laman_resmi or L_surat_resmi", returned 137 matches in 37 different texts, sorted on position +1 (137 hits)
Currently displaying distribution across text classifications.

Display distribution of: Test categories: Show as: Distribution table:
Cross-tabulating against: No crossstats available: Other actions: Choose action:

Based on classification: sumber

Category <input type="text"/>	Words queried within category	Hits in category	Dispersion (no. texts with 1+ hits)	Frequency <input type="text"/> per million words queried within category
A_koran	2,701,024	100	27 out of 6,494	39.24
B_majalah	2,439,621	11	2 out of 2,835	4.51
G_disertas_tesis_skripsi	2,703,612	2	2 out of 85	0.74
H_biografi	2,450,463	15	3 out of 53	6.12
I_populer	2,290,231	1	1 out of 66	0.44
K_laman_resmi	2,440,866	2	2 out of 4,698	0.82
Total:	15,025,817	137	37 out of 14,231	9.12

Based on classification: tahun

Category <input type="text"/>	Words queried within category	Hits in category	Dispersion (no. texts with 1+ hits)	Frequency <input type="text"/> per million words queried within category
2011	1,754,158	6	2 out of 1,390	3.42
2012	1,897,085	6	2 out of 1,530	3.16
2013	2,003,912	10	5 out of 1,552	4.99
2014	1,886,730	20	18 out of 1,340	47.70
2017	1,956,962	1	1 out of 1,332	0.51
2018	2,221,261	13	3 out of 1,619	5.85
2019	2,079,440	7	3 out of 1,609	3.37
2020	2,286,684	4	3 out of 1,874	1.75
Total:	16,066,232	137	37 out of 12,436	8.52

5. KESIMPULAN

Bentuk akronim mungkin telah ditentukan tidak hanya oleh keberadaan kata-kata asli Indonesia tetapi juga oleh nilai-nilai budaya yang berlaku di masyarakat. Kereta mewah berwarna biru yang beroperasi pada malam hari dari Jakarta ke Surabaya, Biru Malam, bisa disebut *Bilam, *Rulam, atau *Ruma, yang semuanya mengikuti aturan fonotaktik bahasa. Namun nama resminya adalah Bima '(harfiah) Biru di Malam Hari', karena, saya kira, kata ini kebetulan adalah nama pahlawan yang paling kuat secara fisik dalam Mahabharata versi Jawa.

Dapat kita anggap bahwa apa yang oleh penutur asli non-linguis disebut "enak di telinga" sebenarnya adalah prinsip linguistik yang sangat mendasar yang juga disebut sebagai kompetensi (*competence*). Kompetensi inilah yang memungkinkan orang Indonesia menghasilkan akronim yang terdengar/

terlihat bagus. Ketika “enak didengar” dianggap sebagai dasar yang cukup untuk akronimisasi, tentu saja konsekuensinya ada banyak inkonsistensi di mana bentuk lengkapnya disingkat menjadi beberapa akronim yang berbeda. Misalnya saja, kata *militer* menjadi *mil*, *m*, dan *mi* masing-masing seperti di *koramil*, *kodam*, dan *mahmilub*. Sementara pemilihan *mi* alih-alih *mil* dalam *mahmilub* mungkin dipengaruhi oleh kehadiran ganda /l/ yang tidak sepenuhnya asing tetapi sangat jarang. Dengan demikian, tidak ada alasan mengapa militer di *koramil* dan *kodam* harus mengambil dua bentuk yang berbeda, terutama ketika istilah-istilah ini ditemukan oleh sumber yang sama--kantor militer di Jakarta. Akronimnya bisa saja menjadi *koram* (meski ada juga akronim *korem*!) dan *kodam*, atau *koramil* dan *kodamil* –yang kesemuanya mengikuti aturan fonotaktik bahasa, dan sama enaknanya dengan akronim yang ada. Namun, pada kenyataannya tidak seperti itu.

Namun, ada sejumlah besar akronim, yang saya berani katakan “berada di luar jaringan”. Kasus-kasus yang telah disebutkan sebelumnya seperti pangkopkamtib, konjenbant, dan Ifalpolekrochsosbud dan bentuk-bentuk lain seperti ditaj (Direktorat Ajudan Jendral) ‘direktorat Ajudan Jenderal’, urhibjah (urusan hiburan dan kesejahteraan) ‘Bagian Hiburan dan Kesejahteraan’, depdag (departemen perdagangan).) jelas tidak enak di telinga orang Indonesia. Karena ini masalahnya, mungkinkah penutur asli dalam contoh khusus ini adalah apa yang disebut Chomsky “tidak menyadari tata bahasa mereka yang terinternalisasi” (Chomsky, 1970, p.194) , atau, sebenarnya mereka mengikuti apa yang diungkapkan Humboldt secara akurat 138 tahun yang lalu, yaitu, “tidak peduli bagaimana bahasa bawaan secara keseluruhan, pada saat yang sama masih memiliki keberadaan eksternal yang tidak bergantung, mengarahkan kekuatan terhadap manusia itu sendiri”.

