



Transformasi Pendidikan Abad XXI *Sebuah Bunga Rampai*

Dwi Sulisworo, Meita Fitriyanawati, Arsyad Cahya Subrata,
Ikmi Nur Oktavianti, Icuk Prayogi, Ika Maryani, Yosi Wulandari,
Syariful Fahmi, Soffi Widyanesti Priwantoro, Hardi Santosa,
Vera Yuli Erviana, Yudhiakto Pramudya,

Siti Partini Suardiman, Nur Fatimah, Achadi Budi Santosa,
Ratri Nur Hidayati, Pratomo Widodo, Agus Widyantoro, Gunadi,
Sudaryanto, Marizta Syahda Tiara Yahya, Diah Agustin Ari Priyadi,
Ani Susanti, Fariz Setyawan, Arilia Triyoga, Suyatno, Dewi Eko Wati,
Sri Katoningsih, Febritesna Nuraini, Caraka Putra Bhakti, Sunarti,
Anggit Prabowo, Tatang Herman, Siti Fatimah, Dian Hidayati,
Priska Fadhila, Purwati Zisca Diana, Roni Sulistiyono,
Trisna Sukmayadi, Siti Salamah, Rendra Ananta Prima Hardiyanta,
Wachid Eko Purwanto.

Transformasi Pendidikan Abad XXI Sebuah Bunga Rampai

Dwi Sulisworo, Meita Fitriawanati, Arsyad Cahya Subrata,
Ikmi Nur Oktavianti, Icuk Prayogi, Ika Maryani,
Yosi Wulandari, Syariful Fahmi, Soffi Widyanesti Priwantoro,
Hardi Santosa, Vera Yuli Erviana, Yudhiakto Pramudya,
Siti Partini Suardiman, Nur Fatimah, Achadi Budi Santosa,
Ratri Nur Hidayati, Pratomo Widodo, Agus Widyantoro,
Gunadi, Sudaryanto, Marizta Syahda Tiara Yahya,
Diah Agustin Ari Priyadi, Ani Susanti, Fariz Setyawan,
Arilia Triyoga, Suyatno, Dewi Eko Wati, Sri Katoningsih,
Febritesna Nuraini, Caraka Putra Bhakti, Sunarti,
Anggit Prabowo, Tatang Herman, Siti Fatimah,
Dian Hidayati, Priska Fadhila, Purwati Zisca Diana,
Roni Sulistiyono, Trisna Sukmayadi, Siti Salamah,
Rendra Ananta Prima Hardiyanta, Wachid Eko Purwanto.



Penerbit K-Media
Yogyakarta, 2024

Transformasi Pendidikan Abad XXI : Sebuah Bunga Rampai

Tim Penulis:

Dwi Sulisworo, Meita Fitriawanawati, Arsyad Cahya Subrata, Ikmi Nur Oktavianti, Icuk Prayogi,...[dan 37 lainnya]

ISBN: 978-623-174-334-3

Tim Reviewer:

Sudaryanto, M.Pd., Dr. Suyatno, M.Pd.I., Dr. Ani Susanti, M.Pd.B.I., Dr. Wahyu Nanda Eka Saputra, M.Pd., Kons., Dr. Vera Yuli Erviana, M.Pd.,...[dan 4 lainnya]

Tata Letak: Setia S Putra

Desain Sampul: Setia S Putra

Diterbitkan oleh:



Penerbit K-Media

Anggota IKAPI No.106/DIY/2018

Banguntapan, Bantul, Yogyakarta.

WA +6281-802-556-554, Email: kmedia.cv@gmail.com

Cetakan pertama, Januari 2024

Yogyakarta, Penerbit K-Media 2024

15,5 x 23 cm, xii, 494 hlm.

Hak Cipta dilindungi Undang-Undang

All rights reserved

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apapun tanpa izin tertulis dari Penulis dan Penerbit.

Isi di luar tanggung jawab percetakan

DAFTAR ISI

PRAKATA TIM <i>REVIEWER</i>	iii
KATA PENGANTAR.....	vii
DAFTAR ISI	ix

BAB 1 TEKNOLOGI DIGITAL DALAM MULTIPERSPEKTIF PENDIDIKAN

- **Transformasi Pendidikan Melalui Kecerdasan Buatan: Isu-isu dan Implikasi**
Dwi Sulisworo, Meita Fitrianawati, Arsyad Cahya Subrata..... 2
- **Korpus Pembelajar sebagai *Big Data* dalam Pengajaran Bahasa Inggris di Era Masyarakat 5.0**
Ikmi Nur Oktavianti, Icku Prayogi.....27
- **Tren, Peluang, dan Tantangan STEM di Sekolah Dasar Indonesia**
Ika Maryani54
- **Media Pembelajaran Inovatif Berbasis Teknologi: Mewujudkan Pendidik Inspiratif yang Tanggap Budaya**
Yosi Wulandari.....77
- **Multimedia Pembelajaran Interaktif**
Syariful Fahmi, Soffi Widyanesti Priwantoro94

Korpus Pembelajar Sebagai *Big Data* dalam Pengajaran Bahasa Inggris Di Era Masyarakat 5.0

Ikmi Nur Oktavianti

Magister Pendidikan Bahasa Inggris, Fakultas Keguruan dan Ilmu
Pendidikan, Universitas Ahmad Dahlan
ikmi.oktavianti@pbi.uad.ac.id

Icuk Prayogi

Pendidikan Bahasa dan Sastra Indonesia, Fakultas Pendidikan Bahasa
dan Sastra, Universitas PGRI Semarang
icukprayogi@upgris.ac.id

Pendahuluan

Bahasa Inggris di Indonesia masih berstatus sebagai bahasa asing atau *English as a foreign language* (EFL) (Lauder, 2008). Kondisi tersebut menjadikan pengajaran bahasa Inggris di Indonesia sebagai sebuah tantangan tersendiri. Dengan status tersebut, eksposur terhadap bahasa Inggris sebagai bahasa target menjadi terbatas, dominan mengandalkan materi dan interaksi di dalam kelas. Hal ini menjadikan pengajaran bahasa Inggris di Indonesia mempunyai keterbatasan. Maka tidak jarang dijumpai beragam kesulitan dalam mempelajari bahasa Inggris yang tercermin dalam kurangnya motivasi (Hasanuddin & Ciptaningrum, 2021; Nabila et al., 2021) dan kesalahan-kesalahan berbahasa yang diproduksi oleh para pembelajar di Indonesia (Harta et al., 2021; Setiarini, 2018).

