

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Linguistik korpus merupakan studi tentang bahasa yang digunakan melalui korpora yang dikumpulkan untuk berbagai tujuan. Beberapa korpora dimaksudkan dalam penelitian linguistik digunakan untuk menjawab pertanyaan tentang prosodi, leksikal, tata bahasa, pola wacana atau pragmatik bahasa (Kennedy, 1998: 3). Seiring dengan perkembangan teknologi, pada tahun 2018, Kemendikbud Indonesia meluncurkan sebuah situs daring yang bernama Korpus Indonesia (KoIn). Korpus ini dibuat dengan tujuan untuk memenuhi kebutuhan analisis kebahasaan, seperti penyusunan tata bahasa, bahasa acuan penelitian, kamus, dan sebagainya.

KoIN telah memuat 5.488.035 kosakata yang berasal dari 2.250 berkas. Data tersebut bersumber dari teks ilmiah, yang mencakup artikel ilmiah pada jurnal nasional terakreditasi, serta skripsi dan tesis dari beberapa universitas terkemuka di Indonesia. Bidang ilmu yang terdapat dalam korpus ini adalah Ilmu Kesehatan, Ilmu Hayati, Ilmu Fisika, dan Ilmu Sosial. Pembagian bidang ilmu tersebut dilakukan berdasarkan Scopus, yang merupakan penyedia basis data terbesar untuk rujukan ilmiah. Hal ini menunjukkan bahwa ranah akademik, terutama dalam penulisan ilmiah menjadi hal yang urgen saat ini.

Menulis artikel ilmiah memiliki peran penting dalam kehidupan para akademisi dan peneliti. Di Indonesia, mahasiswa strata satu, magister dan doktor diwajibkan untuk menulis artikel jurnal akademik sebagai salah satu syarat

kelulusan. Peraturan tersebut dibuat sebagai hasil keputusan Direktorat Jenderal Pendidikan Tinggi (Dikti) No.152 / E / T / 2012 tentang persyaratan kelulusan untuk sarjana, master dan doktor (Dikti, 2016). Selain itu, Dikti telah mengeluarkan keputusan Permenristekdikti No. 20 tahun 2017 yang mewajibkan dosen dan profesor untuk menerbitkan artikel ilmiah. Artikel ilmiah yang diterbitkan di jurnal nasional maupun internasional diharapkan dapat meningkatkan kuantitas dan kualitas publikasi ilmiah di Indonesia.

Tulisan ilmiah termasuk dalam genre tulisan akademik dan memiliki struktur tersendiri. Jika membahas suatu genre dalam penulisan akademis, Hyland (2008: 550) berpendapat bahwa genre sebagai pengertian umum dari tata bahasa, kosakata, isi, dan sebagainya yang memungkinkan penulis untuk mengekspresikan nilai-nilai dan identitas yang berkaitan dengan disiplin ilmu tertentu. Salah satu karakteristik yang terkait dengan definisi genre tersebut adalah kosakata yang digunakan dalam artikel ilmiah.

Kesesuaian penggunaan kosakata pada penulisan ilmiah dalam suatu bidang ilmu perlu diperhatikan agar fokus dari tema yang dibahas mampu dipahami dan diterima oleh pembaca. Nation (2001) mengklasifikasikan kosakata ke dalam empat kategori; kosakata umum (*high frequency words*), kosakata akademik (*academic words*) kosakata khusus (*technical words*), dan kosakata frekuensi rendah (*low frequency words*). Kosakata yang dimaksud oleh Nation (2001) dalam temuannya merupakan kata-kata dalam Bahasa Inggris.

Kata-kata bahasa Inggris yang sering digunakan dalam komunikasi, bacaan, dan tulisan disebut sebagai kosakata berfrekuensi tinggi. Kosakata yang sering

digunakan ini telah dikembangkan bagi siswa untuk mempelajari kata-kata yang paling umum dalam Bahasa Inggris. Kosakata berfrekuensi tinggi yang paling banyak digunakan adalah *General Service List of English (GSL)* milik West (1953). GSL merupakan daftar kosakata umum yang paling banyak digunakan dalam teks (Nation, 2001). Dalam pembuatannya, West (1953) menggunakan lima juta korpus dan menerapkan beberapa kriteria dalam memilih kosakata tersebut, diantaranya berdasarkan frekuensi, kemudahan belajar, cakupan penggunaan kosakata tersebut dan gaya bahasa. GSL yang telah diproses berisi kurang lebih 2.000 keluarga kata (*word family*).

