

## **Penyusunan Daftar Kosakata Baca Bahasa Indonesia untuk Siswa Sekolah Menengah Pertama (SMP)**

*Preparation of Indonesian Reading Vocabulary List for Junior High School Students*

<sup>1)</sup>Anis Mardiani Astuti, <sup>2)</sup>Furqanul Aziez

<sup>1,2)</sup>Universitas Muhammadiyah Purwokerto

\*email: [anismardiania@gmail.com](mailto:anismardiania@gmail.com), [f.aziez2010@gmail.com](mailto:f.aziez2010@gmail.com)

### **ABSTRAK**

#### **Histori Artikel:**

Diajukan:  
05/09/2021

Diterima:  
09/10/2021

Diterbitkan:  
08/12/2021

*Penelitian ini dirancang untuk menyusun daftar kosakata baca bahasa Indonesia untuk siswa SMP bagi pembelajaran berpendekatan integratif. Penyusunan daftar kosakata baca bahasa Indonesia ini diharapkan dapat membekali siswa dengan kosakata yang memadai sehingga siswa SMP di Indonesia akan dapat membaca buku teks bahasa Indonesia secara mandiri. Untuk pembuatan bahan ajar penelitian ini akan (1) mengembangkan tiga korpora, yaitu korpus buku teks bahasa Indonesia SMP kelas satu, dua, dan tiga, sebagai basis analisis, (2) mengidentifikasi distribusi lexical set untuk tiap korpus, (3) mengidentifikasi total kata minimum untuk dapat membaca buku teks secara mandiri, dan (4) membuat bahan ajar berupa daftar kosakata baca untuk siswa SMP. Untuk kepentingan pembuatan bahan ajar ini, lematisasi dilakukan dengan manual, sedangkan analisis karakteristik kosakata dengan WebVocabProfiler (Cobb, 2009), yang memberikan data- data tentang level kosakata korpus, kandungan kosakata akademik interdisipliner yang ada dalam Coxhead's (2000) AWL, indikator level kosakata tiap kata, dan kosakata yang ada di luar 2000 dan 5000 keluarga kata pertama (asumsi level kosakata masing-masing kelas).*

**Kata kunci:** Korpus, Level Kosakata, Daftar Kosakata Baca, Bahan Ajar

### **ABSTRACT**

*This study was designed to compile a list of Indonesian reading vocabulary for junior high school students for learning with an integrative approach. The preparation of this list of Indonesian reading vocabulary is expected to equip students with adequate vocabulary so that junior high school students in Indonesia will be able to read Indonesian textbooks independently. For the manufacture of teaching materials, this research will (1) develop three corpora, namely the corpus of first, second, and third grade Indonesian junior high school textbooks, as the basis of analysis, (2) identify the lexical set distribution for each corpus, (3) identify the total words minimum to be able to read textbooks independently, and (4) create teaching materials in the form of reading vocabulary lists for junior high school students. For the purposes of making this teaching material, lematization is done manually, while the analysis of vocabulary characteristics is done with the WebVocabProfiler (Cobb, 2009), which provides data on the corpus vocabulary level, the content of interdisciplinary academic vocabulary contained in Coxhead's (2000) AWL, vocabulary level indicators. each word, and vocabulary beyond the first 2000 and 5000 word families (assuming each grade's vocabulary level).*

**Keywords:** Corpus, Vocabulary Level, List Of Reading Vocabulary, Teaching Mater

## **PENDAHULUAN**

Bahasa merupakan dasar bagi pengetahuan manusia, demikian pula dalam pengajaran Bahasa sebagai inti dan dasar bagi mata pelajaran lainnya, khususnya mata pelajaran Bahasa Indonesia yang pada dasarnya secara keseluruhan melibatkan kemampuan berbahasa Indonesia. Kemampuan berbahasa Indonesia dalam dunia Pendidikan berperan penting dalam keberhasilan pengajaran, baik bagi guru maupun siswa. Selain itu, kemampuan ber-bahasa juga bertujuan untuk meningkatkan kemampuan siswa berkomunikasi secara efektif, baik lisan maupun tertulis dengan pihak lain sesuai dengan konteks dan situasinya.

Sejalan dengan hal tersebut, dalam buku materi pelatihan kurikulum SMP tahun 2016 (Kemendikbud 2016:13) dijelaskan bahwa mata pelajaran Bahasa Indonesia secara umum bertujuan agar siswa mampu menguasai empat keterampilan berbahasa. Siswa belajar tentang fungsi Bahasa Indonesia sebagai sarana berinteraksi secara efektif, baik berkomunikasi secara lisan maupun tulis dengan penggunaan tata Bahasa dan tata kalimat. Pemahaman tentang kemampuan berbahasa mampu membekali siswa sebagai seorang individu yang produktif, dalam berbagai aspek terkhusus dalam berbahasa.

Kurikulum Nasional khususnya mata pelajaran Bahasa Indonesia merupakan kurikulum berbasis teks, meskipun sebelum diberlakukan Kurikulum 2013 Revisi sudah berbasis teks. Melalui teks siswa diharapkan mampu lebih aktif dalam proses pembelajaran. Kurikulum 2013 Revisi 2016 Kemendikbud (2016:4) menyatakan bahwa mata pelajaran Bahasa Indonesia pada umumnya bertujuan agar siswa mampu menyimak, mewicara, membaca dan menulis. Inilah yang disebut sebagai keterampilan berbahasa Keempat keterampilan tersebut memiliki keterkaitan yang sangat erat.

