

THE CORPUS DEL ESPAÑOL

- | | |
|---|--------------------------------|
| 1. Basic searches | 5. Links to external resources |
| 2. Examining variation | 6. Virtual Corpora |
| 3. Browsing and searching top 40,000 words | 7. Analyzing entire texts |
| 4. Word sketches (for each of top 40,000 words) | 8. Summary |

The corpora from **Corpus del Español** provide billions of words of recent data from 21 Spanish-speaking countries, and they allow researchers, students, and teachers to gain insight into Spanish in ways that are not possible from any other resource. There are four different corpora in the Corpus del Español:

	Corpus	Size (# words)	Contents	Uses	Created
1	Web / Dialects *	2 billion	Web pages from 21 Spanish-speaking countries, about 50% blogs and 50% general	Comparing dialects	2016
2	NOW (2012 - 2019)	7.3 billion	Web pages from same 21 countries, 2012 to 2019	Changes; dialects	2018
3	Genre / Historical	100 million	Texts from 1200s to 1900s; genre-balanced for the 1900s (spoken, fiction, newspaper, academic)	Changes; genres	2001
4	Google Books	45 billion	Taken from Google Books n-grams. Not the entire text, but rather 1-5 word sequences.	Changes	2011

Of the four corpora, the one that allows the widest range of features is the Web/Dialects corpus (#1 above), and most of the discussion that follows deals with this corpus. The Web/Dialects corpus contains more than two billion words of data from 21 Spanish-speaking countries:

Country	General			Blogs			Total		
	Words	Pages	Web sites	Words	Pages	Web sites	Words	Pages	Web sites
Argentina	93,195,550	82,512	6,791	89,509,348	97,682	7,869	182,704,898	180,194	14,660
Bolivia	19,970,008	21,843	2,528	23,372,055	22,253	920	43,342,063	44,096	3,448
Chile	36,316,495	32,691	5,480	34,281,784	39,925	4,249	70,598,279	72,616	9,729
Colombia	84,285,729	84,349	5,064	95,859,929	104,658	5,855	180,145,658	189,007	10,919
Costa Rica	13,354,311	13,424	3,056	18,338,981	20,242	1,799	31,693,292	33,666	4,855
Cuba	35,842,945	21,666	1,209	31,812,745	32,028	677	67,655,690	53,694	1,886
Rep Dom	18,085,484	20,972	3,032	19,136,866	26,907	1,967	37,222,350	47,879	4,999
Ecuador	24,378,016	27,559	3,217	32,538,210	36,591	2,102	56,916,226	64,150	5,319
España	208,808,667	209,247	10,475	250,504,154	218,067	15,051	459,312,821	427,314	25,526
Guatemala	23,636,182	25,231	2,751	36,254,690	37,613	1,683	59,890,872	62,844	4,434
Honduras	16,420,245	18,129	1,653	22,512,838	26,109	787	38,933,083	44,238	2,440
México	132,651,925	135,344	6,713	127,946,347	156,877	8,379	260,598,272	292,221	15,092
Nicaragua	15,257,754	15,528	1,506	19,379,274	21,226	635	34,637,028	36,754	2,141
Panamá	11,766,575	13,364	2,728	12,932,194	16,355	1,147	24,698,769	29,719	3,875
Perú	47,120,271	53,169	4,262	68,204,165	71,162	4,275	115,324,436	124,331	8,537
Puerto Rico	14,744,742	13,340	2,769	20,911,898	20,946	1,545	35,656,640	34,286	4,314
Paraguay	17,223,991	16,102	2,200	15,766,153	17,791	1,017	32,990,144	33,893	3,217
El Salvador	16,546,308	18,266	2,102	22,600,727	20,538	2,076	39,147,035	38,804	4,178
EEUU	94,603,634	78,225	9,044	85,324,555	87,070	23,008	179,928,189	165,295	32,052
Uruguay	17,419,641	15,418	3,390	23,966,953	21,214	2,051	41,386,594	36,632	5,441