Seperti yang telah disebutkan sebelumnya, selain kosakata umum, terdapat kategori lain, yaitu kosakata akademik. Kosakata akademik pada umumnya merupakan kosakata yang sering digunakan dalam berbagai jenis teks akademik (Nation, 2001: 16). Daftar Kata Akademik atau dalam Bahasa Inggris disebut *Academic Word Lists (AWL)* merupakan daftar kosakata yang dibuat oleh Coxhead (2000) yang berisi 570 keluarga kata (*word family*). Kosakata dalam daftar ini merupakan representatif dari kosakata akademik karena cakupannya yang tinggi dalam teks akademik. Coxhead (2000) membuat daftar kosakata ini dari korpus yang berjumlah 3,5 juta kata yang dipilih dari jurnal akademik dan buku teks universitas dari empat bidang ilmu yaitu, seni; perdagangan, hukum, dan ilmu alam. AWL yang telah dibuat berisi 3.112 kata yang didalamnya tidak mengandung 2.000 kosakata yang telah masuk dalam GSL milik West (1953). Pada Bahasa Indonesia, daftar kosakata akademik sudah tersedia, yaitu Daftar Kosakata Akademik yang disusun oleh Rhubido (2017) yang terdiri dari 250 keluarga kata.

Kategori selanjutnya adalah kosakata khusus (*technical words*). Kosakata khusus terkait erat dengan topik atau disiplin ilmu suatu teks. Kosakata ini cukup umum digunakan dalam bidang suatu disiplin ilmu namun tidak ditemukan di disiplin ilmu lain. Kosakata khusus ini berbeda pada tiap disiplin ilmu dan hanya muncul sekitar 5% dari sebuah teks akademik (Nation 2001: 16). Seperti yang telah disebutkan sebelumnya, terdapat kategori lain yaitu kosakata berfrekuensi rendah. Untuk disiplin ilmu tertentu, koskata ini merupakan kosakata yang bukan merupakan kosakata berfrekuensi tinggi, bukan juga kosakata akademik, dan juga bukan merupakan kosakata khusus. Kosakata ini biasanya merupakan kosakata khusus dari bidang lain, kata benda, dan kosakata bahasa asing (Nation, 2001: 16).

Dalam artikel ilmiah, kategori kosakata yang sering ditemui adalah daftar kosakata akademik. Beberapa disiplin ilmu telah memiliki daftar kata akademik, seperti *Chemistry Academic Word List (CAWL)* atau Daftar Kosakata Akademik Kimia yang dibuat oleh Valipouri dan Nassaji (2013). Daftar kosakata ini dibuat dari korpus sebanyak empat juta kata yang berasal dari artikel penelitian di bidang kimia atau yang disebut juga *Chemistry Research Article (CRAC)*. Mereka mengembangkan CAWL untuk mengetahui kata-kata yang paling sering digunakan dalam bidang kimia.

Meskipun daftar kosakata yang mencakup disiplin ilmu tertentu telah dibuat, beberapa kosakata terkadang saling tumpang tindih, seperti *Nursing Academic Word List (NAWL)* dan *Medical Academic Word List (MAWL)* milik Wang, et al., (2008). Yang (2015) melakukan sebuah penelitian dengan membuat korpus dari artikel penelitian ilmu keperawatan atau *Nursing Research Articles*

Corpus (NRAC) yang kemudian membandingkan daftar kosakata tersebut dengan AWL, NAWL, dan MAWL untuk melihat mana yang paling baik dalam mencakup kosakata dalam NRAC.

Berdasarkan penelitian tersebut, muncul perbandingan 676 kosakata di NAWL dan 623 keluarga kata di MAWL yang menunjukkan kejadian tumpang tindih antara kosakata di NAWL dan MAWL adalah sebanyak 429 kata (63,46%). Sementara itu, perbandingan antara 676 kosakata dalam NAWL dan 570 kosakata dalam AWL menunjukkan kejadian tumpang tindih sebanyak 378 kosakata (55,92%), sedangkan hanya sekitar 342 kosakata (54,90%) dari 623 kosakata di MAWL yang tumpang tindih dengan 570 kosakata di AWL (Wang, et al., 2008: 451). Dari hasil yang muncul dapat dilihat bahwa terdapat banyak kemiripan kosakata antara dua disiplin ilmu yaitu keperawatan (NRAC) dan kedokteran (MAWL).

Meskipun ilmu keperawatan dan kedokteran termasuk dalam ilmu kesehatan, keduanya tetap memiliki perbedaan dalam praktik penelitian yang menyebabkan terdapat perbedaan dalam penggunaan kosakata. Menurut Hyland dan Tse (2007), bermacam-macam praktik dan metode dalam penelitian dapat dikaitkan dengan penyampaian argumen dan penggunaan kata.