Dari keempat keterampilan ber-bahasa tersebut, membaca merupakan salah satu keterampilan yang penting untuk dikuasai, karena dalam proses pembelajaran siswa melakukan kegiatan membaca. Namun kemampuan membaca masih sangat rendah.

Abidin (2012: 9) berpendapat bahwa rendahnya kemampuan membaca siswa merupakan cerminan kegagalan pembelajaran membaca di sekolah. Maksud dari pernyataan tersebut adalah betapa pentingnya peranan guru dalam merencanakan, melaksanakan, dan mengevaluasi pembelajaran membaca. Jika pembelajaran membaca dapat berlangsung dengan baik, kemampuan membaca pun akan meningkat.

Meningkatkan kemampuan membaca siswa erat kaitannya dengan penguasaan kosakata yang dikuasai. Banyak penelitian yang membuktikan bahwa rendahnya kemampuan membaca disebabkan oleh lemahnya penguasaan kosakata. Sejalan dengan pendapat tersebut, Alwasilah (2012: 169) mengungkapkan “di Indonesia hanya tercatat 2% siswa yang prestasi membacanya masuk ke dalam kategori sangat tinggi, 19% masuk ke dalam kategori menengah, dan 55% masuk ke dalam kategori rendah.”. Jika demikian maka menurut Laufer (1989) sulit sekali bagi mereka untuk menjadi pembaca yang mandiri.

Penguasaan kosakata merupakan hal terpenting dalam keterampilan berbahasa, tanpa penguasaan kosakata yang memadai, maka tujuan pembelajaran Bahasa tidak akan tercapai. Sejalan dengan hal tersebut Tarigan (2015:2) menjelaskan bahwa kualitas keterampilan berbahasa seseorang bergantung pada kuantitas dan kualitas kosakata yang dimilikinya. Semakin banyak kosakata yang dimiliki, semakin besar pula kemungkinan untuk terampil berbahasa.

Kosakata bahasa Indonesia adalah kosakata yang tumbuh dan berkembang sesuai dengan perkembangan zaman. Perkembangan kosakata berarti menempatkan konsep-konsep baru dalam tatanan yang lebih baik atau ke dalam urutan-urutan dan susunan tambahan. Penguasaan kosakata dapat dipengaruhi oleh latar belakang pengetahuan, usia, dan pendidikan serta jumlah referensi bacaan. Kemampuan penguasaan kosakata mampu mengekspresikan ide maupun gagasan secara lisan maupun tulis, hal tersebut membutuhkan pilihan kata yang tepat. Bagaimanapun, meski fakta menunjukkan kuatnya pengaruh kosakata terhadap membaca, perhatian terhadap

kosakata dalam pembelajaran bahasa masih rendah. Hal ini setidaknya dapat teridentifikasi dari bahan ajar yang digunakan di lembaga-lembaga pendidikan formal. Bahan ajar yang digunakan di Indonesia, sebagai contoh, belum memberikan perhatian yang cukup kepada penguasaan kosakata siswa. Persoalan bahan ajar sendiri berawal sebagaimana tercermin pada teks-teks ujian mereka (Aziez, 2011). Kedua, kosakata selama ini hanya diajarkan dengan pendekatan *meaning-focused* atau *implicit-instruction*. Praktik ini bertentangan dengan hasil-hasil penelitian yang menunjukkan bahwa pengajaran yang mengintegrasikan *form-focused* dan *meaning-focused* terbukti efektif meningkatkan penguasaan kosakata (Ellis, 2001; Hulstijn, 2003; Loschky & Bley-Vroman, 1993; Skehan, 1998).

Berangkat dari permasalahan tersebut nampak jelas bahwa perlu sekali dibuatkan bahan ajar yang dikembangkan dari korpora buku teks. Penelitian ini menekankan pada penyusunan daftar kosakata baca Bahasa Indonesia yang terdapat dalam buku teks siswa *Bahasa Indonesia Edisi Revisi* SMP/MTs kelas VII, VIII, dan IX karangan Tim Penyusun Kemendikbud dan diterbitkan oleh Kemendikbud (2017).

Penyusunan daftar kosakata baca sebagai bahan ajar yang akan dikembangkan memerlukan adanya pelaksanaan pendekatan pembelajaran integratif yang efektif untuk diterapkan dalam pembelajaran di Sekolah Menengah Pertama khususnya muatan pelajaran Bahasa Indonesia. Oleh karena itu, pada dasarnya pelaksanaan kegiatan pembelajaran diarahkan agar siswa mencapai kompetensi yang diharapkan.

Berdasarkan permasalahan di atas, perlu dilakukan sebuah penelitian, penelitian tersebut dirancang untuk mengembangkan bahan ajar berupa daftar kosakata baca setidaknya dari dua hal. Pertama, penggunaan teks-teks autentik sebagai sumber bahan ajar, yang nota bene ditulis dan ditujukan untuk penutur asli, yang karenanya memuat banyak sekali kosakata di atas level kosakata siswa, Bahasa Indonesia dengan tujuan untuk membantu penguasaan kosakata baca siswa SMP serta meningkatkan kemampuan

membaca siswa secara mandiri. Penelitian ini ditujukan untuk menghasilkan sejumlah bahan ajar kosakata baca bahasa Indonesia untuk siswa SMP. Secara rinci bahan-bahan ajar tersebut adalah:

1. korpora buku-buku teks SMP berdasar hasil penyusunan penelitian ini.
2. tingkat keterbacaan masing-masing korpus dengan menggunakan Coleman Liau Index, Flesh Kincaid Grade Level, dan ARI (*Automated Readability Index*), dan SMOG. Analisis dilakukan dengan bantuan *Readability Online-Test*,
3. daftar kosakata baca untuk tiap kelas SMP.

