

# Kolokasi Bahasa Inggris dalam Tugas Akhir Mahasiswa: Analisis Berbasis Frekuensi dengan Google sebagai Korpus

Tri Nuraniwati<sup>1</sup>, Ashfa F. Lathifah<sup>2</sup>

<sup>1</sup>*Prodi Bahasa Inggris Sekolah Vokasi Universitas Gadjah Mada*  
trinuraniwati@ugm.ac.id

<sup>2</sup>*Prodi Bahasa Inggris Sekolah Vokasi Universitas Gadjah*  
ashfaf1@gmail.com

## Abstract

This study aims to analyze the search results' frequency of collocations used by students of English Program, Vocational College UGM in their graduating papers (Tugas Akhir) and to show how Google can be used as a practical online corpus for collocations query. Collocations are words that co-occur together in any text and have fixed association. The use of collocations is part of linguistic awareness which determines how language is used in its natural state. Through corpus-based analyses, 1000 collocations are randomly chosen from 10 graduating papers (100 collocations per graduating paper). Using Google search engine, each collocation is entered as key words to find out the frequency of its search hits. The frequency is subsequently categorized into high frequency, medium frequency, and low frequency. From the 1000 collocations, 802 (80.2%) collocations belong to high-frequency collocations, 120 (12%) collocations are in the category of medium frequency, and 78 (7.8%) collocations are categorized as low-frequency collocations. The results indicate that students have shown good level of accuracy in using collocations in their academic writing.

**Key words:** graduating paper, collocations, frequency, Google

## Intisari

Penelitian ini bertujuan untuk menganalisis frekuensi hasil pencarian dari kolokasi yang dipakai mahasiswa Prodi Bahasa Inggris Sekolah Vokasi UGM dalam tugas akhir mereka serta menunjukkan bagaimana mesin pencari Google dapat dipakai sebagai korpus daring yang praktis untuk pencarian kolokasi. Kolokasi merupakan kumpulan kata yang cenderung dipakai secara bersama. Penggunaan kolokasi menunjukkan kesadaran metalinguistik di mana sebuah aspek kebahasaan dipakai secara alam oleh penuturnya. Dengan analisis berbasis korpus, 1000 kolokasi yang dipilih secara acak dari 10 tugas akhir (100 kolokasi per tugas akhir). Menggunakan mesin pencari Google, masing-masing kolokasi dimasukkan sebagai kata kunci untuk mengetahui jumlah kemunculannya sebagai hasil pencarian yang kemudian dikategorikan menjadi kolokasi berfrekuensi tinggi, menengah, dan rendah. Hasil yang didapat menunjukkan sebanyak 802 (80,25%) masuk kategori frekuensi tinggi, 120 (12%) frekuensi menengah, dan 78 (7.8%) frekuensi rendah. Hasil ini menunjukkan level akurasi penggunaan kolokasi yang bagus dari mahasiswa dalam tulisan akademik mereka.

**Kata kunci:** tugas akhir, kolokasi, frekuensi, Google

## Pendahuluan

Istilah *collocation* (kolokasi) dikenalkan pertama kali oleh Palmer, seorang pakar pengajaran bahasa Inggris pada tahun 1933, untuk menjelaskan sukseksi 2 kata atau lebih yang harus dipelajari sebagai kesatuan yang tak terpisah dan tidak bisa didefinisikan berdasarkan makna masing-masing komponennya (Kennedy, 2002). Istilah kolokasi kemudian dipopulerkan dalam konteks ilmu linguistik oleh Firth pada tahun 1957 untuk menjelaskan kombinasi kata yang mempunyai asosiasi tetap (Gabrielatos, 1994). Contoh yang terkenal untuk menjelaskan asosiasi ini adalah kolokasi ‘*to take a photo*’ (memfoto/mengambil foto) di mana kata kerja ‘*to take*’ tidak bisa digantikan oleh kata kerja lain semisal ‘*to make*’ atau ‘*to get*’ karena kata kerja ‘*to take*’ dan kata benda ‘*a photo*’ sudah memiliki asosiasi tetap.

Dari sudut pandang penutur asli, kemampuan menggunakan kolokasi merupakan aspek alamiah kebahasaan dan merupakan bagian dari kesadaran metalinguitik (*metalinguistic awareness*), sebuah frasa yang dipakai Koda untuk menjelaskan “*the knowledge of rules about*

*language, parts of language, and how language works, or even a simple self-recognition of one's own language and the forms being used*” (Shaw, 2011). Kesadaran metalinguitik muncul karena faktor kebiasaan dan intensitas paparan bahasa. Karena paparan yang terbatas, sebagian besar orang Indonesia yang masuk kategori penutur dan pembelajar bahasa Inggris sebagai bahasa asing (*English as a foreign language learners*) memerlukan usaha keras untuk mempelajari bentuk dan menghafalkan makna kolokasi sesuai konteks penggunaan baik dalam ragam lisan maupun tulis.

Penggunaan kolokasi dalam ragam lisan sama pentingnya dalam ragam tulis. Dalam ragam lisan, perkataan seseorang akan lebih mudah dipahami apabila penutur menggunakan kolokasi dengan tepat. Dalam ragam tulis, kelancaran penulisan secara gramatikal (tata bahasa) maupun leksikal (kosa kata) salah satunya dipengaruhi oleh apakah penulis menggunakan kolokasi yang tepat atau tidak. Seringkali, penulis yang dalam ini mahasiswa, mengacu pada istilah bahasa Indonesia, kemudian diterjemahkan secara literal ke bahasa Inggris yang

mengurangi tingkat idiomasi (*idiomacy*) kebahasaan sebuah tulisan.

May Fan (2009) dalam penelitiannya tentang pemakaian kolokasi oleh pembelajar bahasa Inggris menyimpulkan bahwa penutur asing bahasa Inggris menggunakan kolokasi dengan jumlah lebih sedikit daripada penutur asli dan mengalami kesulitan dalam menggunakan kolokasi secara tepat. Fan juga mendapati adanya pengaruh kombinasi kata pembentuk kolokasi dari bahasa ibu mereka ketika menggunakan Bahasa Inggris. Kondisi seperti ini menurut Fan harus diakomodasi oleh pengajar bahasa Inggris dengan pengajaran kosa kata yang meningkatkan kesadaran pembelajar bahasa akan adanya aspek kebahasaan yang idiomatik dan problematik.

