

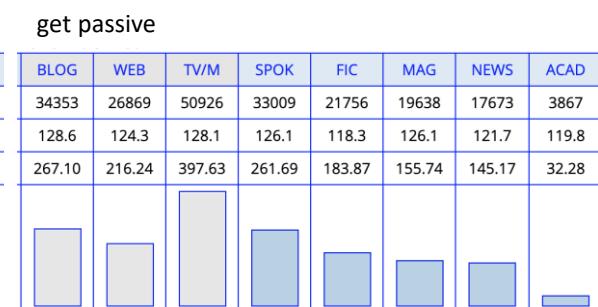
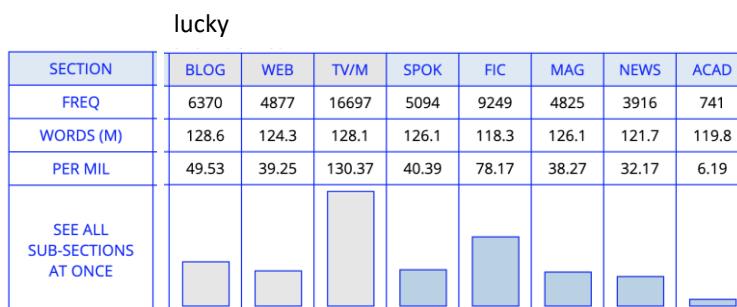
The COCA corpus (new version released March 2020)

The corpora from English-Corpora.org are the world's **most widely-used corpora**. The Corpus of Contemporary American English (COCA) is by far the most widely-used of these corpora. In early 2020, we dramatically expanded the scope and size and features of COCA to make it even more useful for researchers, teachers, and learners.

The corpus contains more than **one billion words** of data, including 20 million words each year from **1990-2019** (with the same genre balance year by year). This makes COCA the only corpus of English that is 1) large 2) recent and 3) has a wide range of genres. The following table shows the genres in the corpus.

Genre	# texts	# words	Explanation
Spoken	44,803	127,396,932	Transcripts of unscripted conversation from more than 150 different TV and radio programs (examples: All Things Considered (NPR), Newshour (PBS), Good Morning America (ABC), Oprah)
Fiction	25,992	119,505,305	Short stories and plays from literary magazines, children's magazines, popular magazines, first chapters of first edition books 1990-present, and fan fiction.
Magazines	86,292	127,352,030	Nearly 100 different magazines, with a good mix between specific domains like news, health, home and gardening, women, financial, religion, sports, etc.
Newspapers	90,243	122,958,016	Newspapers from across the US, including: USA Today, New York Times, Atlanta Journal Constitution, San Francisco Chronicle, etc. Good mix between different sections of the newspaper, such as local news, opinion, sports, financial, etc.
Academic	26,137	120,988,361	More than 200 different peer-reviewed journals. These cover the full range of academic disciplines, with a good balance among education, social sciences, history, humanities, law, medicine, philosophy/religion, science/technology, and business
Web (Genl)	88,989	129,899,427	Classified into the web genres of academic, argument, fiction, info, instruction, legal, news, personal, promotion, review web pages (by Serge Sharoff). Taken from the US portion of the GloWbE corpus.
Web (Blog)	98,748	125,496,216	Texts that were classified by Google as being blogs. Further classified into the web genres of academic, argument, fiction, info, instruction, legal, news, personal, promotion, review web pages. Taken from the US portion of the GloWbE corpus.
TV/Movies	23,975	129,293,467	Subtitles from OpenSubtitles.org, and later the TV and Movies corpora. Studies have shown that the language from these shows and movies is even more colloquial / core than the data in actual "spoken corpora".
	485,179	1,002,889,754	

GENRES Because COCA has so much data from each of these eight genres, it provides useful information about the frequency of words, phrases, and grammatical constructions across the genres – whether they are very informal (e.g. TV and movie subtitles or in spoken transcripts), more formal (e.g. academic articles), or somewhere in between (e.g. magazines and newspapers). For example, the following two charts show the frequency of the word *lucky* in the different genres, as well as the "get passive" (e.g. *he got promoted*).



The great general balance (and large size) of COCA also allows you to see the frequency of related phrases across genres. For example, the following table shows “soft NOUN”.

	CONTEXT	ALL	BLOG	WEB-GENL	TV/MOVIES	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
1	SOFT TISSUE	1120	62	67	74	39	36	100	35	707
2	SOFT DRINKS	1109	123	123	42	90	83	304	296	48
3	SOFT MONEY	790	21	34	12	446	8	78	153	38
4	SOFT SPOT	867	133	110	166	63	159	135	86	15
5	SOFT DRINK	721	43	60	48	77	68	191	199	35
6	SOFT VOICE	546	12	39	10	11	351	53	52	18
7	SOFT POWER	421	49	53	1	64	1	51	32	170
8	SOFT TISSUES	328	13	23	5	6	10	41	6	224
9	SOFT LANDING	274	42	19	18	38	17	56	63	21
10	SOFT LIGHT	247	22	25	9	1	121	38	21	10
11	SOFT PEAKS	216	2	4		1	2	144	62	1
12	SOFT SKIN	207	9	24	28	6	114	19	4	3
13	SOFT TOUCH	213	22	26	28	15	43	51	25	3
14	SOFT SKILLS	219	41	41		7		13	13	104

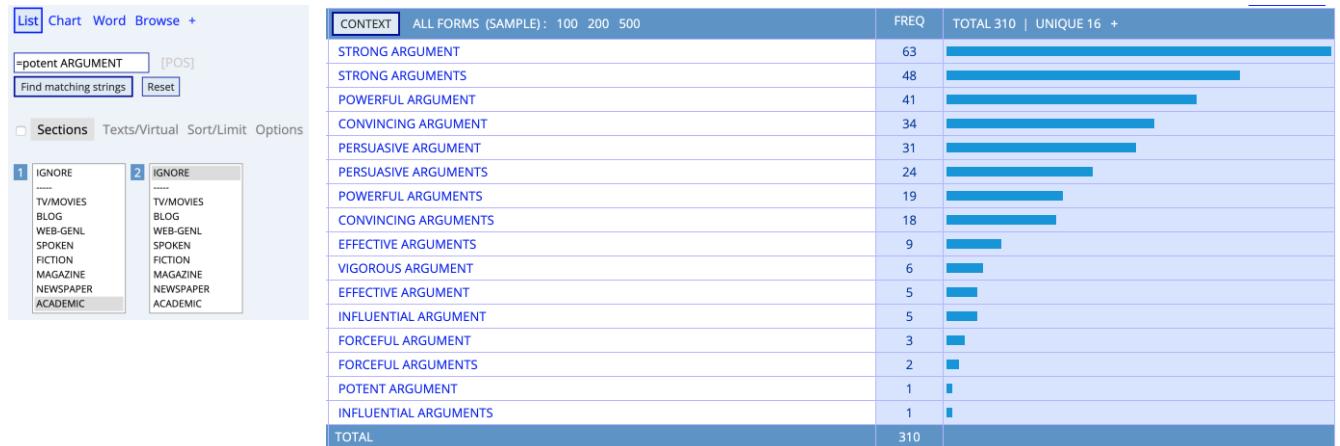
You can also compare two genres (or sets of genres). For example, the following table shows collocates (nearby words) of *chair* in ACADEMIC (left) and FICTION (right).

