

## *Afiksasi Nomina dalam Bahasa Indonesia*

Karlina Denistia<sup>1</sup> dan Gede Primahadi Wijaya Rajeg<sup>2</sup>

<sup>1</sup> Vocational School, Universitas Sebelas Maret

<sup>2</sup> University of Oxford/Universitas Udayana

<sup>1</sup> [karlinadenistia@staff.uns.ac.id](mailto:karlinadenistia@staff.uns.ac.id), <sup>2</sup> [gede.rajeg@ling-phil.ox.ac.uk](mailto:gede.rajeg@ling-phil.ox.ac.uk)

### **English Abstract**

This paper describes noun formation in Indonesian through the affixation process. We collected the data from the corpus of Contemporary Indonesian Grammar, which consists of around 18 million words. This paper employs both qualitative and quantitative methodologies. We describe noun formation by examining the underlying forms' form, meaning, and word class. Furthermore, we analyze the productivity of each noun-forming affix, examine the development of affix productivity per year, and the distribution of word class of the base word for the five most productive noun-forming affixes: *-an*, *ke-/-an*, *peN-*, *peN-/-an*, and *per-/-an*. Using the corpus approach, we found new meanings created from the formation of nouns through the affixation process in Indonesian. The data shows that the affixes *ke-/-an*, *peN-/-an*, and *-an* are the most productive noun-forming affixes compared to other affixes. The quantitative approach also allows us to describe the word-class preferences of noun-forming affixes. Our findings indicate that the affix *ke-/-an* tends to be derived from the adjective word, the affix *per-/-an* has an over-representation for noun-based words, and verbs have an over-representation as a base word for the affixes *-an*, *peN-*, and *peN-/-an*. This research implies benefits in applying corpus, quantitative, and qualitative linguistics methodologies in derivational morphology.

### **CITATION**

Denistia, Karlina & Gede Primahadi Wijaya Rajeg. 2023. *Afiksasi Nomina dalam Bahasa Indonesia* (Noun Affixation in Indonesian). Book Chapter Manuscript. To appear in *Tata Bahasa Indonesia Kontemporer: Morfologi* (Contemporary Indonesian Grammar: Morphology). A project for corpus-based Indonesian grammar by the Language Agency of Indonesian ([recording for the talk](#))

## ***Afiksasi Nomina dalam Bahasa Indonesia***

Karlina Denistia<sup>1</sup> dan Gede Primahadi Wijaya Rajeg<sup>2</sup>

<sup>1</sup> Sekolah Vokasi, Universitas Sebelas Maret

<sup>2</sup> University of Oxford/Universitas Udayana

<sup>1</sup> [karlinadenistia@staff.uns.ac.id](mailto:karlinadenistia@staff.uns.ac.id), <sup>2</sup> [gede.rajeg@ling-phil.ox.ac.uk](mailto:gede.rajeg@ling-phil.ox.ac.uk)

### **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

### **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:

Subkategorisasi nomina		Contoh kata
Nomina bernyawa dan tak bernyawa	bernyawa	<ul style="list-style-type: none"><li>● nama diri seperti <i>Joko</i>, <i>Karlina</i></li><li>● kekerabatan seperti <i>nenek</i>, <i>kakek</i></li><li>● orang atau yang diperlakukan seperti orang, misalnya <i>tuan</i>, <i>nyonya</i>, <i>hantu</i>, <i>raksasa</i></li></ul>

		<ul style="list-style-type: none"> <li>• nama kelompok manusia seperti <i>Jepang, Indonesia</i></li> </ul>
	nomina tak bernyawa	<ul style="list-style-type: none"> <li>• nama geografis seperti <i>Inggris, Bali, Jawa, utara</i></li> <li>• nama lembaga seperti <i>DPR, MPR</i></li> <li>• waktu seperti <i>senin, Selasa, Januari, pukul 8</i></li> <li>• ukuran dan takaran seperti <i>kali, pikul, goni, lusin</i></li> <li>• tiruan bunyi seperti <i>aum, dengung, kokok</i></li> </ul>
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 bagiannya)	<ul style="list-style-type: none"> <li>• nomina dasar seperti <i>tentara, keluarga</i></li> <li>• nomina turunan seperti <i>wangi-wangian, tepung-tepungan, minuman</i></li> </ul>
	nomina bukan kolektif (nomina yang tidak diperinci bagiannya)	<i>asinan, cairan, hadirin, keluarga, kawan, kelompok</i>

Tabel 3. Subkategorisasi kata benda

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.