## METODE PENELITIAN

Dalam penelitian ini, penulis menggunakan pendekatan kualitatif. Penelitian kualitatif merupakan penelitian yang bersifat deskriptif dan cenderung menggunakan analisis. Proses dan makna lebih ditonjolkan dalam penelitian kualitatif dan landasan teori dimanfaatkan sebagai pemandu penelitian. Kemudian untuk metode penelitian, metode penelitian ini juga menggunakan metode deskriptif. Karena hal itu, penelitian ini menggunakan metode penelitian deskriptif kualitatif.

### Teknik Pengumpulan Data

Data penelitian ini dikumpulkan dengan menggunakan dokumentasi.

### Analisis Data

Data yang diperoleh melalui alat penjarang data tersebut kemudian dianalisis dengan langkah-langkah sebagai berikut.

#### 1. Pengembangan Korpus

Korpus tiap buku dikembangkan dengan mengambil teks-teks yang ada pada buku tiap kelas. Teks-teks yang dijadikan korpus adalah (1) teks yang biasanya digunakan sebagai model genre, sebagai bahan ajar membaca, acuan pertanyaan pemahaman, dan sebagai sumber pengajaran komponen-komponen bahasa; teks ini umumnya muncul pada awal unit bab; (2) penjelasan teoretis tentang genre yang tengah diajarkan; dan (3)

perintah dan/ atau tugas dan latihan. Teks-teks ini kemudian diretrieve, discan optik, atau diketik secara manual dan diklasifikasi berdasar kelas dan genre. Dengan klasifikasi korpora seperti ini maka untuk berbagai keperluan akan dapat diperoleh (a) korpus buku teks SMP secara keseluruhan, (Korpus BTS); (b) korpora dari tiga kelas (Korpus BT1, Korpus BT2, Korpus BT3); dan (c) korpora dari genre-genre yang ada di seluruh kelas (Korpus Recount, Korpus Descriptive, dst).

## 2. Pengembangan bahan ajar

Untuk dapat memperoleh bahan ajar dan menyusun tes kosakata, perlu disusun daftar kata baca. Untuk itu kata-kata yang terdapat dalam tiap korpus dilakukan penyusunan dengan mengekstrak kata dasar menggunakan bantuan KBBI Online, dengan menginput secara manual satu per satu kata yang terdaftar. Dengan proses ini bentuk- bentuk kata turunan seperti *menyapu* akan terdaftar di bawah kata dasar *sapu*. Sementara itu untuk penyortiran nama diri, angka-angka, simbol unit ukuran, singkatan, dan kata-kata kali, dilakukan secara manual. Selanjutnya daftar-daftar kata ini dimasukkan dalam program Cobb's (2009) *Web VocabProfiler* pada K1 – K20 untuk menghasilkan daftar yang sudah teranking berdasar frekuensinya. Selain ditentukan frekuensi tiap katanya, daftar kata tiap kelas ini juga diidentifikasi *grade level*-nya dengan mengacu pada *The Basic Elementary Reading Vocabularies*. Terakhir, kata-kata pada tiga daftar kata ini diklasifikasikan kategori kelas kata bahasa Indonesia, yakni kata benda, kata kerja, kata sifat, dan keterangan.

## HASIL DAN PEMBAHASAN

### Hasil Analisis Data

Data dalam penelitian ini diperoleh dari buku buku pelajaran bahasa Indonesia SMP, peneliti menggunakan 3 buah buku pelajaran untuk mendapatkan kosakata tersebut, ketiga buku tersebut merupakan buku siswa *Bahasa Indonesia untuk SMP/MTs kelas VII, VIII, dan IX* yang ditulis dan diterbitkan oleh Kemendikbud. Dari buku pelajaran tersebut peneliti mendapatkan kosakata baik itu dari percakapan atau teks.

## 1. Pengembangan Kosakata Baca Berbasis Korpus untuk Siswa SMP

Pengembangan kosa-kata di ke-hidupan sehari-hari berbasis korpus diambil dari 3 buku pelajaran bahasa Indonesia sekolah menengah pertama yang digunakan dikelas VII, VIII dan IX. Korpus ini di ambil dari teks dan percakapan di 3 buku pelajaran. Kemudian teks dan percakapan di ketik oleh peneliti, setelah semua teks di ketik, kosa kata- kosa kata di kumpulkan.

Kemudian peneliti membuat daftar dari kosa kata kemudian kosa kata diurutkan sesuai dengan hasil analisis yang dibantu program *Text Document Readability*, setelah itu semua tipe dari kosa kata dihitung oleh peneliti. Setelah semua tipe dari kosa kata dihitung hasil dari persentase dari masing-masing buku akan diketahui. Untuk menghitung level tingkatan kosa kata, peneliti menggunakan program *Vocabprofil*, setelah itu membuat urutan dari bahasa dan kategori kedalam kategori kelas kata. Yang terakhir adalah menghitung persentase dari masing-masing bagian kata.