SEC 1 (ACADEMIC): 119,790,456 WORDS						SEC 2 (FICTION): 118,322,084 WORDS						
	WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2		WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1	DIVISION	56	1	0.5	0.0	55.3	KITCHEN	293	3	2.5	0.0	98.9
2	ENGINEERING	53	1	0.4	0.0	52.4	WING	111	0	0.9	0.0	93.8
3	ABA	55	0	0.5	0.0	45.9	WAY	87	1	0.7	0.0	88.1
4	PROGRAM	46	1	0.4	0.0	45.4	FEET	220	3	1.9	0.0	74.2
5	EDUCATION	42	0	0.4	0.0	35.1	FOOT	72	1	0.6	0.0	72.9
6	RESEARCH	34	1	0.3	0.0	33.6	MOTHER	134	2	1.1	0.0	67.8
7	FACULTY	38	0	0.3	0.0	31.7	COMMAND	65	1	0.5	0.0	65.8
8	ASME	36	0	0.3	0.0	30.1	LEATHER	409	7	3.5	0.1	59.2
9	COMMISSION	30	1	0.3	0.0	29.6	CANVAS	58	1	0.5	0.0	58.7
10	ROLE	33	0	0.3	0.0	27.5	BEACH	67	0	0.6	0.0	56.6
11	SUBCOMMITTEE	24	1	0.2	0.0	23.7	FINGERS	53	1	0.4	0.0	53.7
12	DESIGN	28	0	0.2	0.0	23.4	LAWN	318	6	2.7	0.1	53.7
13	MEMBER	45	2	0.4	0.0	22.2	VISITOR	51	1	0.4	0.0	51.6
14	RISE	44	2	0.4	0.0	21.7	CAPTAIN	61	0	0.5	0.0	51.6
15	DEAN	26	0	0.2	0.0	21.7	DAY	50	1	0.4	0.0	50.6

You can even compare synonyms of a given word in different genres. For example, this table shows the synonyms of *strong* in TV/MOVIES (left) and ACADEMIC (right).

SEC 1 (TV/MOVIES): 128,074,534 WORDS						SEC 2 (ACADEMIC): 119,790,456 WORDS						
	WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2		WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1	SPICY	581	53	4.5	0.4	10.3	ROBUST	2829	108	23.6	0.8	28.0
2	BEEFY	64	7	0.5	0.1	8.6	EFFECTIVE	28807	1596	240.5	12.5	19.3
3	TOUGH	11190	1579	87.4	13.2	6.6	DURABLE	683	65	5.7	0.5	11.2
4	STRAPPING	143	21	1.1	0.2	6.4	CONCENTRATED	2750	298	23.0	2.3	9.9
5	HOT	23639	3877	184.6	32.4	5.7	STAUNCH	252	32	2.1	0.2	8.4
6	BRILLIANT	4418	1156	34.5	9.7	3.6	ZEALOUS	217	31	1.8	0.2	7.5
7	GREAT	107836	34522	842.0	288.2	2.9	ARDENT	375	54	3.1	0.4	7.4
8	BITING	658	230	5.1	1.9	2.7	COMPELLING	2845	452	23.7	3.5	6.7
9	BURLY	52	27	0.4	0.2	1.8	FERVENT	210	34	1.8	0.3	6.6
10	BRIGHT	4561	2542	35.6	21.2	1.7	CLEAR-CUT	405	83	3.4	0.6	5.2
11	SOUND	18257	11719	142.5	97.8	1.5	PERSUASIVE	1360	308	11.4	2.4	4.7
12	WELL-BUILT	20	14	0.2	0.1	1.3	DEEP-SEATED	260	61	2.2	0.5	4.6

The ability to focus in on specific genres means that you can find “just the right word” for a particular concept in a particular genre. For example, suppose that you want to know what word related to *potent* is the most frequent with a form of *argument* in academic English. As the following image shows, you simply search for =*potent ARGUMENT* and limit the search to academic, and you would see the following results, and then (as with any search results) can see the matching phrases in context,



1	1990	ACAD	EnvirAffairs	A	B	C	not be reflected in purchase prices. Therefore, affected proprietors can make a strong argument that there is no inherent justice in requiring th
2	2009	ACAD	AmerIndianQ	A	B	C	, which is particularly based in theories of meaning. Craig Womack makes a strong argument that Native literature can not be studied simply as
3	1996	ACAD	SocialStudies	A	B	C	military prowess to guarantee them independence from potential subjugation. In fact, a strong argument could be made that the Comanches si
4	2000	ACAD	Mercury	A	B	C	contained one of the most read passages in all of science fiction and a strong argument to explore space. On the Moon an extraterrestrial beac
5	1991	ACAD	IntlAffairs	A	B	C	everyone and to no one. # Hungarian-born economist Janos Kornai makes a strong argument in favor of selling firms to "flesh and blood" ris
6	1990	ACAD	ForeignAffairs	A	B	C	-- mounted by USSOCOM elements-were used with considerable effect during its course. A strong argument can be made, however, that it wou
7	2017	ACAD	Stanford Law Review	A	B	C	down the statute on First Amendment grounds and finding that there is a " strong argument that prescriber-identifying information is speech. "
8	2016	ACAD	Eighteenth - Century Studies	A	B	C	force behind fulfilling duties of justice across borders. Consequently, d'Holbach offers a strong argument for why fulfilling cosmopolitan duties
9	1994	ACAD	CanadaLaw	A	B	C	, the greater the outflow of payments for copyright works. There is a strong argument for a country with a relatively small economy to be a follo
10	2017	ACAD	Stanford Law Review	A	B	C	extend to prisoners.256 Calvin, anticipating this aspect of Evenwel, set forth a strong argument of its own, explaining that prisoners are differen
11	2014	ACAD	PolSciQuarterly	A	B	C	security issues, but so is the ability of the individual to discern a strong argument from a weak one. People who are better educated will probab
12	2014	ACAD	Ref&UserServQ	A	B	C	librarians. I recognize that this is not currently the case, thus your strong argument to maintain the OPAC. However, if nothing changes, we will
13	2001	ACAD	ArmedForces	A	B	C	local militia forces. He addresses religion only once, but he makes a strong argument for the importance that it played in his combat life, stating