Agar dapat mengetahui kemahiran pembelajar, pengajar lazimnya menyusun dan menyelenggarakan tes, termasuk dalam pengajaran bahasa Inggris. Guru bahasa

Inggris menyusun soal tes untuk diujikan dan memperoleh skor sebagai refleksi hasil belajar dan pengukuran kemampuan peserta didik terhadap materi yang diajarkan. Dalam penguasaan keterampilan berbahasa, hanya mengandalkan skor sebagai refleksi hasil belajar mempunyai kekurangan, yakni guru kurang terinformasi perihal pencapaian kemampuan berbahasa peserta didik secara lebih detail. Callies & Götz (2015) menjelaskan bahwa kemahiran berbahasa seringkali dinilai secara tidak tepat.

Untuk mendukung keberhasilan pengajaran bahasa, dilakukan berbagai kajian pemerolehan bahasa kedua (*second language acquisition research*). Tujuan utama dari kajian-kajian ini adalah untuk mengungkap prinsip-prinsip yang mendasari proses pemerolehan bahasa kedua atau pembelajaran bahasa asing (Ellis, 2015). Mengingat proses pemerolehan dan pembelajaran bahasa bersifat mentalistik, maka pengukuran keberhasilannya dilakukan melalui produk yang dihasilkan oleh pembelajar atau peserta didik, baik berupa produk kebahasaan lisan maupun tulisan (Ellis, 2015; Granger, 1998, 2002).

Meskipun telah banyak dilakukan, kajian-kajian pemerolehan bahasa masih banyak menggunakan data introspektif dan elisitatif yang mempunyai keterbatasan, termasuk dalam kaitannya dengan validitas data. Data elisit yang dihasilkan dari riset eksperimental bersifat terkondisikan sehingga cenderung bersifat artifisial (Granger, 2012; Granger et al., 2015). Sementara itu, data introspektif juga terbatas mengingat kegiatan berbahasa adalah kegiatan kolektif dan memerlukan pengetahuan kolektif, bukan perseorangan (Burkette & Kretzschmar Jr., 2018). Oleh sebab itu, diperlukan data berupa produksi bahasa pembelajar yang bersifat natural dan berjumlah besar untuk memperkuat hasil

kajian pemerolehan bahasa yang selanjutnya digunakan untuk memperbaiki pengajaran bahasa (Ellis, 2015; Granger, 1998, 2008)

Pada era digital dan teknologi informasi dan memasuki era Masyarakat 5.0, data kebahasaan dapat dikumpulkan secara masif dan disimpan dalam *machine-readable format* yang memudahkan proses komputasi. Dengan kata lain, data tersebut disimpan secara digital dan dapat dianalisis secara komputasional. Dalam kaitannya dengan pengajaran bahasa Inggris, data kebahasaan yang berupa produksi tulisan peserta didik dapat dikumpulkan dalam jumlah yang besar sebagai *big data* yang nantinya dapat digunakan untuk menganalisis perkembangan atau pencapaian keterampilan berbahasa peserta didik. Koleksi produksi bahasa yang berukuran besar dan disimpan secara digital ini disebut sebagai sebuah korpus, atau lebih khusus disebut sebagai *learner corpus* atau korpus pembelajar (Flowerdew, 2014; Granger et al., 2015; McEnery et al., 2019).

Walaupun korpus dan korpus pembelajar cukup esensial dalam pengajaran bahasa, akan tetapi keberadaan korpus masih belum banyak dikenal di kalangan praktisi pengajaran bahasa Inggris di Indonesia (Crosthwaite, 2020; Oktavianti, 2022; Oktavianti et al., 2022). Selain penggunaan korpus secara langsung (*Data-driven Learning*) yang belum populer (Crosthwaite, 2020; Nugraha et al., 2017), para pengajar bahasa Inggris masih belum terbiasa mengumpulkan hasil pekerjaan pembelajar secara masif, disimpan dalam format elektronik, dan dianalisis untuk dapat menginformasikan mereka terkait perkembangan bahasa pembelajar (*learner language*). Dengan kehadiran Masyarakat 5.0 dan literasi *big data* yang semakin meningkat, penggunaan *big data* untuk pengajaran bahasa Inggris perlu

diperkenalkan secara luas di Indonesia. Oleh sebab itu, tulisan ini akan menguraikan (i) konsep dan penyusunan korpus pembelajar, (ii) beberapa contoh korpus pembelajar Inggris, (iii) penggunaan korpus pembelajar dalam pengajaran bahasa Inggris, dan (iv) manfaat yang diperoleh dari penggunaan korpus pembelajar dalam pengajaran bahasa Inggris.

Pembahasan

Konsep dan Penyusunan Korpus Pembelajar

Secara umum korpus didefinisikan sebagai kumpulan teks digital dalam jumlah relatif besar (McEnery et al., 2019; Timmis, 2015). Korpus dapat dikelompokkan ke dalam dua tipe, yaitu korpus referensi atau korpus umum dan korpus spesifik (Crawford & Csomay, 2015). Korpus pembelajar merupakan salah satu contoh korpus spesifik dan disusun dari kumpulan produksi bahasa pembelajar baik dalam mode lisan maupun tulisan yang disimpan secara digital (Granger, 2008; Granger et al., 2015). Korpus pada umumnya dikenal sebagai *big data* dalam kajian linguistik dan pengajaran bahasa (Burkette & Kretzschmar Jr., 2018; Oktavianti, 2022). Secara lebih spesifik, korpus pembelajar merupakan *big data* dalam konteks produksi kebahasaan peserta didik sehingga dapat digunakan untuk keperluan penelitian pemerolehan bahasa dan yang lainnya.

Adapun yang dikumpulkan dalam sebuah korpus pembelajar adalah bahasa pembelajar. Bahasa pembelajar (*learner language*) adalah produksi kebahasaan pembelajar ketika mereka menggunakan bahasa kedua atau bahasa non-primer (Granger, 2002; Nesselhauf, 2004). Penelitian-

penelitian pemerolehan bahasa membuktikan bahwa, terlepas dari silabus yang dikembangkan dan digunakan oleh guru, pembelajar atau peserta didik akan mempunyai ‘silabus’ mereka sendiri (urutan perkembangan kebahasaan yang sistematis) (Lightbown & Spada, 2021).

Meskipun korpus pembelajar mempunyai fungsi penting dalam penelitian dan pengajaran bahasa, penyusunannya memerlukan perhatian yang serius. Terdapat beberapa aspek yang harus dicermati saat penyusunan korpus tersebut, khususnya terkait jenis data pembelajar yang akan dikumpulkan. Ishikawa (2023) menyatakan bahwa penyusunan korpus pembelajar perlu memperhitungkan (i) apa definisi pembelajar bahasa kedua atau bahasa asing yang akan dianut, (ii) jenis tulisan atau ujaran apa yang akan dikumpulkan, dan (iii) dan bagaimana proses pengumpulan data korpus akan dilakukan. Tabel 1 menampilkan tipe data korpus pembelajar yang perlu dipersiapkan sebelum melakukan pengumpulan korpus.