Venezuela	41,672,159	45,060	3,331	66,306,730	71,045	3,577	107,978,889	116,105	6,908
TOTAL	983,300,632	961,439	83,301	1,117,460,596	1,166,299	90,669	2,100,761,228	2,127,738	173,970

As far as comparisons with other corpora, the Web/Dialects corpus is about 10 times as large as CORPES, the largest and most recent corpus from the Real Academia Española, which is also based on web pages from 20+ countries. In addition, as we will see, the Corpus del Español (Web/Dialects) allows many types of searches that are [not possible with CORPES](#). There are other corpora of Spanish that are larger than the Corpus del Español, but there are [serious issues](#) with the annotation of these corpora, which makes them very problematic for many types of linguistic research. ([Click here](#) for an overview in Spanish for these three corpora)

Note that in the examples given below, we do not provide instructions on exactly what to enter in the search form to see these results. But there are many "context-sensitive" help files within the corpus itself, which show this type of information. For example, in the main search form, click on any search type (highlighted here in red) to see instructions and sample searches (shown here in green), and related help files (shown here in yellow).

List Chart Word Browse **Collocates** Compare KWIC -

GRUESO Word/phrase [POS] ?

NOUN Collocates [POS]

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates Reset

☐ Sections Texts/Virtual Sort/Limit Options

(HIDE HELP)

COLLOCATES display

See what words occur near other words, which provides great insight into meaning and usage. For example, nouns before **grueso** or after **contar con**, verbs before **brecha**, or any word near **garganta**, **iluminar**, **fuerte**, or **rápidamente**.

More information: [compare to LIST display](#), [types of collocates](#), [direction and distance of collocates](#).

1. Basic searches (go to beginning)

At the most basic level, the Corpus del Español allows users to search by word, phrase, part of speech, lemma (forms of a word), synonyms, and customized word list. In nearly all cases, the queries take just 1-2 seconds to search through the two billion words of data. For example:

NOUN SUAVE

HELP	①	★	ALL FORMS (SAMPLE): 100 200 500 WORDS	FREQ	TOTAL 19,643 UNIQUE 3,136 +
1	①	★	PIEL SUAVE	636	
2	①	★	VOZ SUAVE	616	
3	①	★	MÚSICA SUAVE	569	
4	①	★	FORMA SUAVE	451	
5	①	★	MANERA SUAVE	346	
6	①	★	TEXTURA SUAVE	277	
7	①	★	MOVIMIENTOS SUAVES	270	
8	①	★	FUEGO SUAVE	259	
9	①	★	EJERCICIO SUAVE	250	
10	①	★	COLORES SUAVES	218	

























=ESTUDIANTE =INTELIGENTE

HELP	①	★	ALL FORMS (SAMPLE): 100 200 500 WORDS	FREQ	TOTAL 215 UNIQUE 58 +
1	①	★	ALUMNOS CAPACES	28	
2	①	★	ESTUDIANTES CAPACES	27	
3	①	★	ESTUDIANTES INTELIGENTES	18	
4	①	★	ESTUDIANTE INTELIGENTE	16	
5	①	★	ALUMNOS INTELIGENTES	10	
6	①	★	ESTUDIANTE CAPAZ	9	
7	①	★	ALUMNO CAPAZ	7	
8	①	★	ALUMNO INTELIGENTE	6	
9	①	★	DISCÍPULOS ESPIRITUALES	6	
10	①	★	ESTUDIANTES LISTOS	5	

PRON PONER * @ROPA

HELP	①	★	ALL FORMS (SAMPLE) : 100 200 500 WORDS	FREQ	TOTAL 4,923 UNIQUE 1,960 +
1	①	★	SE PONGA LOS PANTALONES	67	<div></div>
2	①	★	SE PONE EL TRAJE	58	<div></div>
3	①	★	SE PUSO EL TRAJE	57	<div></div>
4	①	★	SE PONE LOS PANTALONES	55	<div></div>
5	①	★	SE PUSO LOS PANTALONES	44	<div></div>
6	①	★	ME PUSE UNA CAPA	34	<div></div>
7	①	★	SE PONE LAS BOTAS	34	<div></div>
8	①	★	SE PUSO UN VESTIDO	33	<div></div>
9	①	★	SE PUSO EL SOMBRERO	31	<div></div>
10	①	★	SE PONGAN LOS PANTALONES	29	<div></div>