## 2. Tingkat Kesesuaian Kosakata untuk Sekolah Menengah Pertama

Seperti yang peneliti katakan pada Bab 2 bahwa level untuk siswa menengah pertama berada pada K-2, itu maksudnya bahwa siswa-siswa sanggup untuk berbicara dan berkomunikasi dengan baik ketika mereka mengetahui 2000 kata pertama dari buku pelajaran. Berdasarkan data dari Corpus Brown yang dikutip dalam Aziez (2014) dua ribu kata pertama sama dengan 79,7%, itu maksudnya bahwa untuk siswa sekolah menengah pertama, 2000 kata pertama cukup untuk memahami isi dari teks dan berkomunikasi dengan baik.

Peneliti akan memaparkan hasil dari K-1 dan K-2 kata di 3 buku pelajaran bahasa Indonesia. Jika hasilnya menunjukkan nomor persentase 80%, berarti siswa sudah mengetahui, berarti siswa sudah mengetahui 80% kat di buku pelajaran, dengan kata lain mereka tidak mengetahui 20 % kata dari buku pelajaran, karena mengetahui 80 % kata dibuku pelajaran.

Tabel 4.1 Persentase bagian bahasa dari 3 buku pelajaran

No	Nama buku	Kata benda (%)	Kata kerja (%)	Kata sifat (%)	Keterangan (%)
1	Bahasa Indonesia SMP/MTs kelas VII Edisi Revisi 2017	33,64%	31,33%	29,12%	5,89%
2	Bahasa Indonesia SMP/MTs kelas VIII Edisi Revisi 2017	32,96%	34,08%	29,50%	3,45%
3	Bahasa Indonesia SMP/MTs kelas IX Edisi Revisi 2018	38,78%	38,57%	18,17%	4,46%

Hasil dari buku pelajaran Bahasa Indonesia kelas VII

#### Rasio dan indeks terkait

Berkenaan dengan seluruh teks

Kata-kata dalam teks (token) 38337

Kata-kata berbeda (tipe) 5292

Mengenai data saja

Token 25642

Tipe-tipe 1515

Kelompok 1446

Token per kelompok 17.73

Tipe per kelompok 1.05

#### Tabel persentase kosakata buku Bahasa Indonesia kelas VII

Tabel 4.2

Persentase dari K-1 dan K-2 untuk buku Bahasa Indonesia SMP/MTs kelas VII

Fre. Level	Kelompok (%)	Tipe (%)	Token (%)	Cumul. token %
K-1 Words :	1247 (86.2)	1249 (23.60)	24327 (63.5)	63.5
K-2 Words :	36 (2.5)	36 (0.68)	95 (0.2)	63.7

K-1 Words berarti bahwa daftar 1000 kata pertama yang muncul dalam teks, itu sama dengan k-2 Words, di sini peneliti hanya mengambil sampai K-2 Words karena untuk berkomunikasi dalam kehidupan sehari-hari di

SMP cukup memiliki 2000 kata pertama dari teks pada buku pelajaran. Daftar dari 2000 kata pertama. Daftar dari 2000 kata pertama terlampir dalam laporan, ada dua warna dari

kosa kata itu, ada biru dan hijau, Biru untuk K-1 dan Hijau untuk K-2.

**Buku Bahasa Indonesia SMP/MTs kelas VIII Edisi Revisi 2017**

Hasil dari buku pelajaran bahasa Indonesia kelas VIII

**Rasio dan indeks terkait**

Berkenaan dengan seluruh teks	
Kata-kata dalam teks (token)	40130
Kata-kata yang berbeda (tipe)	5613
Mengenai daftar saja	
Token	23899
Tipe	1447
Kelompok	1345
Token per kelompok	17.77
Tipe perkelompok	1.08

**Tabel Persentase kosakata dari buku bahasa Indonesia kelas VIII**

Tabel 4.3.

Persentase dari K-1 dan K-2 untuk buku bahasa Indonesia kelas VIII

Freq. Level	kelompok (%)	Tipe (%)	Token (%)	Cumul. token %
K-1 Words :	1172 (87.1)	1175 (20.93)	22530 (56.1)	56.1
K-2 Words :	29 (2.2)	30 (0.53)	167 (0.4)	56.5

**Daftar dari 2000 kata pertama**

Seperti yang peneliti katakan di atas, ada dua warna dari kosakata tersebut, yang biru untuk K-1 dan yang hijau untuk K-2.

**a. Bahasa Indonesia SMP/MTs kelas IX Edisi Revisi 2018**

Hasil dari buku pelajaran bahasa Indonesia kelas IX

**Rasio dan indeks terkait**

Berkenaan dengan seluruh teks	
Kata-kata dalam teks (token)	20294
Kata-kata yang berbeda (tipe)	3951
Mengenai daftar saja	
Token	8511
Tipe	830
Kelompok	775
Token per kelompok	10.98
Tipe perkelompok	1.07

**Tabel Persentase kosakata dari buku bahasa Indonesia kelas IX**



Tabel 4.4  
Persentase dari K-1 dan K-2 dari buku bahasa Indonesia kelas IX

Freq. Level	kelompok (%)	Tipe (%)	Token (%)	Cumul. token %
<b>K-1 Words :</b>	612 (79.0)	614 (15.54)	7667 (37.8)	37.8
<b>K-2 Words :</b>	31 (4.0)	33 (0.84)	95 (0.5)	38.3

#### Daftar dari 2000 kata pertama

Pada daftar dari kosakata itu, ada dua jenis warna, ada biru dan hijau, Yang biru berarti kosakata tersebut termasuk dalam 1000 kata pertama dalam teks (K-1 Words). Yang lainnya berwarna hijau, artinya kosakata tersebut termasuk dalam K- 2 Words.

