

**APLIKASI SOFTWARE CONCORDANCE PROGRAM
DALAM PENGAJARAN DAN PENELITIAN BAHASA
(studi kasus pada mahasiswa semester 6
Jurusan Bahasa dan Sastra, Universitas Pendidikan Indonesia)**

YANTY WIRZA

*Jurusan Pendidikan Bahasa Inggris
Universitas Pendidikan Indonesia*

This study examines the students' response towards the importance of and challenges faced in using concordance software for language studies. There were 68 students of two classes participated in the study and were assigned to conduct research using the concordance software. The findings revealed that the students found that the software was functional and practical for the use of research in language studies as they only needed very short time to process the data. However, more practice is required so they can develop their analysis ability. The respondents also urged that facilities needed to support the proper and convenient use of the program should be made available.

Key words: *corpus linguistics, concordance software, language studies.*

Pendahuluan.

Perkembangan teknologi dalam dunia pendidikan dan pengembangan bahasa yang biasa dikenal dengan istilah CALL (*Computer-Assisted Language Learning*) dalam berbagai aplikasinya telah banyak merubah bentuk dan cara pengajaran dan penelitian bahasa. Salah satu aplikasi teknologi yang digunakan terutama dalam bidang linguistik adalah berbagai software corpus linguistics yang digunakan dalam berbagai penelitian, pengajaran dan, dan pengembangan bahasa. Sebagai jurusan yang *core business*-nya bergerak dalam bidang pendidikan yang juga mencakup bidang linguistik murni dan terapan, Jurusan Pendidikan Bahasa Inggris Universitas Pendidikan Indonesia telah merespon perkembangan dalam CALL dengan menawarkan Mata Kuliah Perluasan dan Pendalaman Corpus Linguistics yang termasuk dalam paket Mata Kuliah Linguistik pada mahasiswa semester

enam sejak diberlakukannya kurikulum yang terbaru tahun 2006. Oleh karena itu, studi penajakan respons mahasiswa terhadap matakuliah ini perlu dilakukan untuk mengetahui pendapat dan masukan dari mahasiswa perihal mata kuliah Corpus Linguistics ini.

Definisi dan Sejarah Korpus

Korpus berasal dari bahasa Latin yang berarti "body" atau kumpulan yang dalam makna yang terdekat berarti kumpulan teks dari suatu variasi bahasa yang dapat semaksimal mungkin merepresentasikan suatu kecenderungan atau fenomena dari variasi tersebut (Mc Eney dan Wilson, 2007). Oleh karena itu, kita tidak berbicara tentang *beberapa* teks saja, akan tetapi kumpulan teks yang bisa saja terdiri dari ratusan ribu bahkan jutaan kata untuk dapat melihat suatu fenomena bahasa.

Dalam sejarahnya, korpus dalam bidang linguistic dibagi dalam dua fase: fase awal dan fase modern. Pada fase awal, korpus

dilakukan dengan cara manual dengan mengumpulkan bahasa lisan (misalnya ujaran) dan tertulis dalam waktu yang sangat lama bahkan bertahun-tahun untuk dapat mencapai jumlah data yang cukup representative. Pada fase ini, Chomsky, seorang linguist yang sangat ternama, sempat mengkritik penggunaan korpus karena korpus hanya mewakili *language performance* (penggunaan riil bahasa) dan dianggap tidak mewakili apa yang disebut *language competence* (pengetahuan tentang bahasa), yang justru dibutuhkan untuk membangun suatu teori bahasa.

Kenyataannya, linguist yang meyakini pentingnya korpus tetap melakukan penelitian mereka dan tercatat ada beberapa penelitian yang hingga kini terbukti menjadi dasar-dasar dari pengembangan teori bahasa. Diantara penelitian yang dilakukan pada fase awal korpus adalah studi tentang pemerolehan bahasa (*language acquisition*) pada periode pertengahan abad ke-18 hingga pertengahan abad ke-19 yang didapat dari kumpulan dan rekaman ucapan-ucapan anak serta diari para orang tua. Penelitian lain yang penting adalah tentang pengajaran bahasa (*language pedagogy*) yang sudah dimulai sejak awal hingga pertengahan abad ke-19 yang sangat berpengaruh terhadap pengajaran bahasa asing. Adapun penelitian lain adalah dalam bidang sintaksis, semantic, dan konvensi ejaan.