Selain akomodasi pengajar dalam proses, pembelajar bahasa masa kini juga perlu dibekali dengan kemampuan pencarian kolokasi berbasis data (*data-driven learning*) yang sesuai dengan karakteristik mereka sebagai pribumi digital (*digital natives*) yang tumbuh dengan fasilitas teknologi dan cakap menggunakan teknologi baik untuk gaya hidup maupun proses pembelajaran.

*Digital natives* merupakan istilah yang dipakai pertama kali oleh Marc Pensky pada tahun 2001 untuk generasi pembelajar yang sejak kecil sudah terpapar teknologi digital dan menjadikan teknologi sebagai bagian tak terpisahkan dari kehidupan mereka. Dalam menghadapi tipe pembelajar seperti ini, dosen maupun pengajar pada umumnya tidak boleh skeptis dan mau tidak mau harus menjadi imigran digital (*digital immigrants*) untuk bisa menjadi fasilitator yang baik (Pensky, 2011).

Bagi pribumi digital, kegiatan membuka kamus khusus kolokasi seperti *Oxford Collocations Dictionary* ataupun bertanya kepada penutur asli sudah tidak lagi populer. Data yang sangat besar ukurannya yang tersedia secara daring merupakan kamus praktis bagi pribumi digital. Dengan bantuan mesin pencari Google, mereka dengan mudah mencari tahu makna dan konteks penggunaan suatu kolokasi.

Agar selaras dengan gaya belajar mahasiswa, pengajar bisa mengajak mereka memanfaatkan mesin pencari Google sebagai salah satu alternatif untuk membantu mencari tahu apakah kolokasi yang mereka pakai sudah tepat atau belum.

Dengan basis data yang sangat besar, mesin pencari Google dianggap sebagai sebuah korpus daring untuk pembelajaran kosa kata Bahasa Inggris.

Istilah korpus (jamak: korpora) menurut Leech (McEnery, dkk, 2006) mengacu pada koleksi materi tekstual yang disusun dengan tujuan tertentu. Selain yang berwujud piranti lunak, korpora juga tersedia secara daring yang bisa diakses secara gratis dengan ukuran yang sangat besar seperti British National Corpus (BNC), Corpus of Contemporary American English (COCA), dan Michigan Corpus of American Spoken English (MICASE). Korpora tersebut telah cukup lama dikembangkan dan merupakan korpora yang sangat mapan sebagai sumber pembelajaran dan penelitian bahasa, baik lisan maupun tulis.

Walaupun dikategorikan mapan dan lengkap, korpora seperti BNC atau COCA kurang familiar dikalangan mahasiswa. Fakta bahwa sumber teks BNC dan COCA berasal dari penutur asli Bahasa Inggris juga bisa menyurutkan nyali mahasiswa dalam belajar kosa kata. Oleh karena itu dalam penelitian ini dipilih Google yang dianggap sebagai mesin pencari sekaligus korpus daring yang

sangat dikenal di kalangan mahasiswa dan dapat diakses sewaktu-waktu.

Dewasa ini mahasiswa sangat tergantung dengan mesin pencari ini sehingga kata 'Google' sekarang identik dengan kata 'mencari.' Kita cukup sering mendengar kalimat "Ya nanti saya Google dulu" atau "Coba di-Google saja" dalam pembelajaran di kelas maupun di luar kelas. Selain tingkat kemutakhiran basis data, kelebihan Google yang lain adalah kemampuannya menampilkan semua aspek kebahasaan dari level kata hingga kalimat yang sangat kompleks, termasuk menampilkan data kolokasi. Dalam hitungan detik, Google mampu menampilkan ratusan juta hasil pencarian. Walaupun tidak semua sumber teks yang masuk ke basis data Google berasal dari sumber yang valid secara gramatika kebahasaan (karena Google juga menampilkan tulisan penutur asing bahasa Inggris), jumlah hasil pencarian atau frekuensi kemunculan kolokasi bisa menjadi parameter apakah sebuah kolokasi diterima secara luas dan dianggap tepat. Dalam penelitian berbasis korpus, frekuensi merupakan aspek yang penting, khususnya dalam pembahasan aspek leksikal

Dengan menggunakan mesin pencari Google penelitian ini menganalisis apakah kolokasi-kolokasi yang dipakai mahasiswa dalam makalah tugas akhir mereka sudah masuk dalam kategori kolokasi yang berfrekuensi tinggi atau belum. Penelitian ini juga menunjukkan cara bagaimana mahasiswa sebagai penutur asing Bahasa Inggris bisa memanfaatkan Google untuk pembelajaran kosa kata Bahasa Inggris khususnya untuk mempelajari kolokasi.

## **Metode Penelitian**

Data dari penelitian ini adalah 1000 kolokasi yang dipilih secara acak dari 10 makalah tugas akhir mahasiswa Prodi Bahasa Inggris Sekolah Vokasi UGM angkatan 2010 dan 2011. Nama penyusun tugas akhir dan judulnya tidak akan disebutkan dalam penelitian ini. Masing-masing tugas akhir akan dinamai dengan P1, P2, dan seterusnya sampai P10.

Tugas akhir merupakan karya tulis mahasiswa Prodi Bahasa Inggris yang wajib ditulis dan diujikan sebagai syarat kelulusan. Tugas akhir bisa berwujud laporan observasi ataupun laporan analitik dari praktik kerja lapangan yang dilakukan mahasiswa pada

semester akhir perkuliahan. Tugas akhir umumnya ditulis dalam Bahasa Inggris dengan panjang antara 6000 sampai 10.000 kata yang wajib ditulis mahasiswa sebagai syarat kelulusan. Walaupun sudah melewati proses supervisi dari dosen pembimbing, tulisan mahasiswa tersebut tentu saja tidak sempurna dari aspek gaya, mekanika, dan struktur penulisan. Mahasiswa dengan tingkat kemampuan kebahasaan yang berbeda akan menghasilkan tulisan dengan level gramatikal dan leksikal yang berbeda. Penggunaan aspek idiomatis bahasa khususnya bagaimana mereka menggunakan kolokasi juga akan berbeda level akurasinya.

Dari masing-masing tugas akhir diambil 100 kolokasi secara acak dengan tidak dibatasi bentuk leksikal maupun gramatikalnya. Jumlah kata penyusun per kolokasi juga tidak dibatasi. Ada kolokasi yang tersusun atas 2 kata dan ada yang terbentuk dari 4 kata penyusun.