For any of these searches (or any of the others shown in this overview), it is of course possible to click on any of the words / phrases to see the concordance lines:

CLICK FOR MORE CONTEXT			HELP 		 SAVE	 TRANSLATE	 ANALYZE
1	G US	slideshare.net			niego, me sigo negando, yo quiero a una Inés bella, de piel suave y pechos vivos, los contornosque mis manos siguen soñando, el pelo profundo,		
2	B PE	nosotras.wuole.com			consejos de cuidado de la piel y los secretos de belleza para conseguir una piel suave y radiante. Con unos pocos cambios en la dieta y estilo de vida		
3	B PA	ileanakeup.blogspot.com			y no causa irritación o sequedad. Está diseñado para: Minimizar los poros Piel suave Claro tez Reducir los brotes Restaura el equilibrio del pH salud		
4	G ES	...ntacionysalud.universiablogs.net			dorados, los ojos en forma de almendra y de color miel, una piel suave , de terciopelo. No como la mía que de tan reseca parece de cocodrilo		
5	B CO	fashioncreatortv.com			. Mito: La depilación reseca la piel. Falso: Depilarse deja la piel suave porque elimina la capa superior de piel muerta. Las ceras y cremas con ingred		
6	B PR	elblogiante.blogspot.com			se embarran de cremitas y cositas. Hay que lindo un macho con su piel suave y delicada. Me pregunto qué trabajo de hombre puede hacer, sin jode		
7	G PE	...uthamerica-travel-experience.com			las sirenas y el mas pequeño en la familia de los Trichechidae. Posee piel suave y generalmente una mancha o parche blanco o rosado en su estóm		
8	B AR	waterfrontmakeup.com			hacer lo. Sus pigmentos minerales se funden con la piel para lograr una piel suave , fresca y confortable, pero además uniforme, sin defectos, e incr		
9	B CR	mauspace.com			forma; de ambos los ojos, que son huraños y chispeantes, la piel suave y ajustada al cuerpo, los movimientos a la par saltarines y furtivos.		
10	B PY	madamalynch.blogspot.com			realzaban el primor de sus manos divinas; manos de eucarística blancura bajo cuya piel suave y transparente el azul de las venas evocaba rutas de		

And you can even see more context, if you want:

Source information:

Source	http://elblogiante.blogspot.com/2011/07/que-es-un-macho.html
Country/genre	Puerto Rico: Blog
Title	El Blogiante: Que es un macho?

Expanded context

otro. En la cultura humana, la masculinidad es el reflejo de la actividad hormonal que genera actitudes y cualidades viriles secundarias que no dejan lugar a dudas sobre la definición de el género. La parte clave aquí es No dejan lugar a dudas un hombre que se afeita las piernas, debe crear dudas. A menos que sea ciclista y se las afeite para curar sus heridas fácilmente con menos probabilidad de infección o practique algún deporte. Un hombre que se afeita las piernas porque le molestan los pelos, es un tipo de castración intangible. Hay otras áreas en donde molestan los arbustos y ahí es permisible según el manual de el macho cabrío. Ya que hablamos de pelos en la piel, que tal los que se embarran de cremitas y cositas. Hay que lindo un macho con su piel suave y delicada. Me pregunto qué trabajo de hombre puede hacer, sin joder el humectante que usa en sus cachetes. Lavar un carro? Uy no! después suda y se le llenan los poros de sucio y tiene que hacer se un facial. Trabajo de manufactura? Jamás! Eso jode las manos y da tunel carpal! Más te pone la piel dura y callosa, es como masturbar un flamboyán 40 horas semanales. En Los gustos, un macho va a el cine a ver sangre, acción, violencia, y a Megan Fox. Si el tipo quiere ir a ver twilight, sabes que de macho solo tiene el género. Un macho no quiere ver películas de amor, ni romances pendejos, un macho quiere una película en la que el protagonista acabe con una manada de zombis y la tipa quede eternamente enamorada de lo macho que es. Por eso es que Silvester Stallone es uno