Karena munculnya tumpang tindih dan cakupan yang lebih rendah oleh AWL dalam disiplin ilmu tertentu, Kwary & Artha (2017) membuat Daftar Kata Artikel Akademik atau *Academic Article Word List* (AAWL) untuk Ilmu Sosial. Mereka merumuskan *Academic Article Word List-Social Science* (AAWL-SS) berdasarkan korpus 1.040.259 kata dari artikel ilmu sosial di Elsevier. Setelah

membuat AAWL-SS, mereka menguji AAWL-SS dengan membandingkannya dengan tiga daftar kosakata lain, yaitu AWL, *New Academic Word List* (NAWL), dan *British National Corpus-Corpus of Contemporary American English* (BNC-COCA3) dengan 122 artikel jurnal dalam bidang ilmu sosial. Hasil penelitian tersebut menunjukkan bahwa cakupan kata pada AAWL-SS lebih tinggi daripada AWL, NAWL, dan BNC-COCA3.

Dari penjelasan tersebut bisa dikatakan bahwa penggunaan kosakata perlu disesuaikan dengan tema atau konteks yang sedang dibahas. Penggunaan kosakata atau diksi mampu menunjukkan fokus tema penelitian sehingga pembaca mampu memahami isi dari artikel penelitian tersebut. Penggunaan dan pemilihan kosakata dalam penulisan ilmiah memiliki peran penting. Selain penting dalam penulisan ilmiah, pengetahuan akan kosakata akademik dalam Bahasa Indonesia juga penting untuk para pelajar asing. Beberapa universitas di Indonesia menjalin kerjasama dengan universitas luar negeri untuk melakukan pertukaran pelajar. Selain itu, pada Siaran Pers No:41/SP/HM/BKKP/V/2017 disebutkan bahwa jumlah permohonan Izin Belajar bagi Mahasiswa Asing semakin meningkat, yaitu 150 – 500 permohonan setiap minggunya. Jumlah ini menunjukkan meningkatnya minat para pelajar asing untuk menempuh studi di Indonesia.

Salah satu kebutuhan penting bagi pelajar asing yang akan menempuh studi di Indonesia adalah kemampuan berbahasa. Hal ini tentunya berkaitan erat dengan pengetahuan kosakata yang dimiliki para pelajar asing agar mampu berkomunikasi dengan baik. Status sebagai pelajar asing menuntut para pelajar untuk menguasai Bahasa Indonesia. Salah satu program pemerintah Indonesia

dalam memfasilitasi para pelajar asing untuk menguasai Bahasa Indonesia, yaitu Bahasa Indonesia untuk Penutur Asing (BIPA). BIPA merupakan salah satu program pemerintah yang mengajarkan Bahasa Indonesia bagi penutur asing, salah satunya adalah pelajar asing. Materi yang diberikan dalam program BIPA yaitu mengenai Bahasa Indonesia dalam kehidupan sehari-hari dan pengenalan budaya Indonesia.

Untuk efisiensi pembelajaran bahasa, adanya daftar kosakata dapat membantu para penutur asing dalam menambah perbendaharaan kata. Seperti yang telah disebutkan pada beberapa paragraf sebelumnya, dalam Bahasa Indonesia terdapat *Indonesian High Frequency Word List* (Kwary, 2013) yang dapat digunakan untuk menambah wawasan tentang kosakata yang umum digunakan dalam Bahasa Indonesia. Bagi pelajar asing yang memiliki tujuan untuk belajar di Indonesia, pengenalan kosakata-kosakata akademik juga diperlukan atau dapat disisipkan dalam pembelajaran Bahasa Indonesia sehingga pengetahuan kosakata para pelajar asing terkait bidang ilmu yang sedang dipelajari dapat bertambah.

Pada paragraf sebelumnya telah disebutkan bahwa pada Bahasa Indonesia telah tersedia Daftar Kosakata Akademik (DKA) yang dibuat oleh Rhubido (2017), namun belum tersedia daftar kosakata yang ditujukan untuk bidang ilmu tertentu. Oleh sebab itu, pada penelitian ini, penulis membuat daftar kosakata yang dapat mewakili bidang ilmu tertentu yang akan dinamai Daftar Kosakata Artikel Akademik (DKAA) yang dapat digunakan sebagai rujukan para penutur asing yang ingin menambah pengetahuan atau mempelajari Bahasa Indonesia pada ranah akademik secara spesifik.