Penelitian ini menggunakan metode frekuensi berbasis korpus dengan menggunakan mesin pencari Google sebagai korpus daring yang dipadukan dengan metode *data-driven learning*. Metode frekuensi berbasis korpus merupakan salah

satu metode yang dipakai dalam ilmu linguistik korpus (*corpus linguistics*). Cabang dari ilmu linguistik ini mempelajari bagaimana analisis kebahasaan diterapkan untuk kumpulan teks yang menyusun suatu korpus (Meyer, 2002). Metode berbasis frekuensi yang dipakai dalam *corpus linguistics* terkait erat dengan *metalinguistic awareness* yang sudah disebutkan dalam bagian sebelumnya. Frekuensi tinggi bisa mencerminkan akurasi suatu aspek kebahasaan oleh penutur asli yang polanya bisa ditiru oleh penutur asing. Kosakata berfrekuensi tinggi dalam sebuah korpus bisa membantu mahasiswa untuk memahami makna dan penggunaan kata-kata Bahasa Inggris tertentu secara mendalam untuk keperluan reseptif dan produktif mereka (Romer, 2008).

Terkait *data-driven learning*, metode pembelajaran yang digagas oleh Johns di tahun 1994 ini dipakai untuk memberikan gambaran kepada mahasiswa bagaimana fenomena kebahasaan dapat dilihat dalam konteks yang lebih nyata melalui korpora yang sumber teksnya berasal dari penutur Bahasa Inggris, bukan dari kamus ataupun buku yang ditulis para ahli yang bersifat didaktif (Meyer, 2002). Hal ini sesuai dengan

pernyataan Anderson dan Corbett (2009) “*The careful analysis of corpora can give insights into (i) how language is really used, rather than how people think it is used and (ii) how it is commonly and typically used.*”

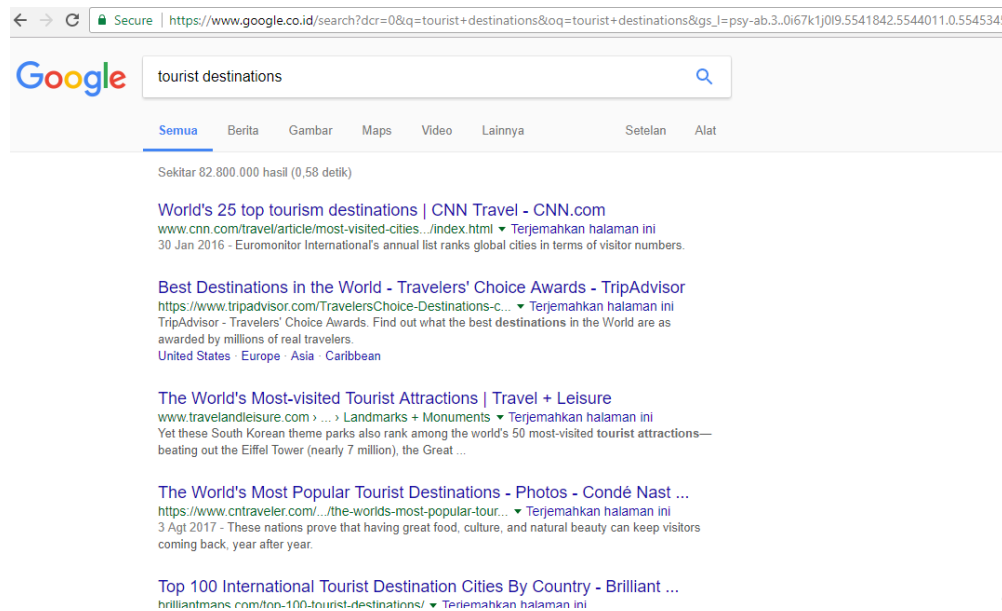
Pembelajaran berbasis korpus berbeda dengan pembelajaran melalui buku teks maupun kamus yang cenderung menunjukkan bagaimana seharusnya aspek kebahasaan dipakai, bukan bagaimana pada umumnya aspek tersebut dipakai oleh penggunanya.

Data berupa 1000 kolokasi dipilih dari 10 tugas akhir yang dipilih secara acak dan kemudian dimasukkan satu per satu ke mesin pencari Google untuk dicari jumlah hasil pencariannya. Pencarian tidak menggunakan metode khusus seperti penggunaan metode Boolean operator (penggunaan AND, OR, NOT or AND NOT) maupun penggunaan tanda kutipan langsung (“...”) untuk membatasi hasil pencarian. Kolokasi diketik secara langsung di laman pencarian Google dan dicatat jumlah frekuensi kemunculannya (*search hits*). Walaupun tidak sulit, proses ini memerlukan waktu yang cukup lama karena terdapat 1000 kolokasi yang harus dicari frekuensinya. Kecepatan mendapatkan data

frekuensi juga sangat tergantung dengan kecepatan akses internet.

Seperti kita ketahui, Google melakukan pencarian sebanyak mungkin terhadap jutaan situs yang terdapat dalam *world wide web* yang jumlahnya akan berkurang dan bertambah sewaktu-waktu, maka frekuensi hasil pencarian untuk kolokasi tertentu juga akan berbeda setiap saat.

