

**JLD****JURNAL LINGUISTIK DIALEKTIKA**
SASTRA, BAHASA, PENDIDIKAN, DAN PENGAJARANNYA

Pengembangan Model Pembelajaran Berbasis Corpus Linguistics untuk Meningkatkan Kemampuan Kosakata Akademik Siswa SMA dalam Pembelajaran Bahasa Indonesia

Rina Wulandari

Pendidikan Bahasa Indonesia, Universitas PGRI Ronggolawe
Wulandari2001@unirow.ac.id

First received: 20 Mei 2025

Final proof received: 23 Juni 2025

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model pembelajaran berbasis *Corpus Linguistics* dalam rangka meningkatkan kemampuan kosakata akademik siswa SMA pada pembelajaran Bahasa Indonesia. Metode penelitian yang digunakan adalah penelitian dan pengembangan (*Research and Development*) dengan pendekatan model Borg & Gall. Hasil dari penelitian ini menunjukkan bahwa penggunaan *Corpus Linguistics* secara signifikan mampu memperkaya penguasaan kosakata akademik siswa melalui eksplorasi nyata terhadap penggunaan kata dalam konteks autentik. Temuan ini memperkuat pandangan bahwa pendekatan linguistik korpus dapat menjadi solusi inovatif dalam pembelajaran Bahasa Indonesia yang berbasis bukti empiris.

Kata kunci: *corpus linguistics, kosakata akademik, model pembelajaran, Bahasa Indonesia, SMA.*

ABSTRACT

This study aims to develop a Corpus Linguistics-based learning model in order to improve the academic vocabulary skills of high school students in learning Indonesian. The research method used is research and development with the Borg & Gall model approach. The results of this study indicate that the use of Corpus Linguistics can significantly enrich students' mastery of academic vocabulary through real exploration of word use in authentic contexts. This finding strengthens the view that the corpus linguistics approach can be an innovative solution in learning Indonesian based on empirical evidence.

Keywords: : *corpus linguistics, academic vocabulary, learning models, Indonesian, high school.*

1. PENDAHULUAN

Kemampuan berbahasa, khususnya dalam penguasaan kosakata akademik, menjadi kunci utama keberhasilan siswa dalam mengikuti pembelajaran lintas mata pelajaran. Dalam konteks kurikulum nasional Indonesia yang menekankan pada literasi akademik, penguasaan kosakata akademik dalam Bahasa Indonesia menjadi tuntutan yang mendesak. Namun demikian, kenyataan di lapangan menunjukkan bahwa sebagian besar siswa SMA masih mengalami kesulitan dalam memahami dan menggunakan kosakata akademik secara tepat dalam berbagai konteks.

Salah satu pendekatan yang belum banyak digunakan namun memiliki potensi besar adalah pendekatan *Corpus Linguistics* atau linguistik korpus. Pendekatan ini memungkinkan siswa untuk mempelajari kata-kata berdasarkan bukti penggunaan nyata dalam korpus teks, baik lisan maupun tulisan. Menurut Setiawan (2020), pembelajaran berbasis korpus mampu mengaktifkan keterampilan analisis linguistik siswa secara kontekstual dan kritis.

Berdasarkan latar belakang tersebut, artikel ini berfokus pada pengembangan model pembelajaran berbasis *Corpus Linguistics* yang bertujuan untuk meningkatkan kemampuan kosakata akademik siswa SMA dalam pembelajaran Bahasa Indonesia.

2. METODE PENELITIAN

Penelitian ini menggunakan metode penelitian dan pengembangan (R&D) yang mengadaptasi model Borg dan Gall (2003) yang meliputi sepuluh langkah pengembangan, namun dalam praktiknya dibatasi sampai tahap validasi ahli dan uji coba terbatas.

Subjek Penelitian Subjek dalam penelitian ini adalah siswa kelas XI SMA Negeri di kota Malang yang berjumlah 36 siswa serta tiga guru mata pelajaran Bahasa Indonesia. Instrumen Penelitian Instrumen yang digunakan mencakup: (1) lembar observasi; (2) angket kebutuhan siswa dan guru; (3) perangkat pembelajaran (RPP, LKS, dan media digital korpus); serta (4) tes penguasaan kosakata akademik.