Tabel 1. Tipe data korpus pembelajar (Ishikawa, 2023)

Category	Major subcategories
(i) Age/School Year	Children Secondary school students College students Adult speakers
Proficiency	Beginner (Pre-A1), Novice (A1-A2) Intermediate (B2-B2) Advanced (C1-C2) Professional (English as a lingua franca)
Area	Local/domestic International

Category	Major subcategories
(ii) Production Mode	Written (academic essays, business letters, research papers) Spoken (monologue, dialogue, interview)
(iii) Data Modality	Mono-modal (text only) Multi-modal (text + audio/video)
Period	Cross-sectional (data collected at one time) Longitudinal
Annotation	Unannotated Annotated (POS, error, speech feature)

Sementara itu, Granger (2002) dan Granger et al. (2015) menegaskan bahwa korpus pembelajar perlu disusun dengan sistematis dan memperhatikan dua variabel utama, yakni pembelajar dan pengaturan tugas untuk memperoleh produksi bahasa pembelajar tersebut. Tabel 2 memaparkan rincian komponen tiap variabel yang perlu diperhatikan dalam penyusunan korpus pembelajar.

Tabel 2. Desain korpus pembelajar (Granger, 2002; Granger et al., 2015)

Learner	Task Setting
<ul style="list-style-type: none"> ● Learning context ● Mother tongue ● Other foreign languages ● Proficiency 	<ul style="list-style-type: none"> ● Time limit ● Use of reference tools ● Exam ● Audience/interlocutor

Beberapa Korpus Pembelajaran Bahasa Inggris

Adapun beberapa contoh korpus pembelajaran dalam bahasa Inggris adalah *International Corpus of Learner English* (ICLE), *Michigan Corpus of Upper-level Student Paper* (MICUSP), *British Academic Written Corpus* (BAWE), dan *International Corpus Network of Asian Learners of English* (ICNALE).

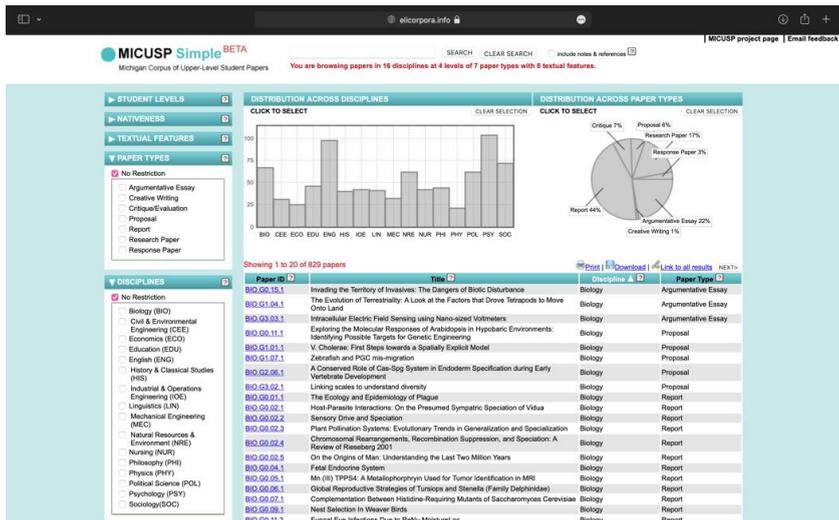
International Corpus of Learner English (ICLE)

International Corpus of Learner English (ICLE) adalah korpus pembelajar yang terdiri dari esai pembelajar bahasa Inggris pada tingkat *upper intermediate* hingga *advanced* yang disusun oleh Sylviane Granger di *University of Louvain* pada 2002. Korpus pembelajar ini berukuran 2,5 juta token saat pertama kali diluncurkan dan bertambah menjadi 3,9 juta token pada versi keduanya tahun 2009. Sementara itu, pada versi terakhirnya, yakni versi ketiga, ICLE berukuran 5,5 juta kata dan mencakup pembelajar dengan 25 ragam bahasa ibu. Di samping itu, versi ketiga juga dapat diakses dalam sebuah website yang memudahkan akses dan memuat lebih banyak subkorpus baru. Korpus ini memerlukan pembelian *license* sebelum dapat digunakan sehingga aksesnya terbatas, meskipun menjadi pelopor korpus pembelajar bahasa Inggris di dunia.

Michigan Corpus of Upper-level Student Papers (MICUSP)

Michigan Corpus of Upper-level Student Papers (MICUSP) merupakan salah satu bagian dari korpus pembelajar yang disusun oleh *Michigan University* dan merupakan kumpulan data tulisan. Adapun kumpulan data tuturan atau ujaran dikumpulkan dalam *Michigan Corpus of Spoken Academic English*. Keduanya dikembangkan oleh *English Language*

Institute pada universitas tersebut. Para guru dan pembelajar dari seluruh dunia dapat memanfaatkan korpus ini karena tersedia secara daring dan tidak berbayar. Gambar 1 merupakan tampilan MICUSP yang dapat diakses secara daring.



Gambar 1. Tampilan MICUSP

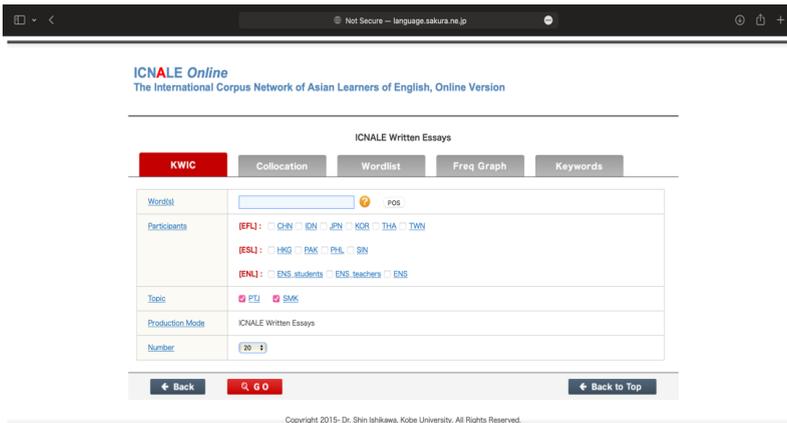
British Academic Written English (BAWE)

British Academic Written English Corpus (BAWE) dikumpulkan sebagai bagian dari proyek bertajuk *An Investigation of Genres of Assessed Writing in British Higher Education*. Proyek ini didanai oleh Dewan Penelitian Ekonomi dan Sosial dengan Nomor proyek 2004 - 2007 RES-000-23-0800 (Nesi, n.d.). BAWE terdiri atas esai mahasiswa dengan kemampuan bahasa Inggris tingkat mahir di universitas pada pergantian abad ke-21. Korpus ini berisi kurang dari 3000 tugas siswa berstandar baik dengan jumlah 6.506.995 kata. Adapun cakupan esai adalah empat disiplin

ilmu, antara lain Seni dan Humaniora, Ilmu Sosial, Ilmu Hayati dan Ilmu Fisika) dan tingkat sarjana dan magister.