In addition to comparing across genres, you can also compare words, to “tease out” differences between related words. For example, the following chart shows words that occur with *deep* and *profound*, showing that (for example) *deep breath* is common but *profound breath* never occurs, or that *profound effect* is common but *deep effect* is not. You can also compare other related words, such as adjectives used with the words *men* and *women*, or verbs occurring near *Obama* or *Trump*.

WORD 1 (W1): DEEP (7.33)

	WORD	W1	W2	W1/W2	SCORE
1	BREATH	7687	0	15,374.0	2,096.7
2	SPACE	1292	0	2,584.0	352.4
3	BREATHS	1285	0	2,570.0	350.5
4	POCKETS	764	0	1,528.0	208.4
5	WATER	1224	1	1,224.0	166.9
6	THROAT	513	0	1,026.0	139.9
7	OCEAN	414	0	828.0	112.9
8	VOICE	799	1	799.0	109.0
9	SNOW	360	0	720.0	98.2
10	RED	337	0	674.0	91.9
11	SEA	671	1	671.0	91.5
12	FREEZE	327	0	654.0	89.2
13	EYES	318	0	636.0	86.7
14	ROOTS	615	1	615.0	83.9

WORD 2 (W2): PROFOUND (0.14)

	WORD	W2	W1	W2/W1	SCORE
1	HEARING	54	0	108.0	791.9
2	DISABILITIES	25	0	50.0	366.6
3	AFFECT	23	0	46.0	337.3
4	EFFECT	695	16	43.4	318.5
5	EFFECTS	214	7	30.6	224.2
6	DISABILITY	14	0	28.0	205.3
7	REVELATION	14	0	28.0	205.3
8	LESSON	27	1	27.0	198.0
9	MANNER	11	0	22.0	161.3
10	RETARDATION	11	0	22.0	161.3
11	SHIFT	100	5	20.0	146.7
12	GIFT	10	0	20.0	146.7
13	IMPAIRMENT	10	0	20.0	146.7
14	STATEMENT	57	3	19.0	139.3

HISTORICAL COCA is the only large corpus of English that has extensive data from the entire period of the last 30 years – 20 million words per year from 1990-2019 (with the same genre balance year by year). This means that in addition to seeing variation by genre, you can also map out recent changes in English in ways that are not possible with any other corpus – such as with the frequency of *awesome* from 1990-2019.,.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	24063	9902	6800	8494	2740	957	1819	1352	350	940	1127	1688	3075	4523	4359
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	139.1	147.8	146.6	144.9	145.3	144.7
PER MIL	24.23	76.99	54.73	66.32	21.72	8.09	14.43	11.11	2.92	6.76	7.63	11.52	21.21	31.14	30.12
SEE ALL SUB-SECTIONS AT ONCE															

And of course you can look at much more than just simple words or phrases. COCA is the only corpus that allows you to map out changes in syntactic constructions over the past 30 years, as with the “like construction” (*and I’m like, no way*) or the “end up V-ing” construction (*you’ll end up paying way too much*) – both of which have increased in each five year period since the early 1990s.

SECTION	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	140	393	639	1145	1780	2581	1826	2340	2489	2849	2949	3293
WORDS (M)	139.1	147.8	146.6	144.9	145.3	144.7	139.1	147.8	146.6	144.9	145.3	144.7
PER MIL	1.01	2.66	4.36	7.90	12.25	17.83	13.13	15.83	16.98	19.66	20.30	22.75
SEE ALL SUB-SECTIONS AT ONCE	like construction (CONJ PRON BE like ' .)						END up V-ing					

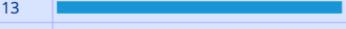
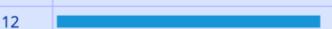
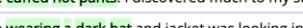
Just as we compared collocates (nearby words) in different genres above (for *chair* in ACAD and FIC), we can also compare collocates in different periods, to look at semantic change. For example, the following table shows collocates of *green* in the 1990s (left) and the 2010s (right).

SEC 1 (1995-1999, 1990-1994): 286,833,557 WORDS						SEC 2 (2010-2014, 2015-2019): 290,003,115 WORDS								
WORD/PHRASE		TOKENS 1	TOKENS 2	PM 1	PM 2	RATIO	WORD/PHRASE		TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO	
1	GREEN RANGER	187	0	0.7	0.0	65.2	1	GREEN GAZETTE	96	0	0.3	0.0	33.1	
2	GREEN PEPPER	215	35	0.7	0.1	6.2	2	GREEN JOBS	87	0	0.3	0.0	30.0	
3	GREEN CROSS	52	9	0.2	0.0	5.8	3	GREEN PRACTICE	60	2	0.2	0.0	29.7	
4	GREEN VEGETABLES	73	32	0.3	0.1	2.3	4	GREEN ENERGY	170	7	0.6	0.0	24.0	
5	GREEN PEPPERS	94	47	0.3	0.2	2.0	5	GREEN ARROW	192	8	0.7	0.0	23.7	
6	GREEN MAN	58	31	0.2	0.1	1.9	6	GREEN BUILDING	130	18	0.4	0.1	7.1	
7	GREEN ACRES	52	31	0.2	0.1	1.7	7	GREEN SCREEN	85	12	0.3	0.0	7.0	
8	GREEN BELL	154	92	0.5	0.3	1.7	8	GREEN ZONE	118	21	0.4	0.1	5.6	
9	GREEN PLANTS	61	37	0.2	0.1	1.7	9	GREEN LANTERN	96	21	0.3	0.1	4.5	
10	GREEN GLASS	61	37	0.2	0.1	1.7	10	GREEN SPACES	97	23	0.3	0.1	4.2	

We can also see what is being said about a topic over time. For example, the following shows the collocates of *crisis* in each five year period (and genre) from 1990-2019, which shows what we were worrying about in these different periods.