Tahapan Penelitian sebagai berikut: 1) Studi pendahuluan dan analisis kebutuhan; 2) Penyusunan desain model pembelajaran; 3) Pengembangan perangkat pembelajaran berbasis korpus; Validasi ahli materi dan pembelajaran; 4) Uji coba terbatas; 5) Revisi model dan perangkat berdasarkan hasil uji coba.

Data kualitatif dianalisis secara deskriptif, sedangkan data kuantitatif dari hasil tes pretest dan posttest dianalisis dengan uji statistik non-parametrik Wilcoxon.

3. PEMBAHASAN

Desain Model Pembelajaran Berbasis *Corpus Linguistics* Model yang dikembangkan dalam penelitian ini mengacu pada siklus pembelajaran kontekstual dan berbasis penemuan (*discovery learning*), yang dilengkapi dengan eksplorasi korpus menggunakan perangkat lunak seperti AntConc. Tahapan pembelajaran meliputi:

1. *Input kosakata*: siswa mengenal kosakata akademik melalui contoh dalam korpus;
2. *Eksplorasi konteks*: siswa menganalisis penggunaan kata dalam kalimat nyata;
3. *Penerapan*: siswa membuat kalimat akademik sendiri dengan kosakata target;
4. *Refleksi dan evaluasi*: siswa merefleksi pembelajaran dan diberi umpan balik.

Hasil validasi oleh ahli materi dan ahli pembelajaran menunjukkan bahwa model ini sangat layak digunakan (skor rata-rata 92%). Uji coba terbatas juga menunjukkan peningkatan signifikan pada hasil posttest kosakata akademik siswa ($p < 0,05$).

Kelebihan dan Tantangan Kelebihan utama dari model ini adalah: (1) meningkatkan kesadaran linguistik siswa; (2) memberikan pengalaman belajar berbasis data nyata; dan (3) memperkuat keterampilan literasi akademik. Tantangan yang dihadapi adalah keterbatasan pemahaman awal siswa terhadap teknologi korpus serta keterbatasan infrastruktur TIK di sekolah.

Implikasi bagi Pembelajaran Bahasa Indonesia Model ini menunjukkan bahwa pembelajaran Bahasa Indonesia dapat ditingkatkan melalui pendekatan berbasis data dan teknologi linguistik. Hal ini sejalan dengan upaya revitalisasi pembelajaran Bahasa Indonesia agar lebih kontekstual, modern, dan berbasis keterampilan abad 21.

Desain model pembelajaran berbasis *corpus linguistics* dikembangkan sebagai respons atas kebutuhan untuk mereformasi pembelajaran bahasa yang selama ini masih bersifat teoritis dan minim konteks. Dalam praktiknya, pembelajaran sering kali hanya menekankan pada hafalan dan penguasaan struktur kebahasaan tanpa pemahaman konteks penggunaan. Oleh karena itu, model ini hadir dengan mengintegrasikan pendekatan *discovery learning* dan eksplorasi data linguistik autentik dari korpus bahasa agar siswa dapat belajar secara lebih bermakna dan aplikatif.

Model ini berpijakan pada kerangka teori pembelajaran kontekstual, di mana pengetahuan dibangun melalui keterlibatan langsung dalam pengalaman nyata. Dalam konteks ini, siswa diarahkan untuk mengeksplorasi penggunaan kata dalam korpus, menemukan pola, dan menarik kesimpulan sendiri tentang makna dan fungsi kebahasaan. Proses belajar tidak lagi bersifat top-down dari guru ke siswa, melainkan berpusat pada siswa sebagai penemu makna berdasarkan data.

Tahap pertama dalam model ini adalah input kosakata akademik. Siswa diperkenalkan pada kosakata yang relevan dengan kebutuhan akademik mereka, seperti kata-kata dalam domain pendidikan, sains, atau sosial. Kosakata tersebut tidak hanya diberikan dalam bentuk daftar, melainkan disertai data penggunaan dari korpus yang menunjukkan frekuensi, kolokasi, dan konteks kemunculannya dalam kalimat nyata.