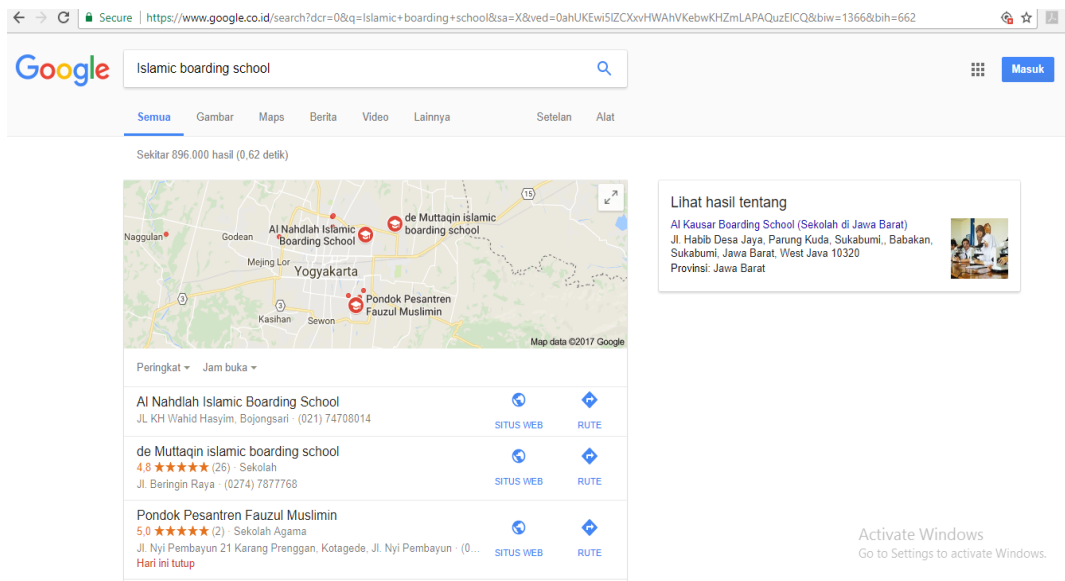
Data jumlah hasil pencarian dalam penelitian ini diambil pada bulan Februari 2017 dan contohnya dapat dilihat di Gambar 1 berikut ini. Gambar 1 merupakan hasil tangkapan layar dari hasil pencarian untuk kolokasi *tourist destinations* yang berasal dari tugas akhir 1 (P1):



Gambar 1. Hasil Tangkapan layar untuk kolokasi 'tourist destinations'

Dalam waktu 0,58 detik, Google mampu menampilkan 82.800.000 hasil pencarian untuk kolokasi tersebut.

Contoh lain bisa dilihat di gambar 2 untuk pencarian kolokasi *Islamic boarding school* berikut ini:



Gambar 2. Hasil tangkapan layar untuk kolokasi 'Islamic boarding school'

Dalam waktu 0,62 detik, Google mampu menampilkan 896.000 hasil pencarian.

Hasil pencarian yang berupa frekuensi kemunculan dari 1000 kolokasi yang digunakan sebagai data penelitian ini selanjutnya diklasifikasikan menjadi tiga kategori: frekuensi tinggi (*high frequency*), frekuensi menengah (*medium frequency*), dan frekuensi rendah (*low frequency*). Kategori frekuensi tinggi mengacu pada

jumlah hasil pencarian di atas 10 juta. Untuk frekuensi menengah, hasil pencarian berkisar antara 1 sampai 10 juta. frekuensi rendah memiliki jumlah hasil pencarian kurang dari 1 juta. Kategorisasi ini dibuat berdasarkan pola pencarian umum yang ada di mesin pencari Google.

Tabel 1 berikut ini memuat contoh jumlah frekuensi hasil pencarian untuk 10 kolokasi yang diambil dari tugas akhir 1 (P1):

Table 1. Contoh hasil pencarian kolokasi tugas akhir 1 (P1)

No.	P1	Jumlah Hasil Pencarian Google
1	tourist destinations	82.800.000
2	religious tourism	1.770.000
3	culinary tourism	2.350.000
4	handicraft tourism	500.000



5	educational tourism	1.960.000
6	overall area	839.000.000
7	land area	62.600.000
8	service point	51.300.000
9	natural attraction	116.000.000
10	tourist view	119.000.000

Sumber: Data Primer

Dari tabel 1 terlihat kolokasi *tourist destinations, overall area, natural attraction, land area, service point, dan tourist view* masuk kategori kolokasi berfrekuensi tinggi. *Religious tourism, culinary tourism, dan educational tourism* termasuk kolokasi kategori frekuensi menengah. Sedangkan *handicraft tourism* masuk ke kategori kolokasi dengan frekuensi rendah. Contoh lain dari frekuensi hasil pencarian untuk 10 kolokasi dalam tugas akhir 5 (P5) dapat dilihat di tabel 2 berikut ini:

Tabel 2. Contoh hasil pencarian kolokasi tugas akhir 5 (P5)

No.	P5	Jumlah Hasil Pencarian Google
1	brief profile	174.000.000
2	modern Islamic boarding school	<b>480.000</b>
3	story of establishment	282.000.000
4	Islamic boarding school	<b>896.000</b>
5	boarding school	18.400.000
6	greatly indebted	3.150.000
7	royal title	322.000.000
8	fourth generation	6.990.000
9	three kilometers away	25.900.000
10	forested area	8.420.000

Sumber: Data Primer

Dari tabel 2 bisa dilihat kolokasi *brief profile, story of establishment, boarding school, royal title, dan three kilometers away* masuk kategori frekuensi tinggi. *Greatly indebted, fourth generation, dan forested area* masuk kategori frekuensi menengah, sedangkan

*modern Islamic boarding school* dan *Islamic boarding school* berada pada kategori kolokasi berfrekuensi rendah. Contoh hasil pencarian yang lebih lengkap untuk tugas akhir 1 (P1) sampai tugas akhir 10 (P10) dapat dilihat di lampiran akhir penelitian ini.

## Hasil Dan Pembahasan

Setelah penelusuran jumlah hasil pencarian masing-masing kolokasi

didapatkan dan dikategorikan level frekuensinya, peneliti selanjutnya menghitung berapa jumlah kolokasi berfrekuensi tinggi, menengah, dan rendah dari masing-masing tugas akhir yang menjadi sampel. Hasil pencarian frekuensi kolokasi yang digunakan mahasiswa dalam tugas akhir mereka dapat diringkas dalam tabel 3 berikut ini:

Tabel 3. Jumlah hasil pencarian kolokasi

FREKUENSI	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	TOTAL	%
Frekuensi tinggi	80	76	83	75	74	81	86	87	81	79	802	80,2
Frekuensi menengah	13	10	7	18	16	13	8	11	11	13	120	12
Frekuensi rendah	7	14	10	7	10	6	6	2	8	8	78	7,8

Sumber: Data Primer

Dari 1000 kolokasi yang dianalisis, 80,2 % (802 kolokasi) masuk kategori kolokasi dengan frekuensi tinggi; 12 % (120 kolokasi) dikategorikan sebagai kolokasi berfrekuensi menengah; dan 7,8 % (78 kolokasi) masuk kategori kolokasi berfrekuensi rendah. Masing-masing tugas akhir menunjukkan

penggunaan frekuensi tinggi dengan persentase lebih dari 70 %. Frekuensi tinggi dengan jumlah paling rendah ditunjukkan oleh tugas akhir 2 (P5) sejumlah 74 kolokasi.