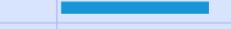
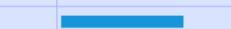
CONTEXT	ALL	BLOG	WEB-GENL	TV/MOVIES	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC	1990-1994	1995-1999	2000-2004	2005-2009	2010-2014	2015-2019
DEBT	1513	370	419	4	264	2	76	203	175	115	63	25	29	420	72
CRISIS	1088	142	146	64	268	24	140	128	176	194	128	76	116	162	124
HEALTH	917	65	92	36	209	7	214	177	117	149	84	84	133	115	195
ENERGY	771	87	96	32	129	10	135	196	86	94	43	245	123	48	35
GULF	656		5	3	267	1	122	194	64	629	11	3	2	4	2
OIL	437	58	38	8	76	8	89	69	91	132	46	64	53	28	18
REFUGEE	381	10	16	6	122	6	57	65	99	47	48	19	15	27	199
SOLUTION	381	49	49	4	97	7	42	67	66	116	31	33	36	40	27
MIDLIFE	348	32	33	115	62	24	40	21	21	25	31	45	26	104	52
CREDIT	385	86	82	2	46	1	50	94	24	3	9	2	138	50	15
CLIMATE	396	114	80	2	65	4	86	14	31	2	1	12	48	33	106

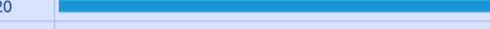
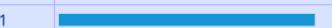
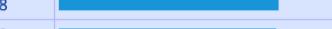
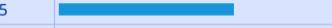
As with all of the other corpora from English-Corpora.org, there is a very **wide range of searches**, including: words, phrases, substrings, lemmas, part of speech, synonyms, and customized wordlists. For example, the search **WEAR * ADJ @CLOTHES** takes just about one second to search through the billion words to find strings like the following (and it doesn't require learning unnecessarily convoluted search syntax).

	<input type="checkbox"/> CONTEXT	ALL FORMS (SAMPLE): 100 200 500	FREQ	TOTAL 1,292 UNIQUE 900 +
1	<input type="checkbox"/>	WEARING A TOP HAT	22	
2	<input type="checkbox"/>	WEARING A SHORT SKIRT	18	
3	<input type="checkbox"/>	WEARING A HARD HAT	14	
4	<input type="checkbox"/>	WEARING A WHITE BLOUSE	14	
5	<input type="checkbox"/>	WEAR A SHORT SKIRT	13	
6	<input type="checkbox"/>	WEARING A WHITE HAT	13	
7	<input type="checkbox"/>	WERE A SHORT SKIRT	13	
8	<input type="checkbox"/>	WEARING A RED TIE	12	

1	2019	FIC	PennLitJournal	A B C know Jessica, the salesgirl with the husky voice who as usual wore cuffed hot pants . I discovered much to my surprise that she lived in southeast, to
2	2019	NEWS	Chicago Tribune	A B C the 2400 block of Partridge Lane told police Feb. 19 that someone wearing a dark hat and jacket was looking into a neighbor's vehicle. # A resident i
3	2019	TV	The Blacklist	A B C to the defendant? It was the only suit in the area wearing a fedora-style hat . I see. And did you arrest the defendant at that point? No
4	2018	NEWS	OregonLive.com	A B C Security disability benefits, punched the guard after he was accused of wearing a stolen hat on May 12, 2017. # Nathaniel M. Williams, 25Columbia
5	2017	FIC	Bk:One-hundred-knuckledFist	A B C because he looks toward the sound, too. # A woman wearing a bright-yellow blouse and black pants hums past his stand, armed with a small vacuu
6	2017	MAG	Fox News	A B C evacuated Tuesday after an explosion was reported and police shot a suspect wearing an explosive belt who shouted " Allahu Akbar, " in what offici
7	2017	NEWS	USA TODAY	A B C in Jerusalem's Old City on May 22, 2017. She wore a white skirt suit by Michael Kors and her customary sky-high heels in a brightly colored striped p
8	2016	FIC	Bk:LoveMaps	A B C my apple! " Maya yelled from the porch. # She wore a white blouse tied at the waist and a blue-and-black plaid skirt. She skipped toward them,
9	2016	FIC	Dialogue: Mormon	A B C .Approaching her, he saw it was Jennie O'Brien. She wore a drab skirt , and her hair was wound into an untidy bun upon her head. #

Because of COCA's advanced architecture, even searches for very general searches like **NOUN + NOUN** or **VERB ADJ NOUN** take just 1-2 seconds to search through the billion word corpus:

	<input type="checkbox"/> USE LARGE NGGRAMS [?]	FREQ	TOTAL 8,108,778 UNIQUE 17,446 +
1	<input type="checkbox"/> HEALTH CARE	64877	
2	<input type="checkbox"/> WORLD WAR	32709	
3	<input type="checkbox"/> LIVING ROOM	23171	
4	<input type="checkbox"/> LAW ENFORCEMENT	23057	
5	<input type="checkbox"/> WEB SITE	22382	
6	<input type="checkbox"/> CLIMATE CHANGE	21967	
7	<input type="checkbox"/> AIR FORCE	19768	
8	<input type="checkbox"/> REPUBLICAN PARTY	16928	
9	<input type="checkbox"/> STATE UNIVERSITY	16926	
10	<input type="checkbox"/> HEALTH INSURANCE	16378	
11	<input type="checkbox"/> FAMILY MEMBERS	16031	

	<input type="checkbox"/> USE LARGE NGGRAMS [?]	FREQ	TOTAL 178,237 UNIQUE 6,395 +
1	<input type="checkbox"/> TAKE GOOD CARE	1226	
2	<input type="checkbox"/> MAKES PERFECT SENSE	1120	
3	<input type="checkbox"/> PAY CLOSE ATTENTION	706	
4	<input type="checkbox"/> FIND COMMON GROUND	683	
5	<input type="checkbox"/> BEGIN AUDIO CLIP	651	
6	<input type="checkbox"/> SAY GOOD NIGHT	641	
7	<input type="checkbox"/> TAKE FULL ADVANTAGE	558	
8	<input type="checkbox"/> GRATED PARMESAN CHEESE	553	
9	<input type="checkbox"/> COME FULL CIRCLE	536	
10	<input type="checkbox"/> COME RIGHT BACK	491	
11	<input type="checkbox"/> WORK FULL TIME	445	
12	<input type="checkbox"/> TRY NEW THINGS	439	

A unique feature of COCA, which makes it very useful for language learners and teachers, is the ability to **browse** through a list of the **top 60,000 words** (lemmas) in the corpus, and then to see an extremely wide range of information on each of these words.