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> <i>-an</i>	tapa	pertapa	agen
	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

Tabel 4. Afiksasi pembentuk nomina dalam bahasa Indonesia

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-menjejakkan-menjejaki*, dan *menapak-menapakkan-menapaki*, (2) perasaan, misalnya *menyesal-menyekali-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 berkaitan 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.

## 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 klitiks 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 alomorf 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	
ajaranaku	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
pengajarlal	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	FreqCerpen	FreqNovel	FreqBukuteks	FreqJurnal	FreqTA	FreqBiografi	FreqPopuler	FreqUU	FreqLaman	FreqSurat
ajaran	66	43	35	36	365	99	257	397	122	4	103	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	100	8
pengajaran	5	9	1	1	117	84	55	63	119	8	54	29
pengajaranku	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
pengajarlal	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
ajaranku	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
pengajaranku	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

### Afiks pembentuk nomina

#### 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:

- a) *pe-* + tugas = petugas  
*Petugas* medis setempat menyatakan, pekan lalu sebanyak 200 orang tewas di Misrata. (sumber: koran, 2011)
- b) *pe-* + renang = perenang  
*Perenang* putri lainnya, Ressa Kania Dewi merebut medali emas di nomor 400. (sumber: koran, 2014)
- c) *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)

2. Imbuhan *peN-*

Imbuhan *peN-* memiliki 6 alomorf atau varian, yaitu *peN<sub>peng-</sub>*, *peN<sub>pen-</sub>*, *peN<sub>pem-</sub>*, *peN<sub>pe-</sub>*, *peN<sub>peny-</sub>*, dan *peN<sub>penge-</sub>*. Kemunculan alomorf tersebut dapat diprediksi secara fonologi sebagai berikut [17], [18]:

- 1) *peN<sub>peng-</sub>* muncul ketika melekat pada kata dasar yang berawalan dengan fonem vokal atau fonem *velar obstruent* /g/, /k/, /h/, or /kh/,
  - *peN-* + huni = penghuni
  - *peN-* + gemar = penggemar
- 2) *peN<sub>pen-</sub>* muncul ketika melekat pada kata dasar yang berawalan dengan fonem alveolar obstruent atau palatal obstruent /d/, /t/, /c/, /j/, /sy/, or /z/,
  - *peN-* + tonton = penonton
  - *peN-* + jual = penjual
- 3) *peN<sub>pem-</sub>* muncul ketika melekat pada kata dasar yang berawalan dengan fonem konsonan *labial* /b/, /p/, atau /f/,
  - *peN-* + balap = pembalap
  - *peN-* + pantau = pemantau
- 4) *peN<sub>pe-</sub>* muncul ketika melekat pada kata dasar yang berawalan dengan fonem nasal, *semivowel*, atau *liquid* /m/, /n/, /ng/, /ny/, /w/, /j/, /r/, or /l/,
  - *peN-* + makan = pemakan
  - *peN-* + luncur = peluncur
- 5) *peN<sub>peny-</sub>* muncul ketika melekat pada kata dasar yang berawalan dengan fonem /s/,
  - *peN-* + saring = penyaring
  - *peN-* + sedap = penyedap
- 6) *peN<sub>penge-</sub>* muncul ketika melekat pada kata dasar yang memiliki satu suku kata.
  - *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-* + latihan = 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  
Dusahakannya 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. 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* = gunung  
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)

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)

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)

#### 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 menggampangkan 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)