Yang perlu diperhatikan adalah daftar kolokasi dengan frekuensi rendah yang dapat dilihat di tabel 4 berikut:

Tabel 4. Daftar kolokasi dengan kategori *low frequency*

<i>Low-frequency collocations</i>									
P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
handicraft tourism	several healers	spirited leadership	talented businessmen	modern Islamic boarding school	reign of Han dynasty	sit calmly	city of gudeg	customer buying habits	ritual sura tradition
favorite tourist destination	miraculously recovered	administratively managed	distributor of consumer electronics	Islamic boarding school	gift shop noodles	broadcasted publicity	indeed purposefully	serat alam home industry	heirloom carrier
souvenir markets	swallows nests hunters	interdisciplinary and condensed research	vendor neutral	survival of Gontor	medium-class restaurant	English news program		handicraft products	ritual procession
perusal of the history	biggest producer	condensed research	optimally comfortable indoor environment	plaited mats	other stalls	commemoration speech		natural fiber crafter	different golden eras
tourism object gate	swallows nests cultivation	institutional cooperation	water chillers	Indonesian personality	noodle stall	VHF transmitter		unavailability of showroom	supernatural and magical phenomena
optimum management	rattled sound	audio-lingual classes	variable refrigerant volume	Islamic brotherhood	wet noodle			batik materials	lineage of kingdom
visitor annularity	strong Javanese culture	phonetics lab	variable refrigerant flow					household ornaments	genie palace
	Javanese poem	Javanese instrumental	routine activities and programs					batik necklace	
	swallow nest	graduate designations	Waqf maintenance and enlargement						
	swallow bird habit	unstructured curriculum	waqf land						
	stiff wing		well qualified Muslim leader						
	chalk cave								
	spacious tail								
	loud and rattled sound								

Sumber: Data Primer

Kolokasi dengan kategori *low frequency* mengindikasikan beberapa hal:

1. Adanya faktor penerjemahan literal yang mempengaruhi bagaimana mahasiswa menyusun kolokasi secara literal kata per kata dalam kalimat-kalimat mereka. Sebagai contoh kolokasi *chalk cave* yang dipakai di tugas akhir 2 (P2). Mahasiswa terpengaruh dengan istilah 'goa kapur' dalam bahasa Indonesia. Kata 'kapur' harusnya diterjemahkan menjadi 'limestone' karena kata 'chalk' mengacu pada 'kapur' yang dipakai untuk menulis.
2. Adanya istilah lokal yang sudah terserap Bahasa Inggris namun belum banyak dipakai karena merupakan *culturally-bound words* seperti kolokasi *Islamic boarding school* yang dalam bahasa Indonesia sering disebut sebagai pesantren dan *batik necklace* (kalung batik) yang keberadaannya dalam basis data Google tidak banyak.

## Kesimpulan

Dengan presentase kolokasi berfrekuensi tinggi sebanyak 80,2%, hasil penelitian ini menunjukkan bahwa mahasiswa penulis tugas akhir di Prodi Bahasa Inggris Sekolah Vokasi telah menunjukkan kemampuan penggunaan kolokasi yang bagus dalam tulisan akademik mereka, yang berarti kemampuan reseptif dan produktif mereka dalam menggunakan kosa kata akademik bahasa Inggris sudah cukup mumpuni.

Bagi mahasiswa yang masih menjalani studi, fasilitasi penggunaan korpus daring harus terus dikembangkan oleh para dosen untuk memperbaiki penguasaan kosa kata mereka. Penggunaan Google sebagai korpus daring dalam penelitian ini terbukti efektif dan dapat diadopsi oleh mahasiswa untuk membantu memastikan level akurasi dan keberterimaan penggunaan kolokasi tertentu. Menurut Fletcher (2002), sumber daring termasuk di dalamnya korpora memiliki keuntungan dalam hal kemutakhiran, multimodalitas, dan ketersediaan data dengan jumlah sangat besar namun dengan biaya minimal. Untuk mesin pencari Google, konten web yang

selalu bertambah setiap saat menambah kemutakhiran basis data Google sebagai mesin pencari. Bentuk data yang beragam

juga mengakomodasi proses pembelajaran kosa kata yang bervariasi.

## Daftar Pustaka

- Anderson W. & J. Corbett. (2009). *Exploring English with Online Corpora: An Introduction*. London: Palgrave.
- Fan, May. (2009). An Exploratory study of collocational use by ESL students—A task based approach. *System*, 37, 110-123.
- Fletcher, William. (2002). Making the Web More Useful as a Source for Linguistic Corpora. *Corpus Linguistics in North America 2002: Selections from the Fourth North American Symposium of the American Association for Applied Corpus Linguistics*.
- <http://kwicfinder.com/AAACL2002whf.pdf> Accessed 16.04.13
- Gabrielatos, Constatinos. (1994). Collocations: Pedagogical implications and their treatment in pedagogical materials. Unpublished essay. Research Centre for English and Applied Linguistics. University of Cambridge.
- Kennedy, Graeme. (2003). Amplifier Collocations in the British National Corpus: Implications for English Language Teaching. *TESOL Quarterly* 37 (3), 467-477.
- McEnery, Tony et.al. (2006). *Corpus-Based Language Studies: An Advanced Resource Book*. London: Routledge.
- Meyer, Charles F. (2002). *English Corpus Linguistics: An Introduction*. Cambridge: CUP.
- Prensky, Marc. (2012) From Digital Natives to Digital Wisdom: Hopeful Essays for 21<sup>st</sup> Century Education. [http://marcprensky.com/writing/Prensky-Intro\\_to\\_From\\_DN\\_to\\_DW.pdf](http://marcprensky.com/writing/Prensky-Intro_to_From_DN_to_DW.pdf)
- Romer, U. (2008) Corpora and language teaching. In A. Ludeling & M. Kyoto (Eds.), *Corpus linguistics: An international handbook* (Vol. 1, pp. 112-130). Berlin: Mouton de Gruyter.
- Shaw, Erin. (2011). *Teaching Vocabulary Through Data-driven Learning*. Idaho: BYU.