Pada fase modern, seiring dengan perkembangan teknologi komputer, penelitian yang berbasis korpus terus berkembang dan semakin banyak dilakukan. Bahkan, saat ini, istilah korpus hampir sinonim dengan istilah mesin pembaca korpus ataupun software korpus. Dalam prosesnya, data dilengkapi dengan informasi linguistik tertentu sesuai dengan tipe anotasinya. Salah satu software korpus yang paling banyak digunakan adalah *concordance program* yang dapat melakukan hal-hal sebagai berikut: mencari (search) suatu kata/

frase, menampilkan (retrieve) kata/frase yang diperlukan, menghitung (calculate) jumlah atau frekuensi penggunaan kata/frase, serta memilih (sort) kata/frase berdasarkan kategori tertentu (Meyer, 2002). Informasi ini dapat tersaji dalam waktu yang sangat singkat bahkan dalam beberapa detik saja. Beberapa korpus yang terkenal adalah Leuven Drama Corpus (1975), LOB Corpus (1978) London-Lund Corpus (1982), Brown Corpus (1982), dan COBUILD (1987).

Penggunaan Software Korpus dalam Pengajaran dan Penelitian.

Sebagaimana telah disebutkan dimuka, meskipun pada masa-masa dimana korpus banyak menuai kontroversi, penggunaan mesin dan software korpus terus meningkat. Pentingnya korpus dalam pengajaran dan penelitian bahasa terutama karena data korpus berasal dari data empiris yang dikumpulkan secara sistematis (McEnery and Wilson, 2007). Ditambah dengan kemampuan komputer untuk memproses data dalam jumlah besar, maka korpus menjadi suatu alat analisis yang penting sekaligus praktis dalam pengajaran dan penelitian bahasa (Milroy, 1987).

Hingga kini, telah banyak penelitian yang berbasis korpus tanpa diragukan lagi keterwakilannya terhadap variasi bahasa yang menjadi object penelitiannya. Hampir semua aspek bahasa seperti studi tentang ujaran (speech), kosa kata (lexicon), makna kata/frase/klausa (semantics), penggunaan bahasa (pragmatics), bahasa dan masyarakat (sociolinguistics), bahasa dan budaya (cultural studies), dan lain sebagainya telah menggunakan korpus sebagai basis data yang valid dan representatif.

Dalam suatu kelas bahasa atau yang berkaitan dengan bahasa, telah banyak digunakan buku-buku yang berbasis korpus. Penggunaan kamus, misalnya, adalah hasil kerja yang didapat dengan

menggunakan software korpus yang harus meng-*up date* dari waktu ke waktu jumlah *entry*-nya seiring dengan perkembangan masyarakat, budaya, dan teknologi. Contoh lain adalah dalam pengajaran ESP (*English for Specific Purposes*). *Collins COBUILD Key Words in the Media*, misalnya, dirancang untuk membantu pelajar untuk lebih familiar dengan penggunaan Bahasa Inggris dalam media surat kabar, radio dan televisi. COBUILD juga menerbitkan *Business Vocabulary in Practice* untuk membantu memahami kosa kata dan penggunaan bahasa Inggris dalam konteks bisnis disamping banyak seri yang lainnya. Dalam korpusnya, COBUILD memiliki tidak kurang dari 500 juta kata yang disebut *Bank of English*.