#### 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)

#### 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)

## 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*.

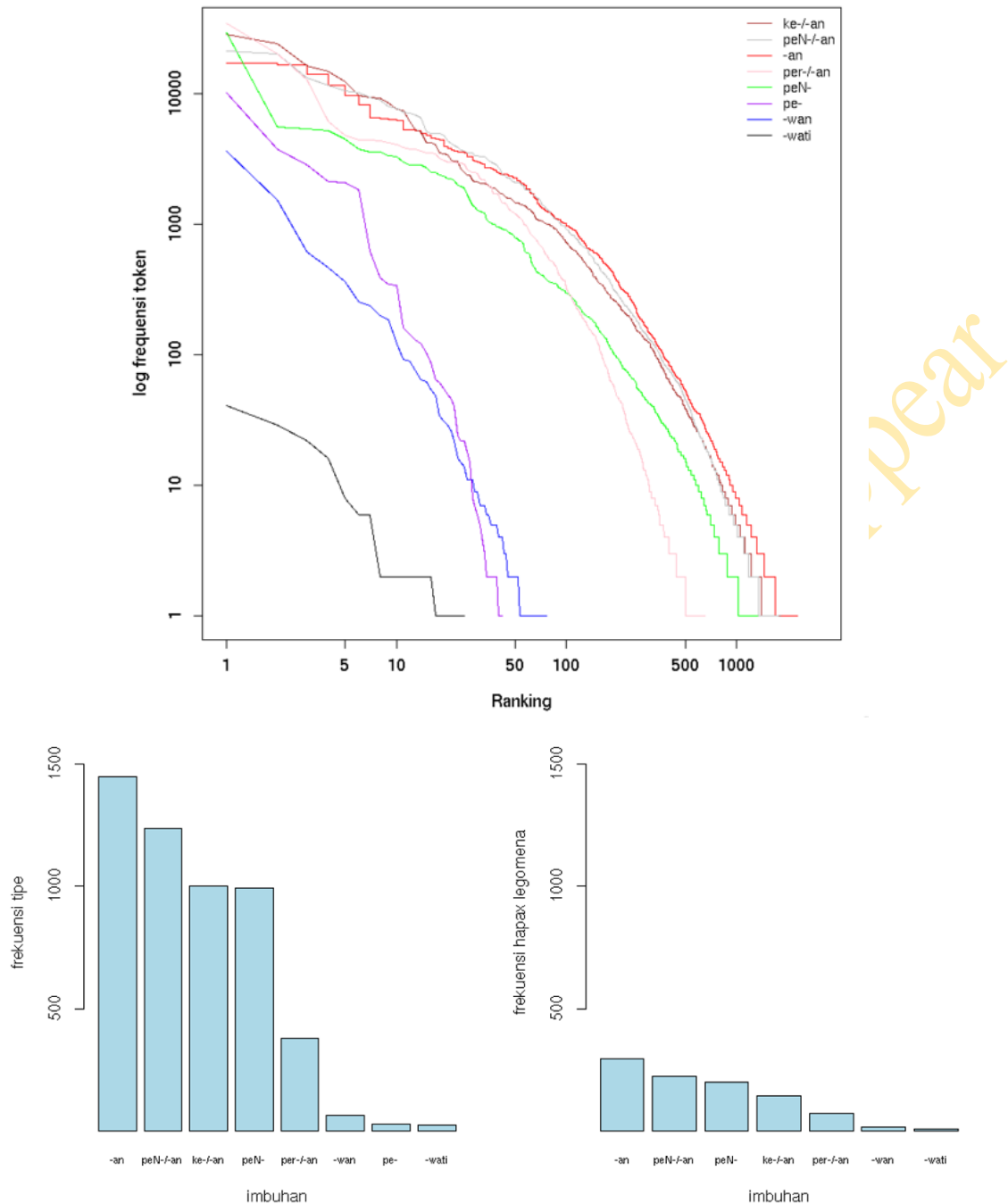
### 1. Distribusi afiks

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 5.

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

Tabel 5. Perhitungan token, tipe, dan hapax legomena untuk setiap imbuhan pembentuk kata benda dalam bahasa Indonesia

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:

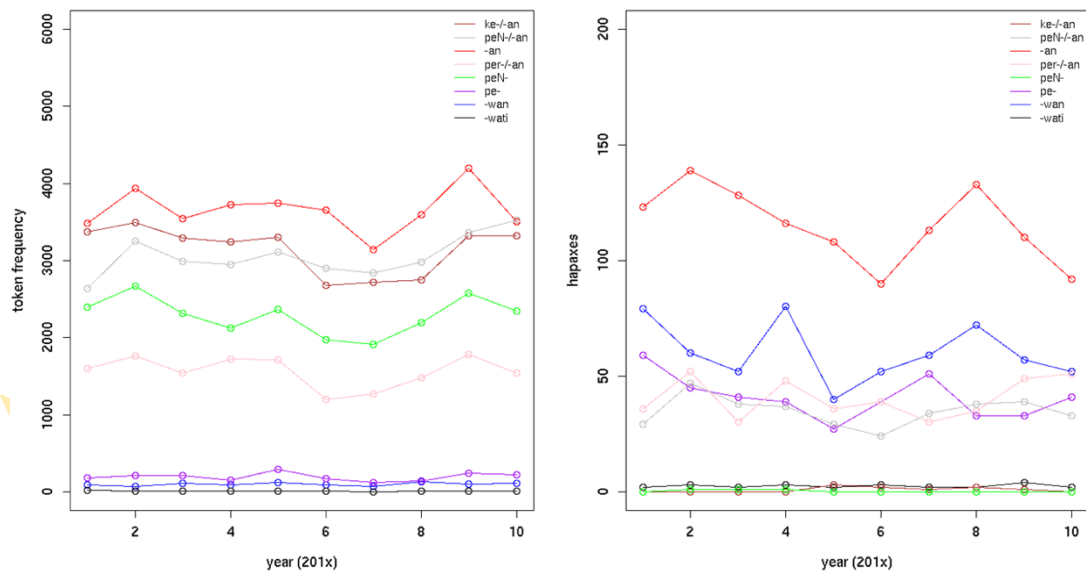
- 1) *peN-/an* : penelitian, pendidikan, pelaksanaan, pengembangan, pembelajaran, pelayanan, pembangunan, pembiayaan, peningkatan, penggunaan
- 2) *-an* : laporan, bagian, lingkungan, tujuan, hubungan, bangunan, jaringan, anggaran, tindakan, makanan
- 3) *ke-/an* : kegiatan, keuangan, ketentuan, kesehatan, 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, pembedanan, 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/) (<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.