Pada pembuatan daftar kosakata, penulis menggunakan artikel jurnal akademik dalam Bahasa Indonesia yang sekaligus sebagai objek penelitian. Jurnal yang akan digunakan dapat dilihat daftarnya di laman SINTA di <http://sinta2.ristekdikti.go.id/>. Di SINTA, skor yang diberikan bervariasi yaitu mulai dari SINTA1 (S1) hingga SINTA6 (S6). Skor ini digunakan untuk menentukan kualitas akreditasi jurnal.

Di antara enam skor yang terdapat pada SINTA, S1 dan S2 merupakan yang tertinggi. Jurnal yang terdapat dalam kategori S1 mendapat akreditasi “A” (diindeks oleh Scopus) yang hampir seluruh artikel di dalamnya menggunakan Bahasa Inggris. Sementara itu, jurnal yang terdapat dalam kategori S2 merupakan jurnal terindeks yang mendapat akreditasi “B”. Beberapa jurnal artikel yang terdapat dalam kategori SINTA2 (S2) merupakan jurnal yang menggunakan Bahasa Indonesia dan Bahasa Inggris. Oleh sebab itu, dalam penelitian ini penulis memilih artikel jurnal yang terdapat dalam SINTA2 (S2) sebagai objek penelitian yang kemudian dikelompokkan menurut tiap bidang ilmu berdasarkan pembagian ilmu pada Garuda (Garba Rujukan Digital) yang serupa dengan pembagian ilmu pada Scopus, yaitu bidang Ilmu Sosial, Ilmu Kesehatan, Ilmu Fisika, dan Ilmu Hayati.

Pemilihan bidang empat bidang ilmu tersebut mengacu pada Daftar Kosakata Akademik (Rhubido, 2017). Dalam pembuatan daftar kosakata akademik, Rhubido (2017) melibatkan artikel akademik dari keempat bidang ilmu tersebut. Salah satu tujuan dari penelitian yang dilakukan oleh Rhubido adalah membuat daftar kosakata yang dimaksudkan dapat mewakili seluruh bidang ilmu yang menggunakan Bahasa Indonesia. Dari celah tersebut, penulis kemudian membuat

daftar kosakata akademik dengan pemilihan bidang ilmu yang sama namun diperuntukkan bagi tiap bidang ilmu yang diharapkan mampu merepresentasi kosakata dengan lebih baik dan efektif sesuai yang memenuhi kriteria yang disebutkan oleh Nation (2001), yaitu sebesar 9%.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian yang telah dijelaskan oleh penulis, pentingnya pembuatan daftar kosakata akademik pada lintas bidang ilmu perlu dilakukan. Selain belum tersedianya daftar kosakata akademik Bahasa Indonesia yang dikhususkan untuk bidang ilmu tertentu, daftar kosakata ini juga diperlukan sebagai rujukan untuk menulis ilmiah, dan rujukan bagi para penutur asing untuk mempelajari Bahasa Indonesia dalam ranah akademik khususnya bidang ilmu yang lebih spesifik. Pembuatan daftar kosakata lintas bidang ilmu dilakukan pada empat bidang ilmu berdasarkan pembagian ilmu dari Garuda yang mengacu pada Scopus, yaitu pada bidang Ilmu Sosial, Ilmu Kesehatan, Ilmu Fisika, dan Ilmu Hayati. Dari beberapa fenomena tersebut, penulis mengajukan tiga rumusan masalah sebagai berikut:

- a. Kosakata akademik apa saja yang terdapat dalam artikel akademik bidang Ilmu Sosial, Ilmu Kesehatan, Ilmu Fisika, dan Ilmu Hayati?
- b. Bagaimana perbandingan cakupan DKAA pada artikel akademik bidang Ilmu Sosial, Ilmu Kesehatan, Ilmu Fisika, dan Ilmu Hayati?
- c. Bagaimana distribusi kelas kata DKAA pada artikel akademik bidang Ilmu Sosial, Ilmu Kesehatan, Ilmu Fisika dan Ilmu Hayati?

1.3 Tujuan Penelitian

Rumusan masalah yang diajukan oleh penulis memiliki beberapa tujuan.