Hasil persentase token kumulatif lebih tinggi pada jumlah kosakata karena peneliti hanya mengambil sebagian bahasa saja, tidak semua kosakata dihitung. Di sini peneliti hanya mengambil Kata K-1 dan K-2 karena bagi siswa SMP mereka dapat berkomunikasi dalam kehidupan sehari-hari jika sudah mengenal 2000 kata pertama pada teks.

### 3. Kelas Kata

Kelas kata merupakan klasifikasi kata, penggolongan kata atau penjenisan kata. Jenis-jenis kelas kata dalam bahasa Indonesia menurut Kridalaksana (1939):

#### Jenis-jenis kelas kata Nomina (Kata Benda)

Nomina adalah kategori yang secara sintaksis tidak mempunyai potensi untuk bergabung dengan partikel tidak dan mempunyai potensi untuk didahului oleh partikel dari. Nomina berbentuk:

- 1) Nomina dasar, seperti *radio, udara, kertas, barat, kemarin, dll.*
- 2) Nomina turunan, terbagi atas:
  - a) Nomina berafiks, seperti keuangan, perpaduan, gerigi.
  - b) Nomina reduplikasi, seperti Gedung-gedung, tetamu, pepatah.
  - c) Nomina hasil gabungan proses, seperti batu-batuan, kesinambungan.
  - d) Nomina yang berasal dari pelbagai kelas karena proses.

#### Verba (Kata Kerja)

Kata dikatakan berkategori verba jika dalam frasa dapat didampingi partikel tidak

dalam konstruksi dan tidak dapat didampingi partikel *di, ke, dari*, atau dengan partikel seperti *sangat, lebih, atau agak*. Berdasarkan bentuknya, verba dibedakan menjadi:

**Verba Dasar bebas** Adalah verba yang berupa morfem dasar bebas

Contoh: *nonton, makan, mandi, minum, pergi, pulang, lari, loncat.*

**Verba Turunan** Adalah verba yang telah mengalami afiksasi, reduplikasi, gabungan proses atau berupa paduan leksem. Bentuk turunannya, yaitu:

**Verba berafiks**

Contoh: *berdandan, terbayang, kerinduan, kecelakaan, memasak, bekerja, menjalani.*

**Verba berduplikasi**

Contoh: *lari-lari, ingat-ingat, maju- maju, semangat-semangat, malas-malas.*

Verba berproses gabungan Contoh: *bercanda-canda, tersenyum- senyum,terbayang bayang, berandai-andai*. Verba majemuk Contoh: *buah tangan, cuci mata, unjuk gigi, adu domba, campur tangan, main hakim.*

#### Ajektiva (Kata Sifat)

Ajektiva adalah kategori yang ditandai oleh kemungkinannya untuk bergabung dengan partikel tidak, mendampingi nomina, atau didampingi partikel seperti *lebih, sangat, agak*, mempunyai ciri-ciri morfologis seperti *-er* (dalam honorer), *-if* (dalam sensitif), dan *-I* (dalam alami), dan dibentuk menjadi nomina dengan konfiks *ke-an* seperti keyakinan. Dari bentuknya ajektiva dapat dibedakan menjadi:

Ajektiva dasar Dapat diuji dengan kata *sangat, lebih*, misalnya: *adil, agung, Bahagia, bersih, cemberut, canggung, dungu, disiplin, dan lain-lain.*

Tidak adapat diuji dengan kata sangat, lebih, misalnya: *buntu, cacat, gaib, genap, interlocal, kejur, lancing, langsung, laun, musnah, niskala, pelak, tentu, tunggal, dan sebagainya*. Ajektiva turunan Ajektiva turunan berafiks misalnya *terhormat*. Ajektiva bereduplikasi, misalnya *ringan-ringan*. Ajektiva berafiks R-an atau ke-an, misalnya *kemalu-maluan*.

Ajektiva berafiks -i, misalnya *alami, alamiah (alam)*. Ajektiva yang berasal dari pelbagai kelas dengan proses-proses.

### Adverbia (Keterangan)

Adverbial adalah kategori yang dapat mendampingi ajektiva, numeralia, atau proposisi dalam konstruksi sintaksis. Adverbial tidak boleh dikacaukan dengan keterangan, karena adverbial merupakan

konsep kategori, sedangkan keterangan merupakan konsep fungsi. Bentuk adverbial: Adverbial dasar bebas, contoh:

*alangkah, agak, akan, belum, bisa.*

Adverbial turunan, yang terbagi atas:

Dari keseluruhan kelas kata bahasa Indonesia, peneliti mengambil 4 kelas kata berikut kategorinya sebagai klasifikasi kata yang dijadikan daftar kosakata. Dalam penelitian ini, peneliti mengklasifikasikan kata yang didapatkan dari hasil analisis korpus-korpus 3 buku teks pelajaran bahasa Indonesia yang digunakan oleh SMP. Buku teks yang digunakan oleh peneliti ialah buku teks siswa bahasa Indonesia kelas VII, VIII, dan IX yang dibuat dan diterbitkan oleh tim kemendikbud serta menjadi salah satu buku wajib siswa di tiap sekolah menengah pertama di seluruh Indonesia.