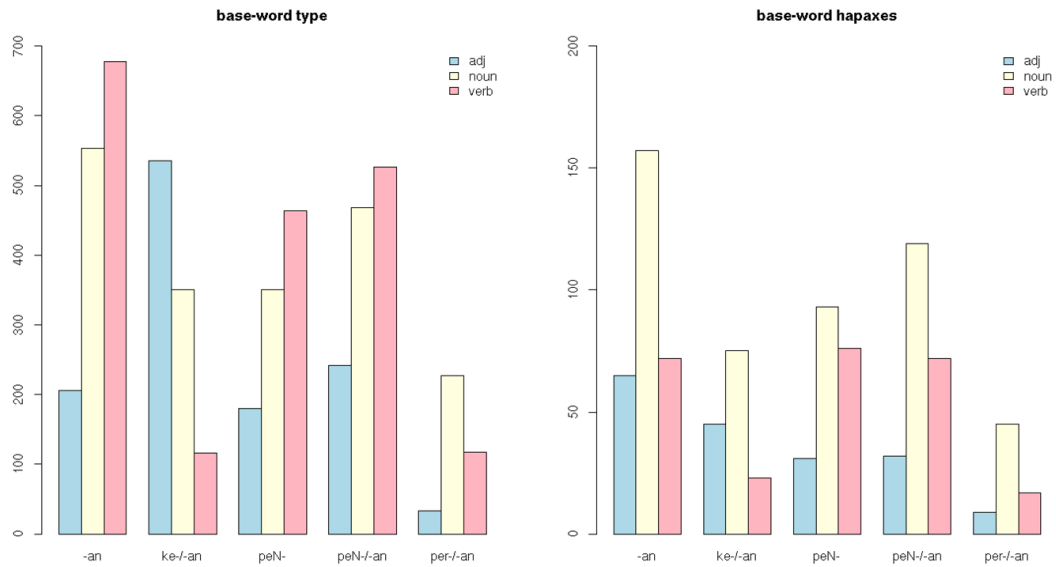
Distribusi imbuhan di dalam korpus yang merupakan produk bahasa tulis tentu saja tidak akan sama dari tahun ke tahun. Bahasa memiliki dinamika, demikian pula dengan dinamika imbuhan yang merupakan salah satu unsur kebahasaan. Gambar 5 menunjukkan perkembangan penggunaan imbuhan pembentuk nomina dalam 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).