Word form	<input type="text"/>							
Part of speech	<input checked="" type="checkbox"/> NOUN <input checked="" type="checkbox"/> VERB <input checked="" type="checkbox"/> ADJ <input checked="" type="checkbox"/> ADV <input checked="" type="checkbox"/> OTHER <input type="checkbox"/> ALL							
Range	5145 - <input type="text"/>							
Pronunciation	Rhymes with <input type="text"/> Type EXACT							
Syllables / stress								

For example, the following are just a few examples of high frequency words (about word #5000 in the 60,000 word list), medium frequency (~25,000), and low frequency (~45,000) words. For each word in the list, users can hear the word **pronounced**, see **videos** with that word in the text, find related **images** from Google Images, and see a **translation** for their preferred target language.

	RANK	FREQ	Word	PoS	Audio	Video	Image	ZH-CN
1	5200	11377	blogger	NOUN				
2	5201	11374	utterly	ADV				
3	5202	11372	trouble	VERB				
4	5203	11368	texture	NOUN				
5	5204	11365	head	ADJ				

	RANK	FREQ	Word	PoS	Audio	Video	Image	ZH-CN
1	25201	577	unappealing	ADJ				
2	25202	577	tryst	NOUN				
3	25203	577	urologist	NOUN				
4	25204	577	scrabble	NOUN				
5	25205	577	demarcate	VERB				

	RANK	FREQ	Word	PoS	Audio	Video	Image	ZH-CN
1	45107	112	incubus	NOUN				
2	45108	112	unquantifiable	ADJ				
3	45110	112	husk	VERB				
4	45111	112	hunky-dory	ADJ				
5	45112	112	jacked	ADJ				

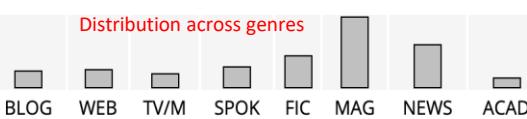
It is even possible to search the 60,000 word list by **pronunciation** (very helpful, because of difficult English spelling). For example, the following is a partial list of two syllable words (accented on the second syllable) that rhyme with *stay*:

Pronunciation	Rhymes with <input type="text"/> Type EXACT							
Syllables / stress								

11	3958	17233	array	NOUN				
12	4122	16221	convey	VERB				
13	4589	13766	survey	VERB				
14	4751	13055	cafe	NOUN				
15	4779	12934	betray	VERB				
16	5536	10325	obey	VERB				
17	5933	9149	ballet	NOUN				

Each of the **top 60,000 words** in the corpus has a “[home page](#)” such as the following, with links to other pages with more information. Users can save words to a “favorites” list for later review, and go back through a history of all of their “word” pages

[trail](#) (NOUN) See: VERB Rank #1-60,000 #2102



Pronunciation Definition

1. a track or mark left by something that has passed 2. a path or track roughly blazed through wild or hilly country 3. evidence pointing to a possible solution D M) C G

PlayPhrase YouGlish Yarn Links to videos

ZH-CN: Google WordRef Reverso Linguee

Google images

Links to translations

TOPICS (more) Words that co-occur in 22 million web pages

hike, hike, mountain, steep, park, terrain, mile, forest, creek, lake, wilderness, route, peak, river, trail, visitor, valley, hiker, snow, climb

COLLOCATES (more) Words that co-occur nearby

NOUN campaign, mile, paper, blazer, park, mountain, blood, road

VERB follow, hike, lead, blaze, ride, climb, cross, wind

ADJ appalachian, steep, narrow, historic, scenic, rocky, indian, rough

ADV left, downhill, steeply, uphill, aptly, underfoot, deftly, briskly

SYNONYMS (more)

path road, track, path, trail, route, trajectory, footpath **track**

trail, trace

CLUSTERS (more) 2, 3, 4 word strings

trail •	trail in • trail to • trail blazers • trail for • trail with • trails in • trail from • trail on
• trail	campaign trail • paper trail • appalachian trail • hiking trails • long trail • blood trail • on trails • money trail
trail ••	trail of tears • trail of blood • trail of destruction • trail that led • trail leads to • trail of smoke • trail of bread • trail of broken
•• trail	on the trail • along the trail • up the trail • off the trail • down the trail • leaving a trail • left a trail • to the trail
trail •••	trail of bread crumbs • trail that led to • trail to the top • trail through the woods • trails in the sky • trail of destruction across • trail that leads to • trail of broken hearts
••• trail	on the campaign trail • out on the trail • from the campaign trail • end of the trail • hot on the trail • side of the trail • on the appalachian trail • miles of hiking trails

TEXTS / VIRTUAL CORPORA (more) Texts that have this as a “keyword”

WEB:nps.gov • WEB:ponderosaca.com • NEWS:AssocPress • MAG:Backpacker • ACAD:InstrPsych • MAG:Bicycling • NEWS:USA TODAY • MAG:Backpacker • FIC:FantasySciFi • MAG:Backpacker • MAG:NatlParks • NEWS:USAToday • MAG:Backpacker • MAG:Backpacker • WEB:...documentaryfilms.com • MAG:NatlParks • MAG:Bicycling • MAG:Backpacker • MAG:Bicycling • MAG:Backpacker • NEWS:Atlanta • MAG:Backpacker • MAG:Backpacker • MAG:NatlParks • NEWS:Atlanta •

CONCORDANCE LINES (more) The word in context (to see patterns of use)