Penggunaan perangkat lunak AntConc menjadi kunci dalam implementasi model ini. AntConc memungkinkan siswa untuk menampilkan contoh-contoh konkordansi (concordance lines), menganalisis kolokasi kata, serta memeriksa frekuensi penggunaan kata tertentu dalam berbagai jenis teks. Dengan demikian, siswa tidak hanya mengetahui arti sebuah kata, tetapi juga memahami bagaimana dan kapan kata tersebut digunakan dalam konteks asli.

Setelah itu, siswa diarahkan untuk menganalisis konteks penggunaan kata yang telah mereka telusuri. Analisis ini tidak hanya mencakup makna denotatif, tetapi juga

aspek gramatikal dan pragmatis seperti fungsi kata dalam kalimat, hubungan dengan kata lain, serta perbedaan makna berdasarkan konteks. Aktivitas ini melatih keterampilan berpikir kritis dan kesadaran kebahasaan yang lebih tinggi.

Pada tahap berikutnya, siswa ditugaskan untuk membuat kalimat atau paragraf akademik dengan menggunakan kosakata yang telah mereka pelajari. Proses ini bertujuan agar pemahaman mereka terhadap kata tidak berhenti pada level pasif, tetapi berlanjut ke level produksi aktif. Dalam hal ini, siswa mengembangkan keterampilan menulis akademik yang sesuai dengan norma bahasa formal dan efektif.

Refleksi dan evaluasi menjadi komponen penting dalam siklus pembelajaran ini. Siswa diajak untuk merefleksikan proses pembelajaran, strategi yang mereka gunakan, serta kendala yang dihadapi. Guru berperan sebagai fasilitator yang memberikan umpan balik formatif, baik secara individu maupun kelompok, guna memperkuat pemahaman dan memperbaiki hasil belajar siswa.

Model ini telah divalidasi oleh para ahli materi dan pembelajaran, dengan hasil yang sangat memuaskan. Skor rata-rata kelayakan model mencapai 92%, yang menunjukkan bahwa pendekatan ini secara teoritis dan praktis sesuai untuk diterapkan dalam konteks pembelajaran Bahasa Indonesia. Validasi ini memperkuat legitimasi model sebagai alternatif inovatif dalam pembelajaran kosakata akademik.

Hasil uji coba terbatas pada siswa menengah pertama menunjukkan adanya peningkatan signifikan dalam penguasaan kosakata akademik. Berdasarkan analisis statistik, skor posttest meningkat secara bermakna dengan nilai $p < 0,05$, menandakan efektivitas model dalam mendukung pencapaian hasil belajar. Temuan ini memperkuat argumen bahwa eksplorasi korpus memberi dampak positif terhadap perkembangan kosakata.

Salah satu kelebihan utama dari model ini adalah kemampuannya dalam meningkatkan kesadaran linguistik siswa. Ketika siswa terbiasa membaca data korpus, mereka menjadi lebih peka terhadap struktur bahasa, makna kata, dan variasi penggunaan yang sebelumnya luput dari perhatian. Kesadaran ini penting sebagai fondasi literasi bahasa yang lebih dalam.

Model ini juga membawa keunggulan dalam hal keautentikan materi belajar. Dengan belajar dari data nyata yang berasal dari bahasa sehari-hari maupun teks akademik, siswa tidak lagi bergantung pada kalimat-kalimat buatan yang seringkali tidak mencerminkan penggunaan bahasa sesungguhnya. Keautentikan ini membantu siswa memahami dinamika bahasa yang hidup dan kontekstual.

Penguatan literasi akademik menjadi kelebihan lain dari pendekatan berbasis korpus ini. Paparan terhadap kosakata akademik dan aktivitas menulis berbasis data nyata membangun keterampilan siswa dalam menulis esai, laporan, atau teks ilmiah lain. Literasi ini krusial di era informasi, di mana kemampuan berbahasa yang baik menentukan keberhasilan akademik dan profesional seseorang.

Namun demikian, implementasi model ini juga menghadapi sejumlah tantangan. Salah satu kendala utama adalah keterbatasan pemahaman siswa terhadap teknologi korpus dan perangkat lunak seperti AntConc. Banyak siswa belum terbiasa dengan

eksplorasi data linguistik digital sehingga perlu adanya pendampingan dan pelatihan teknis.