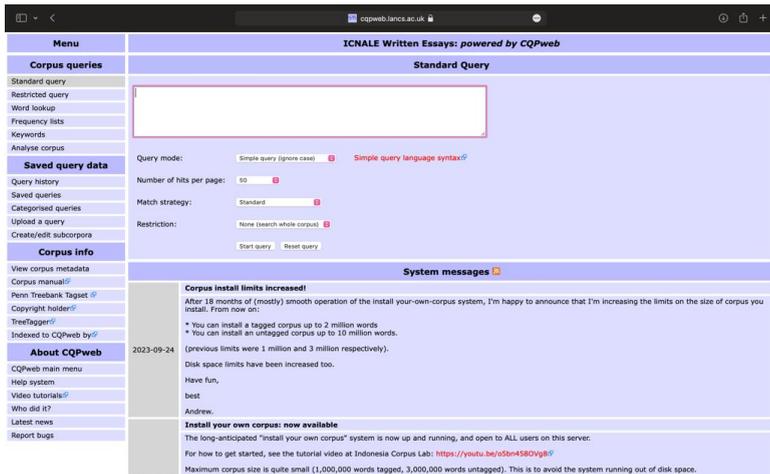
International Corpus Network of Asian Learners of English (ICNALE)

International Corpus Network of Asian Learners of English (ICNALE) terdiri atas 14.000 sampel pidato dan esai yang dihasilkan oleh sekitar 4.000 pelajar di 10 negara dan wilayah di Asia, serta penutur asli bahasa Inggris yang dikumpulkan sejak 2007. Data yang dikumpulkan telah dirilis sebagai modul ICNALE yang berbeda. ICNALE saat ini terdiri dari tiga modul inti: Esai Tertulis, Monolog Lisan, dan Dialog Lisan, serta dua modul tambahan: Esai yang Diedit dan Arsip Pemingkatan Global. Ini telah menjadi salah satu korpora pembelajar (LC) terbesar yang berfokus pada pembelajar bahasa Inggris di Asia. Gambar 2 menunjukkan tampilan *web ICNALE Online*.



Gambar 2. Tampilan ICNALE

Selain dapat diakses pada *web* tersendiri, ICNALE juga dapat diakses dari *Corpus Query Program Web (CQPWeb)* yang dikelola oleh *Lancaster University*. Gambar 3 merupakan tampilan ICNALE dari *CQPWeb*.



Gambar 3. Tampilan ICNALE dalam CQPWeb

Beberapa contoh yang telah dijelaskan di atas (ICLE, MICUSP, BAWE, dan ICNALE) adalah korpora daring yang sudah tersedia atau dapat dibeli izin aksesnya. Para pengajar atau peneliti juga dapat menyusun korpus pembelajar dari produksi pembelajar yang diajarnya. Dengan demikian data yang dikumpulkan dapat menjadi lebih kontekstual dan hasil pencermatannya akan dapat membantu pengajar untuk merumuskan pembelajaran yang lebih efektif. Namun, jenis korpus yang disusun sendiri biasanya berukuran lebih kecil (*small corpus*) dibandingkan dengan yang sudah tersedia secara daring di Internet.

Penggunaan korpus pembelajar dalam pengajaran bahasa Inggris (MICUSP)

Korpora pembelajar yang dibahas pada bagian sebelumnya merupakan korpora yang dapat digunakan sebagai bahan pengayaan pengajaran bahasa Inggris atau sebagai refleksi dalam penyusunan materi ajar bahasa Inggris.

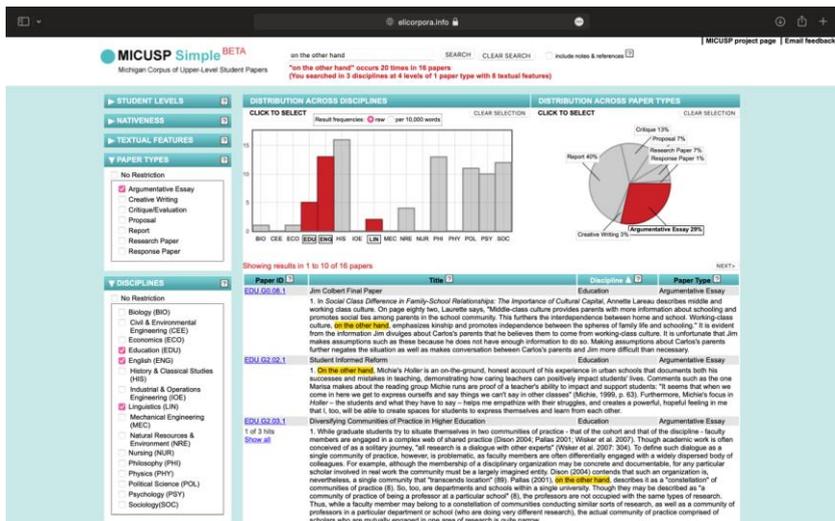
Penggunaan korpus pembelajar yang tersedia daring sebagai bahan pengayaan dan refleksi

Pada dasarnya penggunaan korpus dalam pengajaran bahasa digunakan untuk memberikan informasi terhadap bagaimana bahasa Inggris sesungguhnya digunakan. Materi-materi yang disusun berdasarkan hasil riset korpus disebut sebagai *corpus-informed materials* (McCarthy & McCarten, 2022; Meunier & Reppen, 2015) karena sifatnya yang terinformasi dari hasil pencarian/kueri di korpus. Materi yang disusun berdasarkan hasil korpus diharapkan mendekati penggunaan bahasa Inggris yang aktual. Dengan demikian, kegiatan pembelajaran dengan menggunakan materi tersebut dapat menyiapkan peserta didik untuk berkomunikasi secara kompeten di kehidupan sehari-hari, khususnya di negara/wilayah berbahasa Inggris (Gilmore, 2007).

Tidak jauh berbeda, korpus pembelajar juga dapat digunakan untuk menginformasikan para pengajar, penyusun buku ajar, dan pembuat kebijakan tentang penggunaan satuan lingual atau fitur gramatika yang dihasilkan pembelajar, baik secara lisan maupun tulisan. Beberapa korpora daring yang disebutkan sebelumnya dihasilkan baik oleh penutur jati bahasa Inggris, penutur bahasa Inggris

sebagai bahasa kedua, dan penutur bahasa Inggris sebagai bahasa asing. Korpus pembelajar daring tersebut dapat digunakan sebagai bahan refleksi dan pengayaan pengajar bahasa Inggris. Selain itu, para pengajar juga dapat menjadikan hasil kueri dari korpora tersebut untuk dibandingkan dengan produksi kebahasaan pembelajar yang sedang diajar.