Berikut ini hasil dari persentase 3 buku pelajaran bahasa Indonesia

Tabel 4.5.

persentase dari masing-masing bagian kata buku bahasa Indonesia kelas VII

No	Nama buku	Kata benda (%)	Kata kerja (%)	Kata sifat (%)	Keterangan (%)
1	Bahasa Indonesia SMP/MTs kelas VII Edisi Revisi 2017	33,64%	31,33%	29,12%	5,89%
2	Bahasa Indonesia SMP/MTs kelas VIII Edisi Revisi 2017	32,96%	34,08%	29,50%	3,45%
3	Bahasa Indonesia SMP/MTs kelas IX Edisi Revisi 2018	38,78%	38,57%	18,17%	4,46%

Setelah mengetahui daftar kosakata, peneliti mengetahui masing-masing bagian dari setiap kata, persentase untuk masing-masing bagian dari kata dalam buku ini adalah 33,64% untuk Kata Benda, 31,33% untuk Kata Kerja, 29,12% untuk Kata Sifat, dan 5,89% untuk Keterangan, Buku bahasa Indonesia kelas VIII Setelah mengetahui daftar kosakata, peneliti mengetahui masing-masing bagian dari setiap kata, persentase untuk masing-masing bagian dari kata dalam buku ini adalah 32,96% untuk Kata Benda,

34,08% untuk Kata Kerja, 29,50% untuk Kata Sifat, dan 3,45% untuk Keterangan. Buku bahasa Indonesia kelas IX Setelah mengetahui daftar kosakata, peneliti mengetahui masing-masing bagian dari setiap kata, persentase untuk masing-masing bagian dari kata dalam buku ini adalah 38,78% untuk Kata Benda, 38,57% untuk Kata Kerja, 18,17% untuk Kata Sifat, dan 4,46%. Setelah mengetahui daftar kosakata dan juga persentase dari masing-masing kata tiap buku-buku pelajaran, peneliti menjelaskan tentang kategori kelas kata dari



bentuknya menurut Kridalaksana, yang sudah dipaparkan sebelumnya. Kategori kelas kata yang digunakan ialah kata benda, kata kerja, kata sifat dan keterangan.

Berdasarkan daftar kosakata, ada beberapa kategori pada kata benda, kata kerja, kata sifat, dan keterangan, itu berarti setiap bagian dari bahasa memiliki kategori masing-masing. Pertama, untuk kata benda ada 2 jenis kategori, yakni kata dasar dan turunan, untuk kata dasar ada konkret dan abstrak. Kedua untuk kata kerja ada 2 jenis kategori, yakni dasar bebas dan turunan berafiks. Kemudian untuk kata sifat peneliti membagi kedalam 3 jenis kategori, yakni dasar, turunan, denumeralia. Terakhir adalah keterangan, ada 2 kategori dipandang dari bentuk kelas katanya, yakni kata dasar bebas dan turunan.

### **Pembahasan**

Penelitian ini untuk mengetahui bagaimana menciptakan kosakata kehidupan sehari-hari untuk siswa SMP berbasis Korpus. Kosakata kehidupan sehari-hari berarti kosakata yang digunakan oleh siswa dalam keseharian untuk aktivitas sehari-hari mereka, untuk memperoleh data peneliti menggunakan Dokumenter sebagai sumber utamanya yang digunakan adalah buku pelajaran SMP, peneliti menggunakan 3 buah buku pelajaran untuk mendapatkan kosakata tersebut, buku-buku tersebut merupakan buku pelajaran bahasa Indonesia kelas VII, VIII, IX yang merupakan karangan dan diterbitkan oleh Kemendikbud. Dari buku pelajaran tersebut peneliti mendapatkan kosakata baik itu dari percakapan atau teks.

Setelah mengetahui bagaimana cara untuk menciptakan kosa kata harian berbasis Korpus, peneliti menjelaskan tentang pentingnya level tingkatan kosakata untuk siswa SMP dan juga kategori semantik dalam buku tersebut. Tingkat kosakata di nilai dengan menggunakan program *Vocabprofil*. Melalui program ini peneliti mengetahui level tingkatan kosakata yang tepat untuk siswa SMP dan juga persentase.

Selanjutnya peneliti juga menjelaskan tentang kategori kelas kata berdasarkan bentuknya yang muncul dalam buku-buku pelajaran itu, sebelum peneliti menemukan kategori kelas

kata berdasarkan bentuknya, peneliti membagi kedalam 4 bagian utama kata. yaitu kata benda, kata kerja, kata sifat, dan keterangan. Kemudian untuk masing-masing bagian dari bahasa memiliki kategorinya sendiri, peneliti juga menjelaskan tentang persentase kata dari masing-masing buku.

### **1. Pengembangan kosakata baca untuk siswa SMP berbasis korpus**

Pengembangan kosakata kehidupan sehari-hari berbasis Korpus diambil dari 3 buku pelajaran SMP yang digunakan di kelas VII, VIII dan IX. Korpus ini di buat dari pengumpulan teks dan percakapan di 3 buku pelajaran kemudian teks dan percakapan diketik oleh peneliti, setelah semua teks diketik kosakata dikumpulkan.