You can also find "collocates" (nearby words), which provide insight into the meaning and usage of a word or phrase. For example:


























































































Collocates of GRUESO (all forms of grueso)

HELP	①	★	WORDS	FREQ	ALL	%	MI	
1	①	★	INTESTINO	1724	15185	11.35	9.51	<div></div>
2	①	★	CAPA	838	52947	1.58	6.66	<div></div>
3	①	★	POBLACIÓN	797	376391	0.21	3.76	<div></div>
4	①	★	CALIBRE	760	13679	5.56	8.47	<div></div>
5	①	★	PIEL	480	200777	0.24	3.94	<div></div>
6	①	★	LABIOS	395	47636	0.83	5.73	<div></div>
7	①	★	SAL	379	72886	0.52	5.06	<div></div>
8	①	★	PAREDES	355	58229	0.61	5.29	<div></div>
9	①	★	DELGADO	318	32810	0.97	5.96	<div></div>
10	①	★	PELO	291	105379	0.28	4.14	<div></div>
11	①	★	FINA	288	18817	1.53	6.61	<div></div>
12	①	★	ARMAS	282	191859	0.15	3.23	<div></div>

Collocates of *medio ambiente*

HELP	①	★	WORDS	FREQ	ALL	%	MI	
1	①	★	PROTECCIÓN	3942	220263	1.79	5.53	
2	①	★	MINISTERIO	2709	264932	1.02	4.72	
3	①	★	SALUD	2612	598529	0.44	3.49	
4	①	★	RECURSOS	2588	428311	0.60	3.96	
5	①	★	CUIDADO	2452	172636	1.42	5.19	
6	①	★	NATURALES	2446	171338	1.43	5.20	
7	①	★	MUNDIAL	1539	417635	0.37	3.25	
8	①	★	CONSERVACIÓN	1411	57157	2.47	5.99	
9	①	★	PRESERVACIÓN	1295	15002	8.63	7.80	
10	①	★	RESPECTO	1250	256934	0.49	3.65	
11	①	★	PROTEGER	1195	116196	1.03	4.73	
12	①	★	UNIDAS	1113	96210	1.16	4.90	

You can also see "concordance lines", which show the patterns in which a word occurs. For example, the following are sample concordance lines for RELUCIR:

63	B MX	diarioimagenqroo.mx	 	minuciosa de campesinos de la comunidad de Naranjal				un   en José Maria Morelos por pi	
64	B PR	viveaotronivel.com	 	y que lo distinguió alrededor del 				mi   , porque mientras escribo esta column.	
65	G ES	revistadelibros.com	 	, y con Céline , en los 				tres  ,  de Quer pasticciaccio . En la prin	
66	G HN	iglesia-nazareno.org	 	que es salvo , esto sin lugar a 				en su  . Usted no va ha guardar los mandamientos	
67	B GT	bethania.com	 	. Si no hay retos en tu 					tu  , tu fe , y tu ministerio . Nunca faltara
68	G VE	encuentroastrologicos.com	 	tenga conocimientos sobre la materia , lo primero 				   . Los interesados pedirán que	
69	B MX	papaesceptico.com	 	porque involucrar se en discusiones , la 					 o  . Saludos . Yo nadamas les recuerdo
70	B MX	hechospoliticos.com	 	. Pero ello no es todo , el 				de   . EL PAN ROMPE CON MALOVA P	
71	B CO	taotv.org	 	familia . Pero , la pistola de 					con el  de que el anuncio aparentemente espont
72	G US	lageneraciony.com	 	. . Andy # 2038 La Lajera : 				su  y  de nueva ricachona . Cuidese	
73	B ES	blogoff.es	 	es una cualidad infrecuente , me alegra 					.  no  4sq , principalmente porque no le he vi:

2. Examining variation (go to beginning)

The corpora from Corpus del Español allow you to easily and quickly examine variation in Spanish – between words, between genres, over time, and between dialects. As an example of **comparing words**, the following are the collocates of the synonyms *potente* and *poderoso*:

WORD 1 (W1): POTENTE (0.68)

	WORD	W1	W2	W1/W2	SCORE
1	CONSOLA	146	1	146.0	215.2
2	TOXINA	61	0	122.0	179.8
3	HERRAMIENTA	891	8	111.4	164.2
4	DELANTERO	47	0	94.0	138.6
5	INHIBIDOR	84	1	84.0	123.8
6	ORDENADOR	163	2	81.5	120.1
7	QUIMIOTERAPIA	77	1	77.0	113.5
8	ZURDAZO	38	0	76.0	112.0
9	PS4	32	0	64.0	94.3
10	GAMA	31	0	62.0	91.4
11	HORMONA	29	0	58.0	85.5
12	TORMENTA	28	0	56.0	82.5

WORD 2 (W2): PODEROSO (1.47)

	WORD	W2	W1	W2/W1	SCORE
1	PERDÓN	101	0	202.0	137.0
2	OZ	93	0	186.0	126.2
3	JEFE	92	0	184.0	124.8
4	MUNDIAL	85	0	170.0	115.3
5	NARCOTRAFICANTE	73	0	146.0	99.0
6	PROFETA	68	0	136.0	92.3
7	IMPERIO	648	5	129.6	87.9
8	TERRATENIENTE	63	0	126.0	85.5
9	CLAN	58	0	116.0	78.7
10	HECHICERO	51	0	102.0	69.2
11	INTERÉS	48	0	96.0	65.1
12	MINISTRO	47	0	94.0	63.8

The Historical/Genres corpus allows you to compare the frequency of words or phrases in the four **genres** of spoken, fiction, newspaper, and academic. For example, the following are the charts for three different words:

muchísimo

ACAD	NEWS	FICT	ORAL
7	68	97	764
5.0	5.0	4.8	4.2
1.40	13.70	20.34	180.48

borracho

ACAD	NEWS	FICT	ORAL
4	27	368	21
5.0	5.0	4.8	4.2
0.80	5.44	77.15	4.96

proporcionar

ACAD	NEWS	FICT	ORAL
205	42	10	25
5.0	5.0	4.8	4.2
41.00	8.46	2.10	5.91

You can also compare words, phrases, or collocates in the different genres. For example, the following is a comparison of verbs starting with *des** in fiction (left) and academic (right):

SEC 1 (FICT): 4,769,873 WORDS

	WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2	RATIO
1	DESAYUNAR	48	0	10.1	0.0	1,006.3
2	DESFILAR	27	0	5.7	0.0	566.1
3	DESANDAR	16	0	3.4	0.0	335.4
4	DESPERDICIAR	15	0	3.1	0.0	314.5
5	DESTAPAR	15	0	3.1	0.0	314.5
6	DESLIZAR	10	0	2.1	0.0	209.6
7	DESCORRER	9	0	1.9	0.0	188.7
8	DESAHOGAR	9	0	1.9	0.0	188.7
9	DESMENTIR	9	0	1.9	0.0	188.7
10	DESVANECER	8	0	1.7	0.0	167.7

SEC 2 (ACAD): 4,999,945 WORDS

	WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1	DESTITUIR	10	0	2.0	0.0	200.0
2	DESBLOQUEAR	5	0	1.0	0.0	100.0
3	DESTACAR	279	8	55.8	1.7	33.3
4	DESIGNAR	94	8	18.8	1.7	11.2
5	DESARROLLAR	331	32	66.2	6.7	9.9
6	DESEMPEÑAR	103	13	20.6	2.7	7.6
7	DESCODIFICAR	6	1	1.2	0.2	5.7
8	DESACTIVAR	6	1	1.2	0.2	5.7
9	DESPLAZAR	26	5	5.2	1.0	5.0
10	DESMANTELAR	13	3	2.6	0.6	4.1