Sebagai contoh, pengajar bahasa Inggris yang akan mengajarkan materi pemarkah wacana *on the other hand* dapat melakukan kueri pada MICUSP, dan hasil dari kueri tersebut dapat digunakan sebagai bahan ajar atau bahan refleksi terhadap materi yang akan diajarkan. Gambar 4 menunjukkan hasil kueri *on the other hand* pada esai argumentatif dengan spesifikasi bidang ilmu tertentu, yaitu *Education, English, dan Linguistics*.

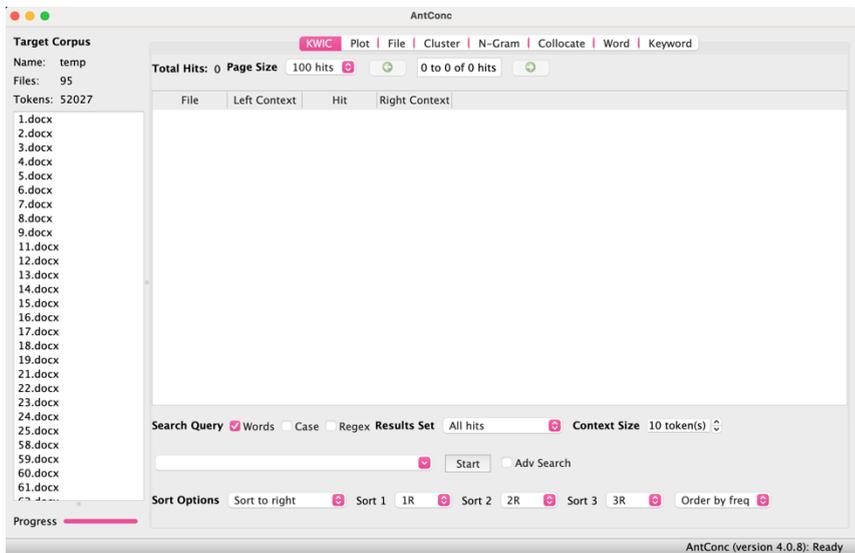


Gambar 4. Hasil kueri on the other hand pada MIUCSP

Dari hasil kueri tersebut nampak bahwa *on the other hand* digunakan tidak hanya di posisi inisial, tetapi juga medial. Selanjutnya, pengajar bahasa Inggris dapat menjadikan temuan tersebut sebagai bahan rujukan, baik dalam penyusunan materi ajar, juga sebagai pembanding dari hasil kueri terhadap korpus pembelajar yang dikumpulkan dari peserta didiknya.

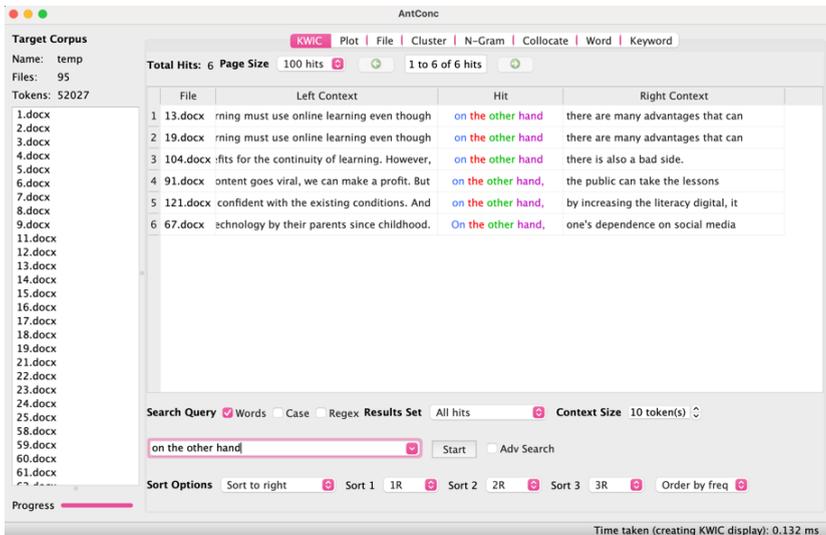
Penggunaan korpus pembelajar yang disusun dari data bahasa pembelajar sendiri

Korpus pembelajar disimpan dalam format elektronik sehingga dapat dianalisis dengan menggunakan *software* atau *tool* yang dikembangkan oleh linguist (Granger, 2002; Granger et al., 2015). Adapun keuntungan penyimpanan data produksi kebahasaan pembelajar secara digital, antara lain data tersebut dapat diatur dan dianalisis dengan lebih mudah (Granger, 2002; Granger et al., 2015). Beberapa *corpus tool* atau piranti korpus yang dapat digunakan untuk menganalisis korpus pembelajar yang sudah dikumpulkan oleh peneliti atau pengajar adalah *AntConc* (Anthony, 2022). Di bawah ini adalah contoh korpus pembelajar yang dikumpulkan dari 95 esai mahasiswa kelas *Academic Writing* dari Prodi Pendidikan Bahasa Inggris. Semua esai tersebut dapat dientrikan ke dalam *AntConc* dengan fitur *Create Quick Corpus*. Setelah tersedia dalam *AntConc*, dapat dilihat bahwa korpus pembelajar tersebut mempunyai 52.027 kata. Berikut Gambar 5 menyajikan tampilan *AntConc* dan korpus pembelajar yang sudah dientri ke dalam *software* tersebut.



Gambar 5. Contoh korpus pembelajaran dienti pada AntConc

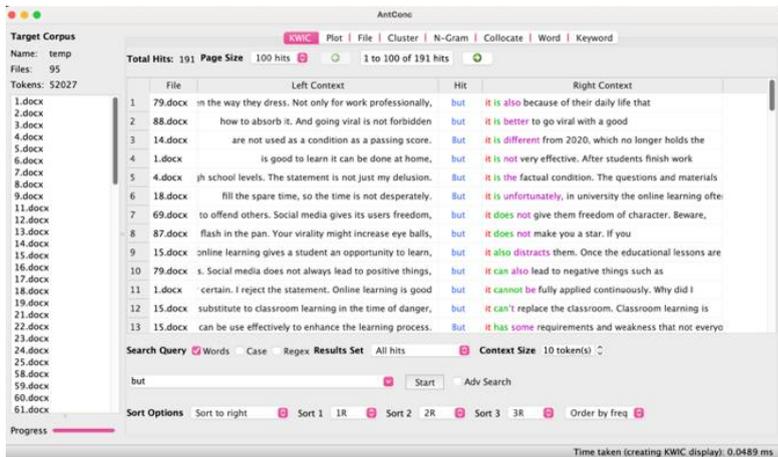
Setelah korpus termuat dalam *AntConc*, analisis kebahasaan dapat dilakukan pada korpus tersebut. Hal ini penting bagi pengajar bahasa Inggris yang ingin mengetahui perkembangan kebahasaan peserta didik: apakah mereka sudah tepat dalam penggunaan fitur lingual atau *lexical bundles* tertentu. Sebagai contoh, pengajar dapat meneliti penggunaan *lexical bundle* yang juga berfungsi sebagai pemarkah wacana *on the other hand* dalam data kebahasaan peserta didik dan diperoleh hasil seperti dalam Gambar 6.