## Lampiran 1

### Daftar frekuensi hasil pencarian P1-P3

No.	P1	Jumlah Hasil Pencarian Google	P2	Jumlah Hasil Pencarian Google	P3	Jumlah Hasil Pencarian Google
1	large number	322.000.000	natural resources	438.000.000	six times	31.900.000
2	tourist destination	8.790.000	tourism places	338.000.000	earliest state university	13.600.000
3	tourist destinations	90.000.000	good potentials	24.800.000	new faculties	21.000.000
4	main tourist destination	52.200.000	good potentials place	825.000.000	most recent development	1.050.000.000
5	capital city	64.600.000	different kinds of business	242.000.000	diploma program	167.000.000
6	historical place	629.000.000	special gifts	230.000.000	eight study programs	214.000.000
7	natural atmosphere	247.000.000	culinary tourism packages	12.200.000	Indonesian culture	29.400.000
8	various thing	480.000.000	home furniture	360.000.000	language learning program	540.000.000
9	tourism potential	108.000.000	economic and cultural significance	110.000.000	foreign students	16.200.000
10	historical tourism	106.000.000	several healers	642.000	another faculty	255.000.000
11	nature tourism	161.000.000	big prize	84.400.000	educational institution	234.000.000
12	cultural tourism	179.000.000	next day	369.000.000	cultural research center	12.000.000
13	religious tourism	1.770.000	miraculously recovered	514.000	International level	1.390.000.000
14	culinary tourism	2.350.000	very healthy	484.000.000	interests of humanity	44.400.000
15	handicraft tourism	500.000	very unique	848.000.000	higher education	915.000.000
16	educational tourism	1.960.000	very well-protected	16.400.000	high quality	904.000.000
17	overall area	839.000.000	swallows nests hunters	523.000	spirited leadership	580.000
18	land area	62.600.000	biggest producer	404.000	cultural sciences	3.090.000
19	things we know	313.000.000	some rituals	44.200.000	progress of education	677.000.000

20	natural attraction	116.000.000	some food	1.470.000.000	national interests and humanity	2.390.000
21	beautiful caves	42.800.000	potentials and cultural values	180.000.000	culture community	953.000.000
22	located approximately	11.300.000	research questions	1.150.000.000	various parties	499.000.000
23	southern village	3.220.000	cultural values	253.000.000	organization and faculty management	136.000.000
24	tourist views	119.000.000	swallows nests cultivation	<b>471.000</b>	educational purposes	409.000.000
25	sacred places	8.890.000	citizens point of view	17.200.000	high integrity	20.200.000
26	tourism department	201.000.000	qualitative and quantitative analysis	2.560.000	changes and progress of knowledge	307.000.000
27	government business charge	516.000.000	certain methods	790.000.000	community problems	769.000.000
28	government business	1.460.000.000	direct observation	10.300.000	knowledge and technology skills	293.000.000
29	culture promotion	320.000.000	rattled sound	<b>498.000</b>	productive activities	110.000.000
30	increasing visitors	103.000.000	additional information	592.000.000	community services	150.000.000
31	development strategies	347.000.000	more detailed information	211.000.000	basic sciences	31.200.000
32	favorite tourist destination	<b>30.400</b>	four chapters	120.000.000	quality of learning process	475.000.000
33	favorite tourism destination	25.300.000	study objectives	232.000.000	learning activities	772.000.000
34	tourism subject	224.000.000	next chapter	39.200.000	administratively managed	<b>437.000</b>
35	completing facility	106.000.000	each species	297.000.000	name management	1.290.000.000
36	more attractive	53.000.000	usually added	629.000.000	university compulsory courses	27.400.000
37	research methods	18.000.000	sign of a place	1.550.000.000	faculty compulsory courses	15.900.000
38	library research	1.210.000.000	strong Javanese cul ture	<b>145.000</b>	each program of study	1.010.000.000
39	accurate database	188.000.000	other region	1.560.000.000	student lectures schedule	99.600.000

40	specific information	1.160.000.000	cultural things	387.000.000	courses of study programs	310.000.000
41	real fact	877.000.000	traditional dances	24.200.000	research funding	276.000.000
42	second chapter	30.100.000	Javanese poem	<b>281.000</b>	various types and qualifications	87.100.000
43	third chapter	446.000.000	administrative logo	404.000.000	research program implementation	7.940.000
44	last chapter	667.000.000	region logo	1.040.000.000	competition system	403.000.000
45	research objectives	236.000.000	capital income	13.000.000	great topic	427.000.000
46	local government	374.000.000	strong bond	211.000.000	interdisciplinary and condescend research	<b>101.000</b>
47	lodging facilities	141.000.000	each other	975.000.000	condescend research	938.000
48	souvenir markets	557.000	aerial bird	27.500.000	implementation of this system	326.000.000
49	certain days	948.000.000	dark black color	317.000.000	other research	1.960.000.000
50	established procedures	358.000.000	fly fast	136.000.000	institutional cooperation	886.000
51	perusal of the history	537.000	medium body size	287.000.000	staff researcher	99.300.000
52	major tourist visits	97.700.000	ceiling of house	130.000.000	study centers	25.000.000
53	tourist visits	67.100.000	good minerals	23.600.000	other studies	781.000.000
54	horse-drawn carriage	1.140.000	six kinds	152.000.000	individual cooperation	231.000.000
55	art potency	16.400.000	swallow nest	671.000	preparatory institution	24.400.000
56	promotional program	5.000.000	swallow bird habit	740.000	largely documented	83.600.000
57	natural potency	18.100.000	many people	2.300.000.000	research activities	2.640.000.000
58	great waves	5.030.000	body size	924.000.000	last three years	354.000.000
59	favorite thing	26.700.000	feathers color	17.300.000	community service activities	61.000.000
60	cultural value	451.000,000	steep cliff	11.900.000	more study programs	754.000.000
61	increasing number	376.000.000	small insects	75.500.000	community service programs	105.000.000