Kemudian peneliti membuat daftar kosakata menurut urutan sesuai hasil analisis yang dibantu program *Text Document Readability*. Setelah itu semua kosakata dihitung oleh peneliti, setelah semua kosakata dihitung hasil persentase dari masing-masing buku akan diketahui, untuk menghitung level tingkatan kosakata peneliti menggunakan

program *Vocabprofil*, untuk mendapatkan persentase dari masing-masing buku pelajaran data dimasukkan kedalam program ini, setelah data dimasukkan kedalam program ini kemudian klik "*submit windows*" dalam program, kemudian hasil dari persentase dari tanda kumulatif dikenal dengan menggunakan program tersebut.

Setelah itu buat berdasarkan bagian kata, daftarkan semua kata kedalam bagian yang sama dari kata, setelah kosakata didaftarkan kedalam bagian yang sama dari kata, berikan kategori kedalam setiap kata dibagian kata. Langkah terakhir adalah menghitung persentase dari setiap bagian kata, persentase dari setiap bagian kata dihitung dari masing-masing buku pelajaran SMP.

### **2. Tingkatan Kosakata yang Tepat untuk siswa SMP**

Untuk mengetahui level tingkatan kosakata peneliti menggunakan program *Vocabprofil*, program ini adalah program online. seperti yang telah peneliti katakan pada bab 2 tentang bagaimana menggunakan

program ini, level tingkatan kosakata dengan otomatis akan dikenali.

Peneliti hanya mengambil K-1 dan K-2 Words, hasilnya memperlihatkan bahwa buku-buku terdiri dari 2000 kata pertama, yang berarti buku itu cocok untuk siswa SMP. Sebagaimana Hirs dan Nation di Aziez (2014), menyatakan bahwa kata 2000 pertama merupakan dasar untuk siswa berkomunikasi di kehidupan sehari-hari mereka.

Hasil yang diperoleh dari semua kata pada 3 buku pelajaran. Pertama peneliti menguraikan tentang hasil dari buku *Bahasa Indonesia untuk SMP/MTs Kelas VII Edisi Revisi 2017*. Persentase kumulatifnya adalah 63.5% itu terdiri dari pengelompokan jenis dan tanda. Untuk K-1 Words tanda kumulatifnya adalah 63.5%, dan untuk K-2 Words hasilnya adalah 63.7%. Untuk K-1 Words kelompok katanya adalah 1247 (86.2%) dan K-2 Words kelompok katanya adalah 36 (2.5%), jumlah jenis dari tanda adalah 95 (0.2%) dan 36 (0.68%) untuk itu kosakata yang siswa ketahui dalam buku ini sudah cukup, cukup di sini maksudnya siswa bisa berkomunikasi di kehidupan sehari-hari menggunakan jenis dan kosakata tersebut.

Hasil yang kedua dari buku *Bahasa Indonesia SMP/MTs kelas VIII Edisi Revisi 2017* dalam buku ini kosakata yang ditemukan peneliti adalah 40130 kata. Itu diambil dari semua kata dalam buku pelajaran ini, kemudian untuk hasil kumulatif token dari K-1 dan K-2 adalah 56.1% dan 56.5%, angka dari tipe kata adalah 1175 dan 30. Kemudian jumlah simbol itu 22530 dan 167 kata. Untuk K-1 Words kelompok katanya adalah 1172 (87.1%) dan K-2 Words kelompok katanya adalah 29 (2.2%).

Hasil terakhir adalah dari buku *Bahasa Indonesia untuk SMP/MTs kelas IX Edisi Revisi 2018*, jumlah kosakatanya adalah 20294 kata, dari angka itu peneliti mendapatkan hasil presentase token kumulatifnya adalah 37.8% dari K-1 Words. Hasil untuk kedua kata token dan jenisnya adalah 37.8% (7667) dan 15.54 % (614) jenis kata, itu hasil dari 1000 kata pertama dari buku ini, hasil kedua untuk K-2 Words jumlah persentase token kumulatif adalah 38.3% sedangkan jumlah dari token kata adalah 6.39% (672) dan untuk jumlah

jenisnya adalah 18.01% (410) tipe kata. Dapat disimpulkan bahwa buku itu dapat digunakan untuk siswa SMP.

Ketiga hasil dari buku pelajaran bahasa Indonesia tersebut, dapat disimpulkan bahwa buku bahasa Indonesia untuk kelas VII dan VIII memiliki persentase dibawah 80% yang berarti bahwa kecakapan penguasaan kosakata dan pemahaman siswa dianggap masih kurang. Sedangkan untuk kelas IX persentase yang dihasilkan ialah kurang dari 50%, jauh dari minimal persentase yang seharusnya siswa kelas IX SMP kuasai. Dari 95 % (kategori ideal bagi siswa yang bisa paham dan berbicara dengan baik. Mereka hanya mengetahui kurang dari 80 %. Itu berarti buku tersebut kurang bagus digunakan oleh siswa SMP.

## KESIMPULAN

Berdasarkan tujuan penelitian, penulis sudah melaksanakan penelitian dengan temuan di atas dapat ditarik kesimpulan bahwa dari 3 buku yang diteliti persentase tertinggi dari bagian bahasa adalah Kata Benda dan persentase terendah adalah Keterangan. Jumlah pengkategorian atau pengklasifikasian kata yang ditemukan dan disusun menjadi sebuah daftar kosakata bervariasi dari tiap bukunya, hal tersebut disebabkan oleh jumlah kosakata temuan pada tiap corpora buku teks. Semakin banyak jumlah kosakata di setiap bagian kata, semakin besar pula persentasenya.