Penggunaan korpus tidak hanya terbatas pada bahasa yang masih aktif digunakan. Bahasa yang sudah punah (*dead language*) pun, selama bukti tertulis masih tersedia, dapat direkonstruksi untuk melihat kosa kata, tata bahasa, susunan sintaksis serta informasi linguistic lainnya yang dapat menjadi cerminan struktur masyarakat pada zaman dimana bahasa tersebut masih digunakan. Namun, karena bahasa yang punah tidak lagi mungki berkembang, korpusnya pun dinamakan *closed corpus*. Hal ini telah dilakukan terhadap Bahasa Yunani dan Bahasa Inggris kuno. Tentu saja, informasi seperti ini sangat bermanfaat untuk pengajaran dan penelitian bahasa.

Dalam kaitan ini, merupakan suatu tantangan bagi mahasiswa yang belajar bahasa untuk dapat menggunakan korpus dalam penelitiannya. Mahasiswa dapat memilih topik-topik yang menarik untuk diteliti dengan menggunakan software korpus. Misalnya, mereka bisa meneliti suatu korpus lisan yang memuat ujaran-ujaran natural yang dapat member gambaran tentang berbagai hal berkaitan dengan genre (jenis teks), gender, umur, dan konteks social lainnya. Atau, penelitian tentang leksikon tertentu pada bahasa surat kabar dapat diambil dari laman suatu surat kabar tertentu dalam

jangka waktu yang cukup panjang. Atau, bisa pula dibuat korpus yang memuat tulisan-tulisan mahasiswa yang kemudian dianalisis penggunaan bahasanya dan dapat ditindaklanjuti untuk memberikan *feedback* kepada pengajaran bahasa.

Pertanyaan Penelitian

Secara spesifik, penelitian ini mengangkat dua isu yang berkaitan dengan mata kuliah Corpus Linguistics yang ditawarkan:

1. Bagaimana respons mahasiswa terhadap pentingnya mempelajari korpus?
2. Masalah atau tantangan apa yang dihadapi mahasiswa dalam melakukan penelitian dengan menggunakan *software concordance*?

Metodologi Penelitian

Penelitian ini menggunakan pendekatan kualitatif, dengan desain studi kasus yang bersifat deskriptif-analitis (Gay et al, 2006) yang bertujuan untuk mendapatkan pemahaman yang baik (*indepth understanding*) mengenai topik yang diteliti. Penelitian dilakukan dalam waktu dua bulan, 8 pertemuan dan melibatkan 68 mahasiswa.

Adapun instrument yang digunakan adalah observasi kelas untuk melihat pemahaman dan kualitas presentasi mahasiswa ketika mempresentasikan hasil penelitiannya, interview terhadap mahasiswa untuk mendalami pengetahuan mereka tentang korpus dan berdiskusi tentang penelitian mereka, serta analisis dokumen untuk menganalisa laporan penelitian mahasiswa yang menggunakan software korpus *concordance program*. Dalam laporan penelitiannya, mahasiswa ditugaskan untuk meneliti suatu topik atas pilihan mereka dengan jumlah kata minimal 250 ribu kata sebagai data korpusnya.

Hasil penelitian.

Dalam penelitian ini terungkap bahwa 44 orang mahasiswa (64,7%) menyatakan bahwa mereka tertarik dengan penggunaan korpus dalam penelitian dan pengajaran bahasa terutama karena software korpus sangat mudah digunakan dan sangat cepat prosesnya. Kecenderungan ini lebih dominan pada responden pria yang sangat mungkin berkaitan dengan tingkat ketertarikan dan pemahaman mereka terhadap komputer dan aplikasinya. Lebih jauh, ketika ditanyakan apakah mereka tertarik untuk menggunakan korpus software penelitian lain, hanya sebagian kecil (12 orang atau 17,6%) responden yang menyatakan bersedia. Hal ini sangat erat kaitannya dengan minat khusus mahasiswa yang berkaitan dengan linguistic terapan terutama yang berhubungan dengan CALL. Sebagian besar responden menyatakan lebih memilih bidang dan minat pada cabang ilmu linguistic lain seperti sociolinguistik dan pragmatik atau bidang lain di luar bidang linguistik seperti sastra dan penerjemahan.