## 2. Distribusi kelas-kata kata-dasar

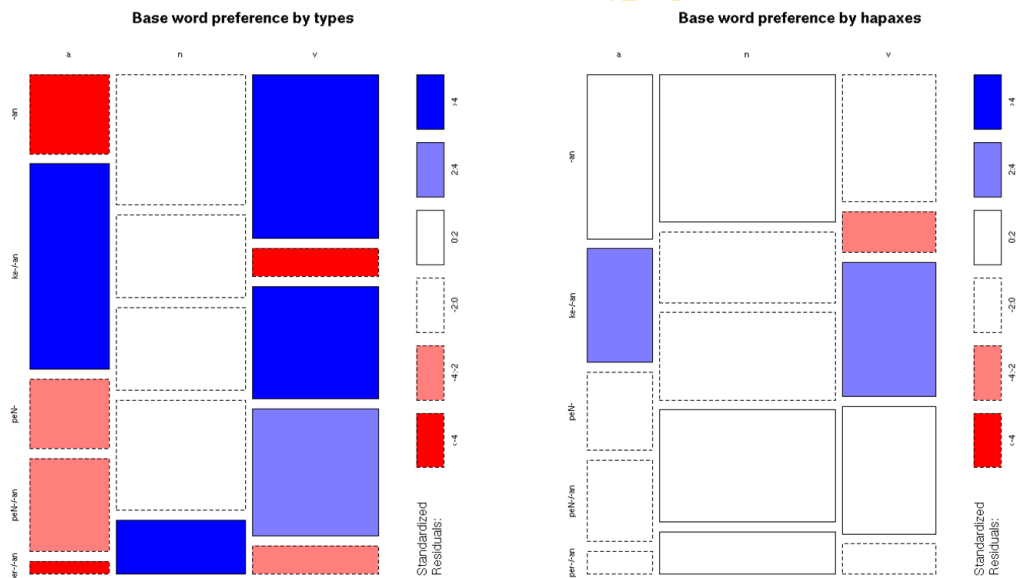
Bagian ini menjelaskan distribusi kelas-kata kata-dasar imbuhan pembentuk kata benda *-an*, *ke-/-an*, *peN-*, *peN-/-an*, dan *per-/-an*. . *Konsep kata dasar dalam penelitian ini adalah kata yang melekat pada imbuhan. Misalnya, kata turunan keterlambatan memiliki kata dasar terlambat.* Kami memilih 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 – kepopolaritasan), *peN-* (koreksi – pengoreksi, desain – pendesain, survey – penyurvei), *peN-/-an* (sifat – penyifatan, bumi – pembumian, asumsi – pengasumsian), dan *per-/-an* (burung – perburungan, kopi – perkopian, limbah – perlimbahan) 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*.

### Imbuhan *pe-* dan *peN-*

Imbuhan *pe-* dan salah satu alomorf *peN-*, *peN<sub>pe-</sub>*, 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<sub>pe-</sub>* yang terdapat pada kata *pelari* dan *pelukis*.

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

Tabel 6. Contoh kata berimbuhan *pe-* dan *peN-* yang memiliki kemiripan bentuk dan makna

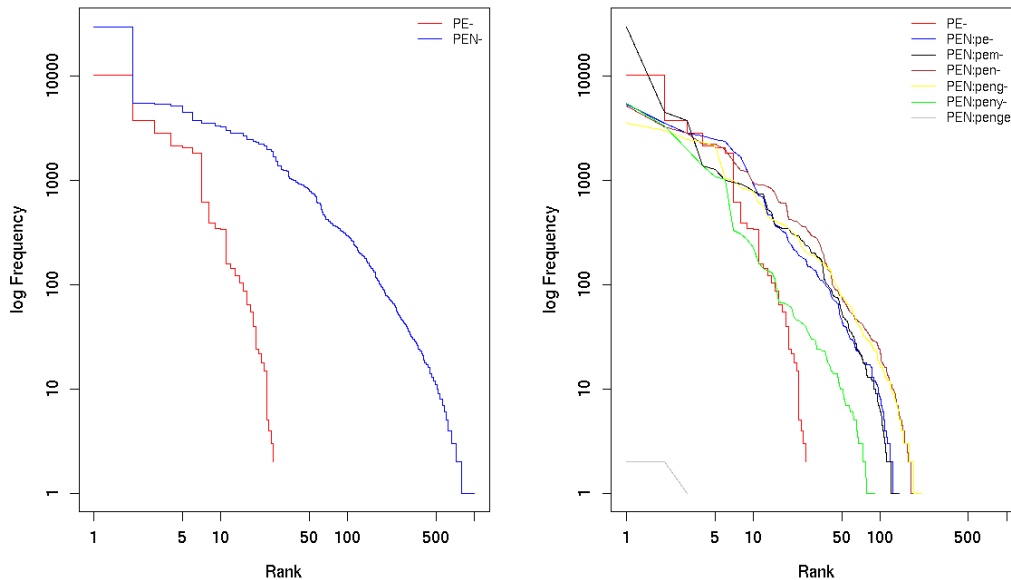
Kemiripan bentuk *pe-* dan *peN-* membuat linguist memiliki dua pendapat yang berbeda mengenai status imbuhan *pe-* dan *peN-* [24]. Sneddon et al. [3] dan Ramlan [13] berpendapat bahwa *pe-* dan *peN-* adalah dua imbuhan yang berbeda, sedangkan Dardjowidjojo [25] dan Kridalaksana [26] berpendapat bahwa *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-* yang 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-*.

### 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.

### Daftar pustaka

- [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. Djenaar, 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.

Accepted Manuscript to Appear