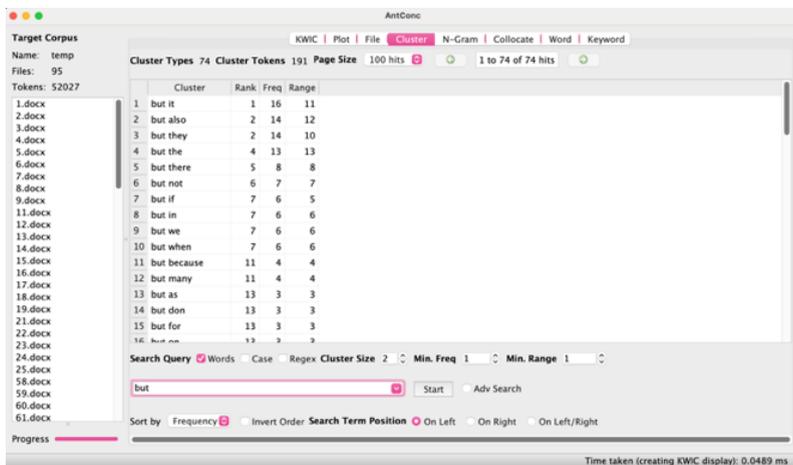
Gambar 6. Hasil kueri on the other hand menggunakan AntConc

Dari data di atas dapat diamati bahwa *on the other hand* tidak hanya digunakan di posisi inisial, tetapi juga posisi medial kalimat. Hasil ini selanjutnya dapat dibandingkan dengan penelusuran pada korpus bahasa Inggris seperti *Corpus of Contemporary American English (COCA)*, *British National Corpus (BNC)*, atau *Global Web-based English (GloWbE)*.

Contoh lain adalah kueri disjungsi *but* pada korpus pembelajar sebagaimana tampak pada Gambar 7(a) dan 7(b) di bawah ini.



Gambar 7(a) Hasil kueri but pada fitur KWIC AntConc

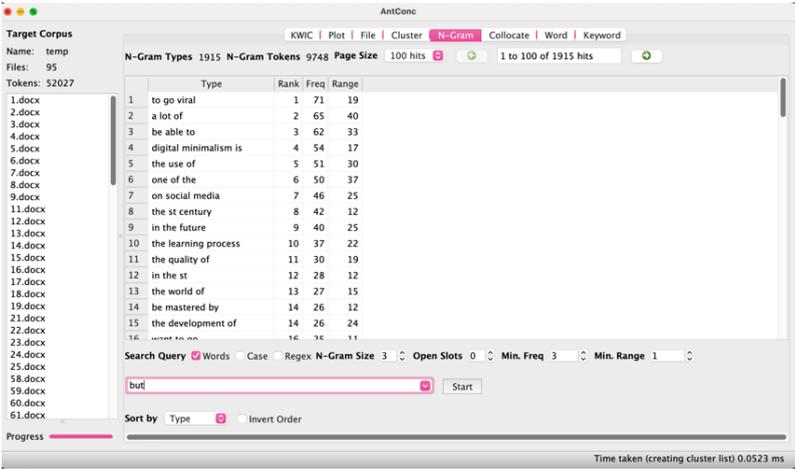


Gambar 7(b). Hasil kueri but pada fitur Cluster, AntConc

Hasil pencarian *but* pada korpus tersebut menunjukkan bahwa pembelajar sudah cukup beragam menggunakan *but* pada posisi inisial dan media, selaras dengan hasil kueri pada MICUSP. Namun, masih perlu pencermatan terhadap kombinasi atau konkordansi *but* pada kalimat yang

diproduksi pembelajar. Dengan menggunakan fitur *cluster* pada Gambar 7, dapat dilihat bahwa disjungsi *but* berkonkordansi atau berkookurensi dengan beragam satuan lingual lainnya, termasuk pronomina, penegasi, kondisional *if*, dan lain-lainnya. Terlepas dari beberapa kookurensi yang lazim, ditemukan penggunaan *but because* yang perlu dicermati lebih lanjut sebagai sebuah temuan. Dengan demikian, pengajar bahasa Inggris dapat membuat keputusan tentang apa yang selanjutnya perlu ditekankan dan apa yang dapat dikurangi dari materi ajar atau kegiatan pembelajaran.

Selain fitur KWIC dan *cluster*, piranti korpus biasanya mempunyai fitur *n-gram* yang memungkinkan hasil kueri berupa kombinasi kata yang sering digunakan bersamaan dan mempunyai fungsi wacana (*lexical bundles*) (Biber et al., 2021; Hyland & Jiang, 2018; Oktavianti & Prayogi, 2022). Gambar 8 berikut menunjukkan hasil kueri *n-gram*.



Gambar 8. Hasil kueri n-gram dengan AntConc

Hasil kueri pada Gambar 8 tersebut dapat menginformasikan pengajar perihal *lexical bundles* yang sering diproduksi oleh peserta didik. Adapun penguasaan *lexical bundles* dapat menunjukkan tingkat kemahiran berbahasa seseorang (Bal Gezegin, 2019; Fajri et al., 2020; Liu & Chen, 2020; Oktavianti & Sarage, 2021). Oleh sebab itu, perlu dipastikan bahwa *lexical bundles* yang diproduksi pembelajar kita sudah tepat dengan memeriksa melalui fitur *n-gram* pada *AntConc*.

Manfaat Korpus Pembelajar

Dengan mengamati pola-pola pada bahasa pembelajar, para pengajar bahasa dapat menyesuaikan dengan materi ajar, silabus, dan aspek pedagogis lainnya sehingga dapat memperkenalkan pembelajaran bahasa yang lebih baik (Ellis, 2015). Korpus pembelajar sebagai sebuah bank data produksi pembelajar dapat membantu para pengajar untuk mencapai tujuan tersebut. Secara lebih terperinci, berikut adalah beragam manfaat korpus pembelajar dalam pengajaran bahasa.