Mengenai kesulitan dan hambatan dalam melakukan riset dengan menggunakan korpus software, sebagian besar mahasiswa menyatakan bahwa menganalisa data korpus cukup sulit karena pada dasarnya mereka harus mencoba "membaca" informasi linguistik itu dengan kerangka keilmuan yang lain yang dibutuhkan untuk menganalisa data. Informasi linguistic yang ditampilkan tidaklah cukup untuk mengambil suatu kesimpulan tentang suatu fenomena bahasa.

Terungkap pula bahwa kemampuan responden yang terbatas untuk memahami jenis data yang diperlukan dan menginterpretasi hasil dari korpus yang sudah dianotasi disebabkan oleh berbagai hal baik yang berhubungan dengan kemampuan akademis ataupun yang berhubungan dengan hal teknis. Kemampuan akademis responden yang

beragam menyebabkan hasil analisa tidak seluruhnya memuaskan. Mesin dan software korpus tidak akan banyak memberikan informasi jika tidak dibarengi dengan kemampuan analisis yang menyeluruh dan mendalam tentang fenomena bahasa yang sedang diteliti. Sedangkan secara teknis, responden banyak mengeluhkan sulitnya mendapat akses computer dan internet karena tidak semua memiliki computer dan akses ke internet.

Analisa terhadap laporan penelitian mereka menunjukkan bahwa sebagian mahasiswa (9 orang atau 13,2%) bahkan tidak dapat memenuhi jumlah kata dalam data mereka. Hal ini menjadi kendala dalam menentukan apakah jumlah data cukup merepresentasikan suatu fenomena bahasa yang dipilih. Adapun alasan utama atas masalah ini, menurut responden adalah tidak tersedianya akses computer dan internet yang memadai sehingga mereka kesulitan mendapat data yang cukup.

Terkait dengan kendala teknis yang secara langsung ataupun tidak langsung menyebabkan responden kesulitan mengakses data ataupun menganalisisnya, responden menuntut agar jurusan Bahasa Inggris harus menyediakan laboratorium computer yang memadai untuk memberikan semua mahasiswa kesempatan yang sama dalam mendapatkan latihan yang cukup dalam mengaplikasikan software korpus serta mengakses informasi terutama yang berhubungan dengan data elektronik.

Kesimpulan

Penggunaan Software korpus linguistic seperti software *concordance program* memberikan kemudahan dengan cara dan metode analisa yang baru yang dapat menjadi lahan penelitian dan bahan pengajaran yang menjanjikan. Jurusan Bahasa Inggris Universitas Pendidikan Indonesia telah mencoba untuk mengikuti perkembangan ini dan diharapkan dapat

dilahirkan linguis-lingui berbakat di masa depan. Hasil penelitian diatas menunjukkan bahwa setidaknya sebagian respoden memiliki minat dan bakat linguistic dan mata kuliah aplikasi seperti korpus linguistic ini sangat menarik bagi mereka. Namun, berdasarkan masukan dari mahasiswa, beberapa hal terutama tentang tersedianya fasilitas yang menunjang bagi keberhasilan mata kuliah ini harus diupayakan untuk diwujudkan.

Referensi

- Gay, L.R. et al.2006. *Educational Research: Competencies for Analysis and Application*. 8th Ed. Pearson Education Ltd. Ohio
- Mascull, Bill. 1995. *Collins Cobuild Key Words in the Media*. Harper Collins Publishser. UK
- McEnery, Tony and Wilson. 2007. *Corpus Linguistics*. Edinburgh University Press. Edinburgh
- Meyer, Charles.F. 2002. *English Corpus Linguistics: An Introduction*. CUP. London
- Milroy, Lesley. 1987. *Observing and Amalysing Natural Language*. Basil Blackwell Ltd. UK
- Robbins, Sue. 2004. | 2nd Ed. Harper Collins Publishser. UK