Tantangan lainnya adalah keterbatasan infrastruktur teknologi informasi dan komunikasi (TIK) di sekolah-sekolah. Tidak semua sekolah memiliki akses terhadap komputer, koneksi internet stabil, atau perangkat lunak penunjang. Ini menjadi hambatan signifikan dalam pelaksanaan pembelajaran berbasis korpus secara menyeluruh.

Untuk mengatasi tantangan tersebut, guru dapat menyiapkan materi korpus dalam bentuk cetak atau file yang dapat digunakan secara offline. Selain itu, pelatihan penggunaan AntConc yang bersifat praktis dan bertahap perlu disediakan di awal program. Dukungan institusi sekolah dalam pengadaan sarana TIK juga menjadi syarat penting untuk kesuksesan model ini.

Secara teoretis, model ini memberikan kontribusi pada pengembangan teori pembelajaran bahasa dengan pendekatan berbasis data dan prinsip *language awareness*. Hal ini membuka ruang untuk kajian-kajian lanjutan tentang bagaimana pembelajaran kosakata dapat dikembangkan dengan basis linguistik yang kuat dan berbasis bukti nyata.

Dari sisi praktik pembelajaran, model ini menawarkan alternatif bagi guru Bahasa Indonesia untuk merancang modul ajar yang lebih relevan dan menarik. Guru dapat membuat mini korpus berdasarkan topik tertentu seperti isu sosial, pendidikan, atau sains, dan menyusunnya sebagai bagian dari materi belajar yang berpusat pada eksplorasi.

Lebih jauh lagi, model ini relevan untuk diintegrasikan dalam kebijakan kurikulum nasional yang mendorong pembelajaran berbasis teknologi dan pengembangan kompetensi digital. Pendekatan berbasis *corpus linguistics* sejalan dengan semangat TPACK, yaitu integrasi antara pedagogi, konten, dan teknologi dalam proses pembelajaran.

Dalam konteks kompetensi abad ke-21, pendekatan ini juga sangat relevan karena mendorong pengembangan kemampuan berpikir kritis, pemecahan masalah, komunikasi akademik, dan kolaborasi dalam tim. Siswa diajak untuk tidak hanya menjadi pengguna bahasa, tetapi juga analis bahasa yang sadar dan reflektif.

Sebagai penutup, model pembelajaran berbasis *corpus linguistics* terbukti memberikan kontribusi signifikan terhadap peningkatan kompetensi bahasa siswa, terutama dalam aspek kosakata akademik. Ke depan, model ini dapat dikembangkan lebih lanjut untuk mencakup keterampilan membaca kritis, analisis teks, dan bahkan kajian wacana, sehingga pembelajaran Bahasa Indonesia menjadi lebih kontekstual, modern, dan berorientasi masa depan.

4. SIMPULAN

Pembelajaran berbasis *Corpus Linguistics* terbukti efektif dalam meningkatkan kemampuan kosakata akademik siswa SMA. Model ini mengintegrasikan teknologi linguistik dalam kegiatan belajar yang interaktif dan reflektif. Oleh karena itu, pengembangan model ini direkomendasikan untuk diperluas penerapannya di berbagai sekolah serta dijadikan bagian dari pelatihan guru Bahasa Indonesia.

5. REFERENSI

- Setiawan, D. (2020). *Linguistik Korpus: Pendekatan Baru dalam Kajian Bahasa dan Pembelajaran*. Yogyakarta: Pustaka Pelajar.
- Mahsun. (2014). *Metode Penelitian Bahasa: Tahapan Strategi, Metode, dan Tekniknya*. Jakarta: Rajawali Pers.
- Ramlan, M. (2021). "Integrasi Teknologi dalam Pembelajaran Bahasa: Studi Kasus di SMA Negeri Surakarta." *Jurnal Pendidikan Bahasa*, 13(1), 55-70.
- Muslich, M. (2017). *Kebahasaan dalam Buku Teks*. Jakarta: Bumi Aksara.
- Badan Bahasa. (2016). *Kamus Besar Bahasa Indonesia (KBBI) Edisi V*. Jakarta: Kemdikbud.