1	WEB: 2012: amazon.com	Along the way , they confront Army Rangers along the Appalachian Trail and conduct a late-night stealth mission to an abandoned NASA
2	BLOG: 2012: ...heother.blogspot....	me , a direct definition of ultra trail running . The trails are amazing DCNR does an excellent job keeping the trails and
3	MAG: 1991: NatlParks	to protect a continuous corridor are ongoing . The North Country Trail Association a 500-member organization established in 1980 ,
4	MAG: 1991: NatlParks	. # For more information , contact the Overmountain Victory Trail Association c/o P.O. Box 632 , Manassas Park , VA 22111
5	MAG: 2003: Bicycling	and anti-bike website operator John Parulis , describing the trails at China Camp , east of San Rafael " Look at this
6	WEB: 2012: yelp.com	try ... # Perks of coming here is hiking the various trails at the top , and taking the ski lift back down .
7	BLOG: 2012: seattlemet.com	that the city did not " re-interview " businesses along the trail before approving its completion ; and the fact that an
8	TV: 2016: Tomorrow, When the W...	OK ? Rain gear 's there too . Should be some trail bikes out the back if they have n't been taken . Awesome
9	NEWS: 1998: Houston	because only a few of the wealthiest teams - the Portland Blazers (owned by billionaire Paul Allen) , the New York
10	MAG: 2010: PopMech	, I also fail to notice that the side of the trail butts a 50-foot cliff off the creek bed . With my blissful
11	MAG: 2009: OutdoorLife	to find one . He showed me some images from a trail camera that was placed on the stand farthest from camp , the
12	TV: 2014: Avengers Assemble	, arizona . How did ... Technopathic static . Covering one trail can sometimes make another . I told you I 'd back your
13	MAG: 2007: AmSpec	first four miles or so were quite steep , as the trail climbed steadily up the drainage through thick timber above the
14	MAG: 1997: Backpacker	and its remote eastern tributaries , the vast and contorted Burr Trail country , the huge and austere Kaiparowits Plateau , and the slot
15	MAG: 2009: Backpacker	, so we divert onto a slippery embankment that rejoins the trail downriver . Later , we recognize what 's left of an elk
16	MAG: 1997: Backpacker	and describes himself as a warm sleeper . Often on the trail for 10 or more days at a time , Mike values a
17	MAG: 2008: Bicycling	paths . Barnack can see riders heading up the Tollgate Trail from his kitchen window . But the Tims also ride . Barnack
18	NEWS: 2018: OregonLive.com	TriMet and PGE are working a section of the Springwater Corridor Trail from Southeast Tacoma to 45th Place and the widely used trail is
19	NEWS: 1992: WashPost	tour , the bus drops hikers off at the Skyline Drive trail head and picks them up at the bottom . A historian and
20	MAG: 1997: Backpacker	a free backcountry trip planner and general map , or call Trails Illustrated (800) 962-1643 for its map to the park

Each of the top 60,000 words also has more detailed pages, including a “dictionary” page, related topics, collocates, clusters, websites, and concordance lines. Samples of each of these pages are given below.

DICTIONARY page

Includes a **definition**, links to Google **Images**, **pronunciation**, **videos**, and **translations** (to desired language) at (up to) four different sites. Also includes **synonyms** and words with more specific meanings (**hyponyms**) and more general meanings (**hypernyms**) (both from WordNet). It also includes **frequency information** (including rank order, number of tokens, and two “range” measures of how well the word is spread throughout the ~500,000 texts, as well as its distribution across the eight main genres in the corpus. Finally, it also includes the frequency of the different **forms** of the lemma (e.g. verb forms), and **related words**.

1. a raised horizontal surface 2. a document stating the aims and principles of a political party 3. a woman's shoe with a very high thick sole 4. any military structure or vehicle bearing weapons 5. the combination of a particular computer and a particular operating system D M O C G

PlayPhrase YouGlish Yarn

ZH-CN: Google WordRef Reverso Linguee

SYNONYMS (more)

policy policy, proposal, platform, manifesto stage stand, platform, podium, dais

MORE SPECIFIC MEANING (click on blue word)

footplate (British) the platform in the cab of a locomotive on which the engineer stands to operate the controls

turntable a circular horizontal platform that rotates a phonograph record while it is being played

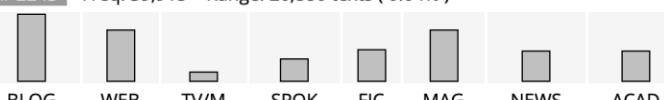
stage a large platform on which people can stand and can be seen by an audience

runway a narrow platform extending from the stage into the audience in a theater or nightclub etc.

footboard a narrow platform on which to stand or brace the feet

FREQUENCY INFORMATION ?

2243 Freq: 39,915 Range: 20,886 texts (0.04%)



WORD FORMS

platform (31,414), platforms (8,501)

RELATED WORDS

platform (n), cross-platform (j), multi-platform (j), platform-specific (j), platform-independent (j)

MORE GENERAL MEANING (click on blue word)

level a flat surface at right angles to a plumb line

structure a thing constructed

document writing that provides information (especially information of an official nature)

construction a thing constructed

shoe footwear shaped to fit the foot (below the ankle) with a flexible upper of leather or plastic and a sole and heel of heavier material

“One click” link to Google Images (alternative for China)



“One-click”
links to
translations
(4 different
sites)

On Wikipedia, it's an open platform that
words and images

“One click”
links to
videos from
PlayPhrase,
YouGlish,
and Yarn

Linguee English ↔ German

platform noun

Plattform f

Social networks can serve as a platform for business.
The platform allows the sharing of files.

Bahnsteig m

My train will be departing from the platform soon.
Mind the gap between the platform and the train.

Podium nt

The speaker stood on a platform.

Gleis nt

The train leaves from platform seven.

less common:
Podest nt · Bühne f · Rampe f · Standfläche f · Pritsche f

Soziale Netzwerke können als Plattform für Geschäfte dienen.
Die Plattform ermöglicht das Teilen von Dateien.

Mein Zug wird in Kürze vom Bahnsteig abfahren.
Beachten Sie die Lücke zwischen dem Bahnsteig und dem Zug.

Der Redner stand auf einem Podium.

Der Zug fährt von Gleis sieben ab.

TOPICS page

Shows words that tend to co-occur in the 22 million webpages in the corpus. In many cases these provide better insight into meaning and usage than collocates (the standard tool for finding textually related words), and yet we're not aware of any other corpus that has these. Users can click on any of these words to follow a "chain" of related words.