Evaluasi Kemampuan Belajar Bahasa

Korpus pembelajar umumnya dimanfaatkan untuk mengevaluasi kemampuan belajar bahasa. Evaluasi diperlukan guna memahami tahap-tahap perkembangan keterampilan berbahasa para pembelajar (De Florio, 2023). Dengan korpus pembelajar, kita bisa mendapatkan wawasan yang lebih mendalam tentang sejauh mana pembelajar telah menguasai bahasa sasaran. Selain itu, guru, instruktur, atau peneliti dapat mengetahui secara terukur area-area kebahasaan mana yang memerlukan perbaikan. Evaluasi kemampuan belajar bahasa dengan korpus pembelajar

membantu pengajar untuk memberikan umpan balik yang lebih terperinci (Callies & Götz, 2015). Hal ini juga memungkinkan pengajar untuk menyesuaikan pembelajaran dengan lebih baik sesuai dengan kebutuhan.

Pengembangan Materi Ajar

Materi ajar yang disusun sesuai dengan tahap-tahap perkembangan kemampuan berbahasa dan mempertimbangkan data bahasa dari korpus pembelajar dapat menghasilkan materi yang lebih sesuai dengan kebutuhan pembelajar dan memaksimalkan efektivitas pembelajaran (McCarthy & McCarten, 2022; McEnery et al., 2019; Meunier & Reppen, 2015). Dengan demikian, materi ajar tidak disusun berdasarkan intuisi, imajinasi, atau karangan dari penulis, tetapi didasarkan atas data empiris dari korpus pembelajar. Artinya, materi ajar yang disusun dan dikembangkan sudah sesuai dengan tahap kemampuan belajar si pembelajar sehingga tepat sasaran.

Memahami Kesalahan Berbahasa

Kesalahan berbahasa merupakan hal yang wajar dan pasti terjadi dalam proses belajar bahasa (Ellis, 2015; James, 2013). Dengan menganalisis data yang terdapat pada korpus pembelajar, kita dapat mengidentifikasi kesalahan berbahasa yang kerap dilakukan oleh pembelajar agar selanjutnya dapat membantu pembelajar dan pengajar dalam mengatasi kesalahan-kesalahan tersebut (Flowerdew, 2014; Granger et al., 2015; McEnery et al., 2019). Misalnya, kesalahan gramatika, diksi, bahkan pelafalan. Dengan diketahuinya kesalahan-kesalahan yang acap terjadi, guru dapat mengajar dengan lebih spesifik.

Penelitian Perkembangan Bahasa

Korpus pembelajar sering dimanfaatkan dalam kajian perkembangan bahasa manusia, sejak anak-anak, remaja, dan dewasa. Sebagai contoh, penelitian perkembangan bahasa pembelajar dan dinamikanya, misalnya korpus pembelajar tingkat PAUD, SD, SMP, SMA, perguruan tinggi, atau dalam konteks pembelajaran bahasa asing. Selain itu, tingkat kemahiran seseorang juga dapat lebih terpetakan dengan dukungan data empiris produksi satuan lingual atau konstruksi lingual, khususnya mengenai materi yang diajarkan (contohnya, konstruksi pasif, dll) (Callies & Götz, 2015). Mengikuti tahapan belajar bahasa kedua yang dikemukakan oleh Krashen (1982), data yang ditemukan dari hasil kajian terhadap korpus pembelajar dapat mendukung para pengajar untuk mencermati kemampuan peserta didik lebih komprehensif.

Simpulan

Dari paparan di atas, dapat ditegaskan bahwa korpus pada umumnya memiliki kebermanfaatan yang signifikan bagi pengajaran bahasa Inggris, termasuk korpus pembelajar. Sebagai *big data*, korpus pembelajar dapat memperkuat hasil riset pemerolehan bahasa terkait perkembangan bahasa pembelajar dan dapat menjadi bahan acuan bagi para pengajar, penyusun buku teks, dan pembuat kebijakan untuk merumuskan apa yang terbaik bagi peserta didik, khususnya di konteks EFL. Terdapat beragam pilihan korpus pembelajar daring yang dapat diakses oleh pengajar atau pendidik bahasa Inggris. Selain itu, pengajar juga dapat mengkompilasi korpus pembelajar yang diambil dari produksi lisan atau tulisan peserta didiknya sehingga data

yang didapat lebih kontekstual. Akan tetapi, agar dapat lebih mengenal dan siap menggunakan korpus, pengajar atau pendidik bahasa Inggris perlu diberi bekal yang cukup mengenai hakikat korpus, cara kerja korpus, dan penggunaannya di ruang kelas dan dalam aktivitas pengajaran bahasa lainnya. Oleh sebab itu, keberadaan mata kuliah *corpus linguistics* atau *corpus-based language pedagogy* pada program studi Pendidikan Bahasa Inggris perlu digalakkan. Adapun paparan dalam tulisan ini masih bersifat *preliminary* dan perlu dilanjutkan dengan pembahasan yang lebih mendalam tentang penggunaan korpus pembelajar di ruang kelas dan didukung dengan riset empiris tentang efektivitasnya.

Daftar Pustaka

- Anthony, L. (2022). *AntConc* (4.0.3) [Computer software]. Waseda University. <https://www.laurenceanthony.net/software>
- Bal Gezegin, B. (2019). Lexical bundles in published research articles: A corpus-based study. *Dil ve Dilbilimi Çalışmaları Dergisi*, 15(2), 520–534. <https://doi.org/10.17263/jlls.586188>
- Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of Spoken and Written English*. John Benjamins Publishing Company. <https://doi.org/10.1075/z.232>
- Burkette, A., & Kretzschmar Jr., W. A. (2018). *Exploring Linguistic Science: Language Use, Complexity, and Interaction* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781108344326>
- Callies, M., & Götz, S. (Eds.). (2015). *Learner Corpora in Language Testing and Assessment* (Vol. 70). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.70>
- Crawford, W., & Csomay, E. (2015). *Doing Corpus Linguistics* (0 ed.). Routledge. <https://doi.org/10.4324/9781315775647>
- Crosthwaite, P. (2020, October 29). *Trainee EFL teachers' DDL lesson planning: Improving corpus-focuses TPACK in Indonesia*. UCREL CRS Webinar, Lancaster. <https://www.youtube.com/watch?v=ZQs-Um9PLWo>
- De Florio, I. (2023). *From Assessment to Feedback: Applications in the Second/Foreign Language Classroom* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009218948>
- Ellis, R. (2015). *Understanding second language acquisition*. Oxford University Press.
- Fajri, M. S. A., Kirana, A. W., & Putri, C. I. K. (2020). Lexical bundles of L1 and L2 English professional scholars: A contrastive corpus-driven study on applied linguistics research articles. *Journal of Language and Education*, 6(4), 76–89. <https://doi.org/10.17323/jle.2020.10719>