62	tourism object gate	<b>429.000</b>	moist place	53.700.000	qualified proposals	61.100.000
63	service point	51.300.000	second similarity	45.300.000	community service funding	35.800.000
64	great influence	445.000.000	all swallow	59.300.000	cost of implementation	324.000.000
65	safety point	962.000.000	more number	3.850.000.000	field of study	1.160.000.000
66	appropriate tourism	88.600.000	same tasks	233.000.000	relevant program of study	502.000.000
67	visitor origin	47.200.000	similar body shape	4.920.000	synergy of cooperation	7.450.000
68	tourist interests	32.800.000	good colony	19.300.000	community service program interdisciplinary	2.220.000
69	optimum management	922.000	strong claw	4.520.000	some facilities	600.000.000
70	closely related	82.300.000	house ceiling	159.000.000	books collection	639.000.000
71	promotional strategy	1.160.000	good environment	944.000.000	several computers	141.000.000
72	main tourism destination	71.600.000	next generation	233.000.000	language lab	17.200.000
73	tourism area	573.000.000	white color nest	3.550.000	audio-lingual classes	<b>101.000</b>
74	continuing efforts	17.600.000	white nest	27.200.000	student capacity	323.000.000
75	tourism products	9.850.000	white swallow bird	8.110.000	computer lab	15.900.000
76	promotion staff	480.000.000	brown and black feather	3.700.000	computer system	1.260.000.000
77	operational phases	48.700.000	grey and brown feather	5.410.000	phonetics lab	<b>391.000</b>
78	promotion activities	480.000.000	high voice	433.000.000	Javanese instrumental	<b>251.000</b>
79	various promotion activities	167.000.000	dark brown eyes	13.400.000	photocopy service	21.900.000
80	relevant colleges	116.000.000	black beak and foot	4.140.000	education process	1.230.000.000
81	various exhibition	122.000.000	stiff wing	672.000	four semester	135.000.000
82	national level	1.580.000.000	fly strongly	59.900.000	minimum credit	640.000.000
83	promotion activity	268.000.000	low place	59.700.000	maximum time limit	38.600.000
84	first businesses	690.000.000	tall tree	20.900.000	study period	17.400.000
85	before doing	1.150.000.000	reef space	35.800.000	second semester	6.430.000

chalk cave	developing promotional activities	32.300.000		748.000	other academic activities	25.900.000
87	promotion strategy	3.030.000	white eggs	26.000.000	thesis supervisor	20.400.000
88	target audience	19.800.000	long shape	422.000.000	scholarship studies	158.000.000
89	second step	28.300.000	seasonal bird	32.900.000	study extension	376.000.000
90	positive impact	842.000.000	certain season	175.000.000	graduate designations	533.000
91	visitor annularity	514.000	second kinds	353.000.000	master program	730.000.000
92	next step	215.000.000	dark black feather	6.100.000	doctorate program	24.300.000
93	promotional messages	32.400.000	brown feather	20.400.000	interest study	856.000.000
94	general action	1.610.000.000	spacious tail	619.000	unstructured curriculum	<b>455.000</b>
95	prospective consumers	41.200.000	loud and rattled sound	624.000	structured curriculum	42.000.000
96	loyal customers	46.300.000	biggest species	78.400.000	master program life cycle	10.800.000
97	high involvement	3.870.000	fly faster ad higher	10.500.000	internship report paper	1.560.000
98	many times	311.000.000	small cave	81.400.000	graduating paper	28.400.000
99	promotional budget	869.000.000	small gap	272.000.000	job descriptions	19.000.000
100	appropriate media	596.000.000	bowl-shaped nests	1.350.000	academic administration staff	364.000.000

Sumber: Data Pribadi

## Lampiran 2

### Daftar frekuensi hasil pencarian P4-P6

No.	P4	Jumlah Hasil Pencarian Google	P5	Jumlah Hasil Pencarian Google	P6	Jumlah Hasil Pencarian Google
1	economic activity	14.900.000	brief profile	174.000.000	main staple foods	6.170.000
2	production and consumption activities	144.000.000	modern islamic boarding school	<b>480.000</b>	Chinese influence	155.000.000
3	usefulness of goods and services	614.000.000	story of establishment	282.000.000	many varieties	72.700.000
4	everyday life	104.000.000	islamic boarding school	893.000	chicken noodle	4.590.000
5	good distribution channels	5.030.000	boarding school	18.400.000	Chinese descent	4.050.000
6	availability of products	617.000.000	greatly indebted	3.150.000	historical records	236.000.000
7	production and distribution activities	312.000.000	royal title	322.000.000	mainland China	11.200.000
8	various factors of production	45.400.000	fourth generation	6.990.000	reign of Han dynasty	793.000
9	good distribution activities	5.850.000	three kilometers away	25.900.000	European continent	60.000.000
10	distribution company	26.400.000	forested area	8.420.000	continental Europe	7.120.000
11	distribution process	741.000.000	early days	54.400.000	gift shop noodles	837.000
12	group of persons	412.000.000	number of students	1.060.000.000	well-known	372.000.000
13	other entities	43.900.000	dedicated students	392.000.000	street food	72.100.000
14	economic needs	643.000.000	survival of Gontor	<b>36.000</b>	small restaurant	470.000.000
15	human beings	58.300.000	surrounding villages	47.900.000	medium-class restaurant	958.000
16	existence of companies	168.000.000	previously very religious	281.000.000	high-class restaurant	10.600.000
17	key players	220.000.000	bad thing	515.000.000	much money	993.000.000

18	Indonesian economy	19.400.000	thought hard	535.000.000	highly recommended	377.000.000
19	community members	1.320.000.000	worked really hard	328.000.000	original and delicious taste	58.200.000
20	new order	94.500.000	fifth child	5.000.000	other mie ayam stalls	<b>81.700</b>
21	grassroots communities	30.600.000	other educational institution	226.000.000	best weapons	21.300.000
22	marketing activities	348.000.000	Islamic calendar	8.290.000	increased enormously	13.300.000
23	necessary goods	351.000.000	assembled audience	18.000.000	city center	284.000.000
24	works effectively and efficiently	73.200.000	three brothers	17.500.000	index success of business	202.000.000
25	distribution of goods	206.000.000	first step	249.000.000	affordable price	361.000.000
26	number of important tasks	241.000.000	child educational program	368.000.000	strategic areas	307.000.000
27	critical information	23.400.000	coconut shells	2.140.000	several places	545.000.000
28	planning and launching exchange	67.600.000	plaited mats	<b>188.000</b>	residential areas	395.000.000
29	persuasive information	43.600.000	later developed	668.000.000	boarding house areas	1.510.000
30	service offers	1.290.000.000	many branch	460.000.000	seven branches	87.200.000
31	buyer demands	19.700.000	many regions	328.000.000	good thing	1.010.000.000
32	final agreement	18.100.000	first building	1.610.000.000	one branch	483.000.000
33	transfer of ownership	225.000.000	each boarding school	2.170.000	unique factors	355.000.000
34	physical distribution	4.620.000	some lists	481.000.000	one of popular mie ayam sellers	7.690.000
35	storage of goods	210.000.000	education institution	428.000.000	uniqueness of the products	23.500.000
36	funding requests	10.100.000	leader cadres	1.280.000	eyes of the customers	76.700.000
37	marketing channel	9.850.000	source of Islamic sciences	2.830.000	improvements of the products and services	228.000.000
38	duct work	6.270.000	language of Al-Qur'an	2.360.000	more specific data	1.200.000.000
39	five first tasks	180.000.000	sciences knowledge	493.000.000	three main branches	83.400.000
40	completion of the transaction	28.100.000	religious character	109.000.000	several customers	432.000.000