Dengan kata lain, Penciptaan akronim tampaknya hampir secara eksklusif didasarkan pada norma-norma yang secara inheren ada dalam bahasa dan, oleh karena itu, dimiliki oleh anggota komunitas tutur. Kesan “kedengaran enak di telinga” menunjukkan bahwa pembentukan akronim melibatkan kompetensi dan rasa Bahasa. Hal inilah yang terkadang memunculkan inkonsistensi (bentuk), seperti “koramil” – ‘mil’, “kodam”- ‘m’, dan “mahmilub” – ‘mi’, serta contoh lain, “sprint” (surat perintah). Bentuk akronim adakalanya telah ditentukan tidak hanya oleh keberadaan kata-kata asli Indonesia atau daerah atau asing, tetapi juga oleh nilai-nilai budaya yang berlaku di masyarakat. Kreativitas ini juga membuat penutur seakan menabrak aturan fonotaktik Bahasa Indonesia yang ada sehingga sering kali muncul bentuk akronim dengan

fonotaktik yang tidak lazim, di antaranya “calhaj”, “sprindik”, “baznas”, dll. Di samping itu, meluasnya penggunaan Bahasa di media sosial juga memberikan kesempatan yang lebih luas lagi sehingga muncul kreativitas-kreativitas yang bersifat kekinian identik dengan pengguna Bahasa yang masih muda di media sosial, contohnya “bucin”, “mager”, “cmn” (cemana – ‘macam mana’), “gpp” (gak pa pa – ‘nggak apa-apa’).

Karena penutur asli menurut definisi memiliki kompetensi linguistik, namun dalam kasus akronim Bahasa Indonesia sekarang mereka menciptakan bentuk permukaan yang tidak dapat dilacak ke tata bahasa yang mereka internalisasikan, jelas bahwa “eksistensi eksternal independen” Humboldt harus menjadi faktor, jika bukan faktornya, yang dapat menjelaskan mengapa orang menghasilkan akronim yang menyimpang. Sebagai linguist, kita harus prihatin dengan fenomena ini, karena jika ini terus berlanjut—seperti yang terjadi di Indonesia—sebenarnya kita sedang menyaksikan perkembangan bahasa dari dua polarisasi yang berlawanan. Saya tidak mengatakan bahwa ini disayangkan, tetapi saya mengatakan bahwa ini sangat unik, untuk sedikitnya, dan bahwa percabangan jangka panjang, terutama dalam struktur fonologis, harus diperhatikan dengan sangat cermat.

Referensi

- [1] Abidin, T. F. (2012). Penentuan Secara Otomatis Akronim dan Ekspansinya dari Data Teks Berbahasa Indonesia. *Prosiding SNATIKA Vol 01* (2011), 1(01).
- [2] Anam, Ahmad Khoiril, dkk. (2021). “Pembentukan dan Pembidangan Akronim pada Koran *Pos Kota*.” *Jurnal Deiksis*, Vol. 13 No. 1, Januari—April. Hlm. 12—20.
- [3] Anderson, B. (1966). The languages of Indonesian politics. Indonesia, (1), 89-116.
- [4] Aronoff, M., & Anshen, F. (2017). Morphology and the lexicon: Lexicalization and productivity. *The handbook of morphology*, 237-247.
- [5] Bauer, L. (1983). *English word-formation*. Cambridge: Cambridge University Press.
- [6] Bauer, L. (1988). *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.
- [7] Dardjowidjojo, Soenjono. (1979). *Acronymic Pattern in Indonesian*. In