TOPICS				HORMONE	NOUN	TOPICS				COSMOLOGICAL	ADJ	TOPICS				VIBRATE	VERB
	Same text	Word	PoS		Same text	Word	PoS		Same text	Word	PoS		Same text	Word	PoS		
1	253	estrogen	n	1	108	universe	n	1	40	vibration	n	1	40	vibration	n		
2	245	supplement	n	2	82	bang	n	2	30	frequency	n	2	30	frequency	n		
3	227	stimulate	v	3	73	physics	n	3	25	physics	n	3	25	physics	n		
4	224	therapy	n	4	71	cosmic	j	4	23	atom	n	4	23	atom	n		
5	209	cancer	n	5	70	cosmology	n	5	22	particle	n	5	22	particle	n		
6	203	disease	n	6	69	galaxy	n	6	21	quantum	n	6	21	quantum	n		
7	202	prescribe	v	7	68	quantum	n	7	19	faint	j	7	19	faint	j		
8	199	symptom	n	8	62	physicist	n	8	19	string	n	8	19	string	n		
9	192	diet	n	9	59	particle	n	9	17	molecule	n	9	17	molecule	n		
10	187	breast	n	10	57	astronomer	n	10	17	sound	n	10	17	sound	n		
11	174	stress	n	11	55	gravity	n	11	17	wave	n	11	17	wave	n		
12	171	nutrient	n	12	54	expansion	n	12	16	electron	n	12	16	electron	n		
13	168	obesity	n	13	53	theory	n	13	16	gravity	n	13	16	gravity	n		
14	161	protein	n	14	50	relativity	n	14	16	magnetic	j	14	16	magnetic	j		
15	161	researcher	n	15	48	atom	n	15	16	mathematical	j	15	16	mathematical	j		
16	160	induce	v	16	45	constant	j	16	16	motion	n	16	16	motion	n		
17	160	injection	n	17	45	correspond	v	17	16	physicist	n	17	16	physicist	n		
18	160	tissue	n	18	44	gravitational	j	18	16	vacuum	n	18	16	vacuum	n		
19	158	medicine	n	19	44	mathematical	j	19	15	grin	v	19	15	grin	v		

COLLOCATES page

Words that occur nearby. Click on any word to see the collocates for the new word, or click on the "text" icon to see the word in context. The gray boxes show the most normal placement of the node word (e.g. *Christmas tree*, but *tree trunk*). You can also sort by Mutual Information score and set frequency thresholds (Advanced Options).

COLLOCATES			TREE	NOUN	Advanced options				Clusters	Topics	Dictionary	Texts	KWIC	...
+ NOUN	NEW WORD	?	+ ADJ	NEW WORD	?	+ VERB	NEW WORD	?	+ ADV	NEW WORD	?			
4322	5.32	christmas	1392	4.46	tall	2514	2.95	grow	84	3.28	overhead			
2787	5.81	branch	793	5.50	olive	2282	5.96	plant	50	3.90	eg			
2333	5.98	palm	728	2.58	green	1929	2.59	fall	23	4.78	gnarly			
2021	5.12	fruit	497	6.34	fallen	1711	2.61	cut	20	2.64	gracefully			
1888	6.34	oak	395	2.99	thick	1452	4.64	climb	16	3.64	thickly			
1880	6.39	trunk	387	3.71	nearby	421	3.54	line	15	2.57	densely			
1782	6.14	pine	368	3.24	giant	412	4.83	decorate	15	3.79	headfirst			
1653	4.44	forest	368	4.12	bare	402	2.74	surround	14	2.61	thick			
1488	4.05	apple	365	7.75	phylogenetic	379	2.63	lean	14	3.62	yonder			
1475	4.77	leaf	315	7.79	downed	280	4.61	trim	12	3.32	skyward			
1236	7.73	shrub	278	4.34	mature	240	4.39	bark	10	2.52	organically			
1055	5.30	shade	276	3.26	orange	214	4.40	sway	8	2.91	precariously			
1052	4.22	root	236	8.18	deciduous	203	2.55	crash	7	2.73	headlong			
948	3.42	species	217	6.73	flowering	195	7.40	fell	7	3.21	sustainably			
945	2.71	plant	210	6.16	evergreen	191	5.91	shade						
929	5.56	limb	209	7.25	uprooted	174	3.44	chop						
848	3.99	stand	200	4.05	lone	155	6.47	prune						
754	4.21	grass	193	3.62	tropical	151	4.74	topple						
726	2.87	wind	177	3.81	dense	149	2.82	hug						
711	3.37	flower	174	5.60	towering	148	3.54	harvest						

CLUSTERS page

The most common 2, 3, and 4 word strings. You can decide how “wide” or “tight” you want the clusters to be (e.g. whether they include highly frequent words like articles and prepositions).

CLUSTERS	BREAD	NOUN	See also: VERB	LIMIT: Loose	Medium	Tight	N+N	Collocates	Clusters	Topics	Dictionary	Texts	KWIC
1145	bread crumbs	828	white bread	700	bread and butter	860	loaf of bread	74	bread on the table	133	day our daily bread		
421	bread pudding	544	corn bread	374	bread and wine	368	piece of bread	62	bread for the world	103	thing since sliced bread		
284	bread slices	391	french bread	158	bread and water	250	loaves of bread	25	bread and the wine	35	breaking of the bread		
177	bread dough	374	daily bread	158	bread and cheese	234	slice of bread	24	bread in five minutes	26	slices of white bread		
156	bread cubes	306	wheat bread	128	bread of life	194	slices of bread	16	bread in the toaster	24	top with remaining bread		
128	bread basket	302	fresh bread	84	bread and circuses	167	whole wheat bread	16	bread and butter issues	23	side of each bread		
119	bread flour	233	banana bread	71	bread and milk	109	pieces of bread	15	bread for the city	22	loaf of french bread		
111	bread machine	226	garlic bread	53	bread is buttered	92	gluten free bread	14	bread is buttered on	19	slice of white bread		
94	bread loaf	210	pita bread	50	bread and pasta	62	freshly baked bread	14	bread in the oven	19	loaf of wonder bread		
83	bread knife	205	sliced bread	48	bread and other	56	whole grain bread	14	bread which came down	18	french or italian bread		
79	bread lines	199	rye bread	39	bread and circuses	56	crust of bread	12	bread and peanut butter	16	salad and garlic bread		
78	bread made	197	wonder bread	35	bread and cereals	55	milk and bread	12	bread to the hungry	16	serve with crusty bread		
77	bread baking	190	crusty bread	35	bread and forgive	50	kind of bread	12	bread to sop up	16	smell of baking bread		
76	bread box	171	break bread	35	bread and salt	47	eat the bread	11	bread with olive oil	15	consecration of the bread		
71	bread crumb	166	stale bread	26	bread and then	45	trail of bread	11	bread and wine become	14	feast of unleavened bread		
71	bread alone	157	baking bread	26	bread and roses	44	cheese and bread	10	bread from the earth	14	loaf of white bread		

KWIC / concordance page

See 100 – 1000 random lines, with the surrounding words highlighted for part of speech. You can re-sort the words by words to the left or right, all with the goal of seeing the patterns in which the word occurs.