You can also compare by **time period**. For example, the following are the frequencies of words over time:

gentilidad

1400s	1500s	1600s	1700s	1800s
20	219	453	95	5
9.7	19.7	14.8	11.5	23.1
2.45	12.86	36.69	9.68	0.26

deshonrado

1600s	1700s	1800s	1900s
10	5	9	4
14.8	11.5	23.1	22.8
0.81	0.51	0.47	0.18

controlar

1700s	1800s	1900s
0	1	2209
11.5	23.1	22.8
0.00	0.05	96.79

The NOW-Spanish corpus allows you to compare frequency in the short window from 2012-2019, to see recent change:

influencer

2016-1	2016-2	2017-1	2017-2	2018-1	2018-2	2019-1
346	582	1080	1790	2451	4270	7107
524.2	668.4	700.5	728.9	598.1	848.6	1,102.8
0.66	0.87	1.54	2.46	4.10	5.03	6.44

virtualizad*

2013-1	2013-2	2014-1	2014-2	2015-1	2015-2	2016-1	2016-2
81	65	98	48	62	43	48	68
201.0	230.9	261.9	294.3	331.1	363.6	524.2	668.4
0.40	0.28	0.37	0.16	0.19	0.12	0.09	0.10

Finally, the Web/Dialects corpus (as well as the NOW-Spanish corpus, which has texts from the same countries) allows you to compare the frequency of words and phrases in **21 Spanish-speaking countries**. The following charts are from just a couple of examples:

muy padre: México

MX	GT	SV	HN	NI	CR	PA	PR	DO	CU	VE	CO	EC	BO	PE	CL	PY	UY	AR	ES	US
502	8	9	4	7	6	3	5	10	12	14	13	7	2	20	12	0	2	12	72	55
246.0	54.3	36.5	35.1	32.4	29.6	22.3	32.2	33.7	63.2	98.2	166.5	52.4	39.4	107.3	66.2	29.8	38.7	169.4	426.6	166.0
2.04	0.15	0.25	0.11	0.22	0.20	0.13	0.16	0.30	0.19	0.14	0.08	0.13	0.05	0.19	0.18	0.00	0.05	0.07	0.17	0.33

POLOLO: Chile

MX	GT	SV	HN	NI	CR	PA	PR	DO	CU	VE	CO	EC	BO	PE	CL	PY	UY	AR	ES	US
192	8	16	3	0	4	2	2	17	4	28	89	15	14	36	2361	8	11	57	228	110
246.0	54.3	36.5	35.1	32.4	29.6	22.3	32.2	33.7	63.2	98.2	166.5	52.4	39.4	107.3	66.2	29.8	38.7	169.4	426.6	166.0
0.78	0.15	0.44	0.09	0.00	0.14	0.09	0.06	0.50	0.06	0.29	0.53	0.29	0.36	0.34	35.65	0.27	0.28	0.34	0.53	0.66

ORDENADOR: Spain (vs computadora, which is more common in Latin America)

MX	GT	SV	HN	NI	CR	PA	PR	DO	CU	VE	CO	EC	BO	PE	CL	PY	UY	AR	ES	US
8564	833	705	667	686	439	540	250	810	2201	1674	3275	1207	868	2117	891	565	970	2334	50482	6342
246.0	54.3	36.5	35.1	32.4	29.6	22.3	32.2	33.7	63.2	98.2	166.5	52.4	39.4	107.3	66.2	29.8	38.7	169.4	426.6	166.0
34.82	15.35	19.33	18.99	21.18	14.84	24.26	7.77	24.04	34.80	17.05	19.67	23.05	22.04	19.73	13.45	18.99	25.04	13.78	118.34	38.20

And searches for any of the following words will show similar "spikes" by country.