- Nguyễn Đ.L. editor, *Southeast Asian linguistic studies*, Vol. 3. Pacific Linguistics, The Australian National University.
- [8] Derin, Tatum & Deliani, Susy & Fauziah, Nurul & Afifah, Nur & Hamuddin, Budianto. (2019). INDONESIA'S TENDENCY TO REFER ABBREVIATION AS ACRONYM: TYPES OF ABBREVIATION AS WORD FORMATION PROCESS. *Globish: An English-Indonesian Journal for English, Education, and Culture*. 8. 10.31000/globish.v8i2.1654.
 - [9] Isa, A. A. (2006). Abreviasi dalam Bahasa Inggris. *Wacana*, 8(1), 113-124.
 - [10] Kridalaksana, Harimurti. (2010). Pembentukan kata dalam bahasa Indonesia. Gramedia.
 - [11] Noviatry & Reniwati (2015). Singkatan dan Akronim Dalam Surat Kabar: Kajian Bentuk dan Proses. *Jurnal Arbitret*, 2(1), 28-43.
 - [12] Irawati. 2007. "Singkatan dan Akronim dalam Media *Chatting dan SMS* (*Analisis* Komunikasi Teks dalam Internet dan Telpn Seluler).
 - [13] Permatasari, N. P. (2013). Abreviasi, Afiksasi, dan Reduplikasi Ragam Bahasa Remaja Dalam Media Sosial Facebook. *Suluk Indo*, 2(3), 230-242.
 - [14] Rijal, S. (2015). Hubungan Makna Akronim dan Kata Pembentuknya pada Acara Indonesia Lawak Klub (ILK) di Trans 7. *Aksara*, 27(1), 73-82.
 - [15] Smith & Hefner, N. J. (2007). Youth language, gaul sociability, and the new Indonesian middle class. *Journal of Linguistic Anthropology*, 17(2), 184-203.



Tata Bahasa Indonesia Kontemporer (TBIK): *Morfologi* ini disusun dengan berbasiskan pada korpus yang memanfaatkan data yang diambil dari data bahasa autentik, disusun secara sistematis, dan tersimpan secara digital sehingga dapat dipastikan bahwa data yang digunakan dalam TBIK tersebut memang benar digunakan oleh penutur BI, bukan data rekaan dari penulis.

Secara umum, verba berawalan *meng-* merupakan verba paling produktif, kemudian diikuti oleh verba berawalan *di-*, *ber-*, dan *ter-*. Subtipe yang dominan dan produktif dari keempat awalan verba tersebut adalah *meng-*, *meng-/kan*, *meng-/i*; *di-*, *di-/kan*, dan *di-/i*; *ber-* dan *ber-/an*; serta *ter-*. Dalam ragam bahasa perundang-undangan dan ragam persuratan (surat resmi kedinasan) afiksasi verba ini memiliki frekuensi yang rendah. Hal itu mengindikasikan rendahnya variasi leksikal dalam kedua ragam tersebut. Sementara itu, dalam cerita naratif seperti cerpen dan novel serta koran, awalan verba ini memiliki frekuensi yang paling dominan. Hal itu mengindikasikan tingginya keragaman leksikal dan peran ragam tersebut sebagai wadah bentukan kata baru. Sementara itu, afiks pembentuk nomina yang paling produktif adalah *-an*, *ke-/an*, *peN-*, *peN-/an*, dan *per-/an*. Dua imbuhan pembentuk kata benda yang mirip secara bentuk, tetapi berbeda secara kuantitatif, yaitu imbuhan *pe-* dan *peN-*.

Kata majemuk dibedakan menjadi majemuk kata dan majemuk frasa. Majemuk kata merupakan kata majemuk yang terdiri atas dua unsur yang ditulis serangkai, sedangkan majemuk frasa merupakan kata majemuk yang terdiri atas dua unsur atau lebih yang ditulis terpisah. Di sisi lain, kata majemuk didefinisikan sebagai gabungan dua unsur atau lebih sehingga *di luar*, *ke kanan*, dan *dari samping* dianggap sebagai kata majemuk yang berupa majemuk frasa. Makna kata majemuk tidak selalu keluar dari makna komponen pembentuknya, bisa saja makna kata majemuk itu masih dapat ditelusuri dari makna kata pembentuknya.

Kreasi kata juga dibicarakan dalam buku ini. Kreasi kata mencakup singkatan dan akronim. Singkatan merupakan pemendekan beberapa kata yang lazimnya diambil huruf awal dan dieja huruf per huruf, sedangkan akronim merupakan singkatan yang berupa gabungan pemendekan beberapa kata yang diperlakukan sebagai kata dan dibaca sebagai kata. Frasa Uji Kemahiran Bahasa Indonesia dipendekkan menjadi UKBI. Jika UKBI dibaca *u-ka-be-i*, itu berarti singkatan, tetapi jika dibaca *uk-bi*, itu berarti akronim.

Wisnu Sasangka, Tim Penulis