KWIC	CRACK	VERB	See also as: NOUN ADJ	# lines: 100 200 500 1000	Collocates	Clusters	Topics	Dictionary	Websites	KWIC
------	-------	------	-----------------------	---------------------------	------------	----------	--------	------------	----------	------

WEBSITE	SORT	SORT	SORT
11 TV:1996:	to need some minor adjustments . Want me to take a crack	at it	Vince ? Look , Vince if you insist on
12 MAG:1996:	, and cacao . While there , he took his first crack	at making	wine . But not from grapes-from palm trees . "
13 BLOG:2012:	for both sides " and prodded them to take yet another crack	at negotiating	some sort of agreement while declaring that " it
14 FIC:2012:	along for years , it 's just fair we get first crack	at the	open grazing . Last week Pick had the buckaroos mark
15 MAG:1993:	To all appearances , it was working well enough . Then cracks	began appearing	in its facade . # One day , for example
16 MAG:2007:	Humvee , and Hallway is pulling away when a single gunshot cracks	behind us	" Shot fired ! " Garlick shouts . The
17 WEB:2012:	crack was the problem . But it was n't as if crack	came to	Baltimore for the first time in 1995 . It had
18 ACAD:2007:	initially of the heroin markets of the 1970s and then the crack	cocaine markets	of the 1980s/1990s . What we imagine in terms of
19 NEWS:2001:	it 's tough to quit . " The Net is the crack	cocaine of	sex addiction , " she says . # Sexual addicts
20 SPOK:1992:	Voiceover) A search eventually turned up about 30 rocks of crack	cocaine	Unidentified Man 9 : It do n't look like much
21 NEWS:1995:	auto industry has erected to keep out competition . And those cracks	could help	Ford reach its long-term goal in Japan - a goal
22 SPOK:1997:	and she 's at large right now . She 's a crack	dealer	a very , very violent fugitive . Very violent .
23 SPOK:1996:	: O'Connor 's assessment jibes with the perceptions of former crack	dealers and	users at a Los Angeles area halfway house called
24 NEWS:2007:	# Suddenly , his spine can turn as crooked as a crack	down a	two-lane highway . # He must keep his quadriceps Ziploc
25 FIC:2013:	little , he remembered . Mother always opened the window a crack	even on	the coldest nights so my brother Larson and I would
26 ACAD:2003:	world , we would certainly consider a component worrisome if a crack	grew halfway	through its thickness . # Taking one micrometer
27 BLOG:2012:	referring to the late pop singer Whitney Houston as a "	ho "	in February in the wake of her untimely death .
28 TV:1994:	it 's more dangerous to go to church than to a crack	house	in There 's a woman behind St Lukes in a
29 ACAD:1997:	Software SmartCrack is a Windows based software for analysis of cracks	in elastic	materials . Stress intensity factors for some 50
30 TV:1998:	I guess when your heart gets broken , you see the cracks	in everything	in I guess when your heart gets broken , you

Texts page

Create Virtual Corpus

Freq	Text (click for topics)	Words
160	SPOK: NPR_Science SCIENCE, HISTORY AND CHEMISTRY OF CHOCOLATE M...	chocolate, cocoa, tree, bean, bird, coffee, flavor, migratory, crop, interesting, farmer, forest, tropical, rain, plantation, growing, wine, texture, ferment, cacao,
119	WEB: ...eyourgrandmother.com How To Make Perfect Brownies How To Cook Li...	brownie, recipe, chocolate, butter, chip, add, cocoa, ingredient, sugar, bake, perfect, cup, egg, texture, flour, powder, salt, pan, oven, batch,
69	NEWS: WashPost OOZE AND AAHS;	chocolate, cake, dessert, gm, mixture, egg, chef, pan, sugar, cream, serve, melt, pastry, butter, fat, mg, inn, ounce, cup, warm,
64	NEWS: WashPost Beyond Bittersweet; Once, any kid with a few ...	chocolate, cocoa, percent, bean, dark, bar, manufacturer, flavor, bitter, percentage, sugar, milk, plantation, origin, dark, product, specialty, growing, prefer, association,
61	MAG: CountryLiving Cook Book: A fine Romance	chocolate, cocoa, set, tablespoon, cup, sugar, beat, teaspoon, butter, bowl, minute, aside, vanilla, heat, pan, flour, baking, add, min, salt,
60	NEWS: USA Today A Valentine to Chocolate	chocolate, specialty, cocoa, praline, shop, information, belgian, butter, cream, bean, sweet, fill, sugar, filling, liquor, passion, box, supplier, demonstration, nearly,

WORDs FROM: NEWS: USA Today: A Valentine to Chocolate

Move between page showing the texts in which the word is a keyword (see above), and the page showing all keywords from a specific texts (right).

	Frequency	Word	PoS	Topics	Collocates
1	60	chocolate	n	Topics	Collocates
2	18	specialty	n	Topics	Collocates
3	16	cocoa	n	Topics	Collocates
4	15	praline	n	Topics	Collocates
5	13	information	n	Topics	Collocates
6	13	shop	n	Topics	Collocates
7	10	belgian	j	Topics	Collocates
8	10	liquor	n	Topics	Collocates
9	9	butter	n	Topics	Collocates
10	8	cream	n	Topics	Collocates
11	7	bean	n	Topics	Collocates
12	7	sweet	j	Topics	Collocates
13	6	fill	v	Topics	Collocates

Summary

COCA has a number of features that sets it apart from any other corpus. These include its **size** (1.0 billion words), how up-to-date it is (texts through Dec 2019), the wide range of **genres** (TV/Movie subtitles, spoken, blogs, web, fiction, magazine, newspaper, academic), and its **searches** (range of query types, and the ease and speed of its searches), including the ability to limit to and compare across genres and time periods.

In addition, it is different from most of the other corpora from English-Corpora.org in the attention that it gives to the **top 60,000 words** in the corpus, and the wide range of information for each word, including frequency information, definitions, synonyms, WordNet entries, related topics, concordances (new display in COCA), clusters, websites that have the word as a “keyword”, and KWIC/concordance lines.

All of these features make COCA the ideal corpus for researchers, teachers, and learners.