Berikut merupakan tujuan dari penelitian ini :

- a. Untuk mengidentifikasi kosakata yang terdapat pada bidang Ilmu Sosial, Ilmu Kesehatan, Ilmu Fisika, dan Ilmu Hayati. Banyaknya kosakata akademik yang terdapat pada lintas bidang ilmu perlu dibuat agar penutur asing dapat mempelajari kosakata apa saja yang sering digunakan pada tiap bidang ilmu.
- b. Untuk membandingkan cakupan daftar kosakata yang telah dibuat serta mengetahui penggunaan kosakata tersebut pada artikel ilmiah. Cakupan yang terdapat pada daftar kosakata akademik menunjukkan keefektifan dari daftar kosakata yang telah dibuat. Dari hasil cakupan tersebut dapat diketahui apakah daftar kosakata akademik yang dibuat dapat mewakili kata-kata yang sering digunakan pada tiap bidang ilmu.
- c. Untuk menganalisis distribusi kelas kata daftar kosakata yang telah dibuat pada tiap bidang ilmu. Klasifikasi kosakata akan memudahkan para penutur asing untuk mempelajari jenis kelas kata apa yang perlu dipelajari lebih banyak agar dapat membantu mereka dalam mempelajari Bahasa Indonesia secara efektif.

1.4 Manfaat Penelitian

Penelitian ini merupakan kajian korpus linguistik yang berkaitan dengan daftar kosakata pada artikel ilmiah di Indonesia yang terbagai kedalam empat disiplin ilmu yaitu, Ilmu Sosial, Ilmu Kesehatan, Ilmu Fisika, dan Ilmu Hayati.

Penelitian ini diharapkan bermanfaat bagi pembaca baik secara teoritis dan praktis. Secara teoretis, penelitian ini dapat memperluas kajian dalam pengembangan linguistik korpus, terutama dalam hal kosakata akademik pada bidang ilmu yang lebih spesifik. Penelitian ini juga diharapkan dapat bermanfaat bagi para peneliti dan ahli bahasa yang tertarik dengan topik ini sehingga dapat melakukan penelitian selanjutnya. Terlebih lagi, saat ini pemerintah Indonesia sedang mengembangkan korpus Bahasa Indonesia sebagai kebutuhan penelitian kebahasaan.

Secara praktis, penelitian ini diharapkan dapat memberikan wawasan tentang kosakata akademik dalam penulisan artikel ilmiah. Sebagai bahasa standar yang digunakan dalam lingkup akademik, belum tersedianya kosakata akademik pada bidang ilmu tertentu tidak jarang membuat sebagian penulis mengalami kesulitan dalam menulis artikel yang sesuai dengan bidangnya. Adanya daftar kosakata ini dapat digunakan oleh para penulis sebagai referensi dalam menulis artikel ilmiah yang sesuai dengan bidangnya. Selain itu, daftar kosakata akademik yang dibuat pada penelitian dapat menjadi salah satu rujukan bagi para penutur asing untuk mempelajari kosakata akademik pada bidang ilmu yang lebih spesifik. Dalam pengajaran, para pengajar dapat menggunakan daftar kosakata ini untuk mengajar Bahasa Indonesia kepada mahasiswa dalam kegiatan penulisan secara akademik sesuai dengan disiplin ilmu para mahasiswa serta dapat digunakan sebagai referensi dalam pengajaran Bahasa Indonesia untuk Penutur Asing (BIPA).

1.5 Daftar Istilah

Berikut ini merupakan beberapa daftar istilah yang sering muncul pada penelitian ini. Berikut rincian daftar istilah tersebut beserta penjelasannya :

1. **Linguistik Korpus** : Studi tentang bahasa yang digunakan melalui korpora yang dikumpulkan untuk berbagai tujuan penelitian bahasa.
2. **Keluarga Kata** : Kata dasar yang memiliki kata turunan (Nation, 2001). Kata turunan tersebut dapat berupa penambahan awalan dan akhiran yang berbeda. Misalnya, kata dasar “kerja” yang memiliki kata turunan seperti “pekerja”, “bekerja”, atau “mengerjakan”.
3. **Frekuensi** : Jumlah kekerapan suatu kosakata atau seberapa sering suatu kosakata muncul pada teks.
4. **Range** : Jumlah jangkauan atau seberapa banyak teks yang memuat kosakata tertentu.
5. **Token** : Jumlah keseluruhan kata dalam teks.
6. **TTR** : *Type-Token Ratio* merupakan ukuran variasi kosakata atau keragaman leksikal dalam sebuah teks atau ujaran seseorang
7. **Daftar Kosakata Artikel Akademik (DKAA)** : Nama daftar kosakata akademik yang dibuat pada penelitian ini yang kemudian terdapat keterangan pada dua kata terakhir untuk menunjukkan spesifikasi bidang terkait. Misalnya: Daftar Kosakata Artikel Akademik Ilmu Sosial (DKAA-IS), dan Daftar Kosakata Artikel Akademik Ilmu Hayati (DKAA-IH).