41	company overview	499.000.000	excellent graduates	71.400.000	main factors	441.000.000
42	talented businessmen	<b>444.000</b>	new generation	120.000.000	improvement and development of the products	221.000.000
43	specialized solutions technology	1.550.000	noble character	1.370.000	some information	5.570.000.000
44	industries and business functions	7.900.000	sound body	19.000.000	following methods	51.100.000
45	distributor of consumer electronics	675.000	broad knowledge	13.400.000	direct observation	9.390.000
46	system integrator	2.830.000	independent mind	4.850.000	valid data	622.000.000
47	reliable experience	350.000.000	science religion	228.000.000	main data	1.430.000.000
48	comprehensive capabilities	2.250.000	another science	811.000.000	some photographs	275.000.000
49	high performance	313.000.000	teach proportionally	1.580.000	additional data	2.060.000.000
50	technical team	524.000.000	intellectual Moslem	16.300.000	operational manager	13.400.000
51	project cost	671.000.000	Indonesian personality	740.000	company history and products	504.000.000
52	every organization	807.000.000	five spirit	311.000.000	related books	589.000.000
53	vendor neutral	981.000	boarding system	51.200.000	related products	372.000.000
54	selection of technology solutions	7.060.000	Islamic brotherhood	988.000	objectives of the study	237.000.000
55	audio-visual and multimedia system	1.460.000	organization structure	37.100.000	methods of collecting data	105.000.000
56	refrigeration system	3.120.000	highest position	608.000.000	brief information	2.560.000.000
57	true integrated ordering	64.600.000	boarding guidance	11.800.000	company profile	749.000.000
58	installation of electronic equipment	97.400.000	some institutions	272.000.000	organizational structure	24.000.000
59	all-in-one solution	247.000.000	widely known	14.500.000	three different branches	156.000.000
60	communications company-wide	17.600.000	only Islamic University	3.370.000	strength and weak point	7.800.000
61	Audio-Visual lines up	24.800.000	Arabic and English teaching medium	1.570.000	whole research	1.060.000.000

62	video-conference	33.900.000	two languages	355.000.000	early business	880.000.000
63	tele-conference	4.160.000	medium of communication	242.000.000	culinary business	29.600.000
64	control system	174.000.000	language academic assignments	41.700.000	first agreement	25.000.000
65	touch-screen panel	28.000.000	research report	1.650.000.000	couple times	407.000.000
66	installation accessories	108.000.000	whole facilities	503.000.000	many years	1.820.000.000
67	digital signage	10.700.000	single system	1.610.000.000	actually unprepared specifically	2.990.000
68	service solutions	72.900.000	routine activities and programs	594.000	useful things	318.000.000
69	broadcast station	122.000.000	regular progress report	9.690.000	some activity	1.070.000.000
70	humidity control	83.800.000	regularly forwarded	33.200.000	little business	67.800.000
71	fresh air supply ventilation	1.070.000	director of the school	859.000.000	lack of experience	386.000.000
72	advanced climate control	32.000.000	main duty	422.000.000	culinary background	8.240.000
73	optimally comfortable indoor environment	602.000	main duty of the foundation	102.000.000	existing business	42.800.000
74	real values	603.000.000	Waqf maintenance and enlargement	<b>20.300</b>	run perfectly	1.600.000.000
75	water chillers	604.000	waqf land	<b>271.000</b>	new business	3.050.000.000
76	variable refrigerant volume	<b>137.000</b>	program of the foundation	942.000.000	new system	3.150.000.000
77	other manufacturers	389.000.000	productivity of the land	97.300.000	second choice	45.400.000
78	variable refrigerant flow	<b>339.000</b>	family welfare	10.600.000	some things interestingly	50.200.000
79	extremely efficient	18.000.000	self finance	14.100.000	major priority	259.000.000
80	energy saving ways	4.730.000	family association	26.200.000	close connection	566.000.000
81	types of buildings	301.000.000	meeting place	41.100.000	academic background	227.000.000
82	minimum installation time	226.000.000	different regions	327.000.000	mechanical engineering program	32.800.000
83	best system	1.800.000.000	student organization	37.900.000	human beings	57.800.000

84	currently available	1.320.000.000	media of learning	1.090.000.000	middle and upper class	6.070.000
85	mid to large applications	161.000.000	whole activities of students	175.000.000	restaurant concept	96.400.000
86	mechanical device	17.500.000	main aim	465.000.000	middle-lower class	2.170.000
87	heat exchange	11.200.000	well qualified Muslim leader	602.000	first business	2.370.000.000
88	closed loop system	8.800.000	student union	15.100.000	restaurant business	550.000.000
89	special partners	573.000.000	other university	2.650.000.000	noodle stall	<b>462.000</b>
90	installation devices	233.000.000	student activities	25.300.000	alternative staple food	2.470.000
91	main distribution strategies	7.410.000	one of devotion place	26.200.000	favorite food	382.000.000
92	distribution strategy	9.920.000	management and community development	50.900.000	hard effort	429.000.000
93	major concern	310.000.000	business management	2.280.000.000	simple tools	433.000.000
94	any company	2.640.000.000	rural development	57.200.000	beginning of creation	39.800.000
95	role of distribution	625.000.000	social service	143.000.000	home-made noodle	15.500.000
96	right time	3.280.000.000	surrounding communities	132.000.000	main target	468.000.000
97	minimum cost	654.000.000	students language activities	416.000.000	better person	56.400.000
98	source of production	1.050.000.000	daily activities	567.000.000	true identity	14.300.000
99	analytical plan	124.000.000	department of teaching	533.000.000	biggest company	357.000.000
100	strategy of distribution	310.000.000	learning process	1.310.000.000	wet noodle	<b>464.000</b>

Sumber: Data Pribadi