- Flowerdew, L. (2014). Learner corpus research in EAP: Some key issues and future pathways. *영어학연구*, 20(2), 43–60. <https://doi.org/10.17960/ELL.2014.20.2.003>
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97–118. <https://doi.org/10.1017/S0261444807004144>
- Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA. In S. Granger (Ed.), *Learner English on computer*. Longman.
- _____. (Ed.). (2002). *Computer learner corpora, second language acquisition and foreign language teaching*. Benjamins.
- _____. (2008). Learner corpora. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook volume I* (pp. 259–274). Walter de Gruyter.
- _____. (2012). Learner Corpora. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (p. wbeal0669). Blackwell Publishing Ltd. <https://doi.org/10.1002/9781405198431.wbeal0669>
- _____, Gilquin, G., & Meunier, F. (2015). Introduction: Learner corpus research – past, present and future. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (1st ed., pp. 1–6). Cambridge University Press. <https://doi.org/10.1017/CBO9781139649414.001>
- Harta, I. G. W., Bay, I. W., & Ali, S. W. (2021). An Analysis of Lexical Collocation Errors in Students' Writing. *TRANS-KATA: Journal of Language, Literature, Culture and Education*, 2(1), 15–25. <https://doi.org/10.54923/transkata.v2i1.18>
- Hasanuddin, N., & Ciptaningrum, D. S. (2021). Types of English Teachers' Questioning and Classroom Interaction Affecting Learners' Motivation in EFL Context. *Indonesian Journal of EFL and Linguistics*, 6(2), 455–472.
- Hyland, K., & Jiang, F. (Kevin). (2018). Academic lexical bundles: How are they changing? *International Journal of Corpus*

- Ishikawa, S. (2023). *The ICNALE Guide: An Introduction to a Learner Corpus Study on Asian Learners' L2 English* (1st ed.). Routledge. <https://doi.org/10.4324/9781003252528>
- James, C. (2013). *Errors in language learning and use: Exploring error analysis*. Longman.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition* (1st ed.). Pergamon.
- Lauder, A. (2008). THE STATUS AND FUNCTION OF ENGLISH IN INDONESIA: A REVIEW OF KEY FACTORS. *Makara Human Behavior Studies in Asia*, 12(1), 9. <https://doi.org/10.7454/mssh.v12i1.128>
- Lightbown, P. M., & Spada, N. (2021). *How languages are learned*. Oxford University Press.
- Liu, C.-Y., & Chen, H.-J. H. (2020). Analyzing the functions of lexical bundles in undergraduate academic lectures for pedagogical use. *English for Specific Purposes*, 58, 122–137. <https://doi.org/10.1016/j.esp.2019.12.003>
- McCarthy, M., & McCarten, J. (2022). Writing corpus-informed materials. In J. Norton & H. Buchanan (Eds.), *The Routledge handbook of materials development for language teaching* (pp. 170–184). Routledge.
- McEneaney, T., Brezina, V., Gablasova, D., & Banerjee, J. (2019). Corpus Linguistics, Learner Corpora, and SLA: Employing Technology to Analyze Language Use. *Annual Review of Applied Linguistics*, 39, 74–92. <https://doi.org/10.1017/S0267190519000096>
- Meunier, F., & Reppen, R. (2015). Corpus versus non-corpus-informed pedagogical materials: Grammar as the focus. In *The Cambridge handbook of English corpus linguistics* (pp. 498–514). Cambridge University Press.

- Nabila, A., Cahyono, B. Y., & Khoiri, N. E. (2021). DEMOTIVATION LEVEL AND DEMOTIVATORS AMONG EFL STUDENTS IN HOME ONLINE ENGLISH LEARNING DURING THE PANDEMIC. *JEELS (Journal of English Education and Linguistics Studies)*, 8(2), 243–271. <https://doi.org/10.30762/jeels.v8i2.3567>
- Nesi, H. (n.d.). (BAWE) British Academic Written English Corpus [University website]. *Conventry University*. Retrieved October 21, 2023, from <https://www.coventry.ac.uk/research/research-directories/current-projects/2015/british-academic-written-english-corpus-bawe/>
- Nesselhauf, N. (2004). *Collocations in a learner corpus*. J. Benjamins Pub. Co.
- Nugraha, S. I., Miftakh, F., & Wachyudi, K. (2017). Teaching Grammar through Data-Driven Learning (DDL) Approach. *Proceedings of the Ninth International Conference on Applied Linguistics (CONAPLIN 9)*. Ninth International Conference on Applied Linguistics (CONAPLIN 9), Bandung, Indonesia. <https://doi.org/10.2991/conaplin-16.2017.68>
- Oktavianti, I. N. (2022). Corpora and data-driven learning: Big data for language teachers. In D. Karmiyati (Ed.), *Embracing Society 5.0 with humanity*. CV. Bildung Nusantara.
- _____, & Prayogi, I. (2022). Discourse functions of lexical bundles in Indonesian EFL learners' argumentative essays: A corpus study. *Studies in English Language and Education*, 9(2), 761–783. <https://doi.org/10.24815/siele.v9i2.23995>
- _____, & Sarage, J. (2021). Lexical Bundles in Students' Argumentative Essays: A Study of Learner Corpus. *Indonesian Journal of EFL and Linguistics*, 6(2), 510–534.
- _____, Triyoga, A., & Prayogi, I. (2022). Corpus for language teaching: Students' perceptions and difficulties. *PROJECT (Professional Journal of English Education)*, 5(2), 441. <https://doi.org/10.22460/project.v5i2.p441-455>

- Setiarini, N. L. P. (2018). Collocation Errors by Indonesian EFL Learners: Types of Errors, Translation Techniques, and Causes of Errors. *Proceedings of the Fourth Prasasti International Seminar on Linguistics (Prasasti 2018)*. Fourth Prasasti International Seminar on Linguistics (Prasasti 2018), Surakarta, Indonesia. <https://doi.org/10.2991/prasasti-18.2018.95>
- Timmis, I. (2015). *Corpus Linguistics for ELT* (0 ed.). Routledge. <https://doi.org/10.4324/9781315715537>.

PROFIL PENULIS

Dr. Ikmi Nur Oktavianti, M.A., dosen di Program Studi Magister Pendidikan Bahasa Inggris, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Ahmad Dahlan. Ia merupakan lulusan S-1 Sastra Inggris Universitas Brawijaya, S-2 Ilmu Linguistik Universitas Gadjah Mada, dan S-3 Ilmu-ilmu Humaniora Universitas Gadjah Mada. Bidang penelitian yang ditekuni adalah *corpus-based language pedagogy*, sintaksis, dan semantik.

Icuk Prayogi, M.A., dosen di Program Studi Pendidikan Bahasa dan Sastra Indonesia, Fakultas Pendidikan Bahasa dan Sastra, Universitas PGRI Semarang. Ia merupakan lulusan S-1 Sastra Indonesia Universitas Gadjah Mada dan S-2 Ilmu Linguistik Universitas Gadjah Mada. Saat ini, penulis tengah menempuh studi doctoral di S-3 Ilmu-ilmu Humaniora di Universitas Gadjah Mada. Bidang penelitian yang ditekuni adalah linguistik kognitif dan linguistik korpus.