## DAFTAR PUSTAKA

- Abidin, Yunus. (2012). *Pembelejaran Membaca Berbasis Pendidikan Karakter*. Bandung: Refika Aditama.
- Alwasilah, A. Chaedar. (2012). *Pokoknya Kualitatif*. Bandung: PT. dunia Pustaka Jaya
- Anderson, R., and W. Nagy. 1991. Word meanings. In R. Barr, M. Kamil, P. Mosenthal, and P.D. Pearson, (Eds.), *Handbook of Reading Research*, Vol. 2, pp. 690–724. New York: Longman.

- Aziez, Furqanul dan Feisal Aziez. (2019). *Kosakata Teori, Pengajaran, dan Pengukurannya*. Purwokerto: UM Purwokerto Press. 272.
- Aziez, Furqanul, Maman Suryaman dan Suwanto. (2020). *Ensiklopedia Pendidikan Indonesia*. Depok: PT RajaGrafindo Persada.
- Aziez, Furqanul. (2011). *Examining the Vocabulary Levels of Indonesia's English National Examinatin Texts*. *Asian EFL Journal*. Professional Teaching Articles – CEBU Issue. Vol. 51 April 2011
- Beck, I. L., McKeown, M. G., and Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford
- Biemiller, A. (2005). Size and sequence in vocabulary development: Implications for choosing words for primary grade vocabulary instruction. In E. H. Hiebert and M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 223–242). Mahwah, NJ: Lawrence Erlbaum. Retrieved August 18, 2009, from PsycINFO database
- Brog, W. R and Gall, M.D. (2003). *Educational Research: An Introduction Edition*. London: Longman Inc.
- Chujo, Kiyomi, et al. (2011). Creating a Corpus-Based Daily Life Vocabulary for TEYL. *Asian EFL Journal*. Professional Teaching Articles. Vol. 49 January 2011.
- Cooper, P. (1984). *The Assessment of Writing Ability: A Review of Research*. Princeton, NJ: Educational Testing Services
- Cornu, A.M. (1979). The first step in vocabulary teaching. *Modern Language Journal* Vol. 53. 262-
- Cunningsworth, Alan. (1995). *Choosing Your Course-Book*. Oxford: Heilnemann.
- Depdiknas. (2008). *Panduan Pengembangan Bahan Ajar*. Jakarta: Depdiknas Direktorat Jenderal Manajemen Pendidikan Dasar dan Menengah Direktorat Pembina SMA.
- Durkin, D. (1979). What classroom instruction has to say about reading comprehension instruction. *Reading Research Quarterly*, 14, 481–533.
- Ellis, R. (2001). Introduction: Investigating Form Focused Instruction. *Language Learning*. Vol 51(1) pp. 1-46
- Harjasujana, Ahmad Slamet. (2006). *Keterampilan Membaca*. Makalah disajikan pada Diklat Membaca, Menulis, dan Apresiasi Sastra (MMAS) bagi Guru Bahasa dan Sastra Indonesia Tingkat Sekolah Menengah, Jakarta: Depdiknas.
- Heryadi, Dedi. (2015). *Metode Penelitian Pendidikan Bahasa*. Bandung: Pustaka Billah.
- Hiebert, E. H., & Kamil, M. L. (2005). *Teaching and learning vocabulary*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Hunt, Alan dan David Beglar. (2005). Framework for developing EFL reading vocabulary. *Reading in a Foreign Language*. Volume 17, No. 1, April 2005
- Kamil, M. L. (2004). Vocabulary and comprehension instruction: Summary and implications of the National Reading Panel findings. In P. McCardle and V. Chhabra (Eds.), *The voice of evidence in reading research*. Baltimore, MD: Paul H. Brookes.
- Kasno. (2004). *Kamus sebagai Sumber*

- Rujukan dan Pengajaran Kosakata.*  
Jakarta: Pusat Bahasa.
- Kemendikbud. (2016). *Bahasa Indonesia*.  
Jakarta: Kemendikbud.
- Kridalaksana, Harimurti. (1939). *Kelas Kata dalam Bahasa Indonesia*. Jakarta: Gramedia Pustaka Utama.
- Majid, Abdul. 2007. *Perencanaan Pembelajaran dan Mengembangkan Standar Kompetensi Guru*. : PT Remaja Rosda Karya
- McKay, Sandra L. (1980). Developing Vocabulary Materials with a Computer Corpus. *RELC Journal* Vol. 11 No. 2 Dec 1980
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Salah, Maher S. (2008). The relationship between vocabulary knowledge and reading comprehension of authentic Arabic texts. Brigham Young University.
- Sugiyono. (2015). *Metode Penelitian Pendidikan*. Bandung: Alfabeta.
- Tarigan, Henry Guntur. (1991). *Metodologi Pengajaran Bahasa*. Bandung: Angkasa.
- Tarigan, Henry Guntur. (2015). *Berbicara Sebagai Suatu Keterampilan Berbahasa*. Bandung: Angkasa
- Tomlinson, B. (ed.) 1998. *Materials Development in Language Teaching*. Cambridge: Cambridge University Press.
- Trianto, Agus. (2005). *Pengembangan Model Bahan Ajar: Penelitian dan*
- Pengembangan Model Bahan Ajar Bahasa Indonesia untuk SLTP Kelas 7 sebagai Implementasi Kurikulum Berbasis Kompetensi.*