Puerto Rico *ay bendito, chavos, chiringa, mahones, habichuela* Cuba *guajiro, jimaguas, babalao, bitongo, pedir botella* Rep Dom *mangú, fucú, tutumpote, mangulina*

México *ándale, hijole, órale, güero, (muy) padre, pinche (NOUN), popote, charola* Guatemala *huipil, canche, muchá, patojo, chirmol, canche* El Salvador *cipote, chero, pupusa, cuilio, bayunco, piscucha* Honduras *catracho, papada* Nicaragua *chavalo, pinol, pinolillo, chigüín, vigorón* Costa Rica *chinear, guila, chunche* Panamá *fulo, chombo, guandul*

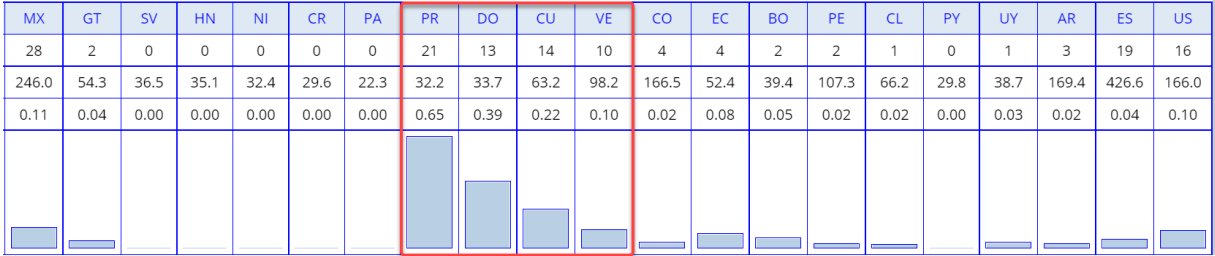
Colombia *cachaco, cachifo, verraquera, estar mamado, guandoca, biche* Venezuela *bojote, coroto, catire, gafo, macundales, arepa, cachapa, cambur, caraotas, jojoto* Ecuador *chumar, chulla, montuvio, omoto* Perú *anticucho, jebe, chupe, pisco, jora* Bolivia *opa, colla, chuño, lagua* Chile *pololo*, pololear, achuntar, bencina, bacán, fome, huaso* Paraguay *ñembo, ñanduti, karai, yopará, mitai* Uruguay *tropero, hacer *sota, con fritas* Argentina *pibe, fiaca, morfar, falopa, sobre el pucho, falluto, cafishio*

España *ordenador, aparcar, enfadar, gafas, zumo, chulo, guay, coger, bolígrafo, patata, melocotón, echar de menos, vale*

The preceding are just single words or short phrases. You can also use this corpus to compare syntactic and morphological features across dialects. For example:

Search string	Example	Where most common
qué tú VERB	<i>¿qué tú quieres?</i>	Caribbean
PREP SUBJ VERB	<i>para ella entender</i>	Caribbean
ART POSS NOUN	<i>una mi amiga</i>	Guatemala (+C Am)
voseo	<i>vos sos</i>	Cono Sur (+C Am)
la las GUSTAR (laísmo)	<i>la gusta el chocolate</i>	España

As an example, the following is the frequency in the corpus for the string *qué tu VERB*, which is more common in the Caribbean region, followed by some of the actual example from Puerto Rico:



1	B PR	blog.cervtech.com	Q	lista de libros sobre Cyberlaw? Relajando le dije al colega: ¿ Para qué tu quieres aprender de Cyberlaw si me tienes a mi? A lo cual me
2	B PR	elpuntos.com	Q	varias veces contestar con sus nombres de pila, Juan le consulta a Alexis ¿ qué tu crees de la pandilla?. Alexis estuvo de acuerdo y el resto es
3	B PR	purrode aqui.com	Q	.. digo, si es que estás dentro del Purismo... Qué tu crees de nuestra Raza... digo, si es que crees que
4	B PR	enhorabuena.net	Q	dice esto y lo otro. Sí, la Biblia dice eso, pero ¿ qué tu piensas al respecto? Es increíble cómo somos tan descarados de limitar a Dios
5	G PR	cuatro-pr.org	Q	yo fui a comprar en una ferretería una herramienta, me dice, ¿ para qué tu quieres eso? y le digo, es que yo construyó guitarras. y
6	B PR	tecnetico.com	Q	que yo tendría con mi computadora llamada Neti (tecNETIca): Neti, ¿ qué tu crees que mi novia estará de acuerdo en que no tengamos recepción y que
7	B PR	blasfemiasyrecuerdos.blogspot.com	Q	me ocurrió preguntar le su pensar en todo este asunto. Oye tío, ¿ qué tu piensas de todas esas farmacéuticas que no quieren que tu te fumes eso que
8	G PR	normavalle.com	Q	modestia aparte, cantaba bien. La gente me decía' Ruth, ¿ por qué tu gastas tanto en ropa? Y yo decía esa es mi pelea,
9	B PR	foodietiendo.com	Q	presentación, servicio etc). ¿ Ya viste los nominados y ganadores? ¿ Qué tu opinas ? Deja tu comentario bajo este post o compártelo con tus amistades or
10	B PR	cientecblog.com	Q	porque soy escogido? ¿ Por qué soy más matemático que tú? ¿ Por qué tu eres mejor músico que yo? ¿ Tu sabes siete idioma y yo apenas
11	B PR	arrozconpinga.blogspot.com	Q	que decir y condenado en este mundo / Hear me now!!! ¿ Qué tu pensaste ? ¿ Qué todo terminaba? / Estupido, pendejo, si la
12	G PR	blog.masslive.com	Q	mi congregación, en esta portada dice que es mejor dar que recibir, ¿ Qué tu prefieres , dar o recibir? Desde luego, esa pregunta no es fácil

In the examples above, the search is for a particular word or phrase, which you have already selected. But the corpus can also show you **all matching words, phrases, or collocates that are more common in one country (or set of countries) than another**. For example, the following are **ismo* words that are more common in Venezuela (left) than in Colombia, México, or España (right):

SEC 1 (Venezuela): 98,170,248 WORDS							SEC 2 (Argentina, Colombia, España...): 1,008,426,240 WORDS						
	WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2	RATIO		WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1	JALABOLISMO	44	1	0.4	0.0	452.0	1	MASSISMO	1249	1	1.2	0.0	121.6
2	ESCUALIDISMO	21	1	0.2	0.0	215.7	2	PRIISMO	362	0	0.4	0.0	35.9
3	MADURISMO	141	10	1.4	0.0	144.8	3	EMPRENDEDORISMO	251	0	0.2	0.0	24.9
4	GOMECISMO	41	3	0.4	0.0	140.4	4	GARANTISMO	254	1	0.3	0.0	24.7
5	VENTAJISMO	812	63	8.3	0.1	132.4	5	PARKINSONISMO	219	0	0.2	0.0	21.7
6	PUNTOFIJISMO	165	14	1.7	0.0	121.1	6	NEOPLATONISMO	218	1	0.2	0.0	21.2
7	RIQUISMO	22	3	0.2	0.0	75.3	7	CHARRISMO	204	1	0.2	0.0	19.9
8	HIPISMO	295	41	3.0	0.0	73.9	8	SOCORRISMO	196	0	0.2	0.0	19.4
9	CASTRO-COMUNISMO	55	8	0.6	0.0	70.6	9	LIBERTARISMO	184	0	0.2	0.0	18.2
10	ECO-SOCIALISMO	27	4	0.3	0.0	69.3	10	CRISTINISMO	557	3	0.6	0.0	18.1

And the following are collocates of MANEJAR that are more common in either México (left) or España (right):

SEC 1 (México): 245,956,621 WORDS

SEC 2 (España): 426,578,900 WORDS

	WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2	RATIO		WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1	POLÍTICAS	109	1	0.4	0.0	189.0	1	FECHA	102	3	0.2	0.0	19.6
2	CAMIONETA	31	1	0.1	0.0	53.8	2	COTARRO	98	3	0.2	0.0	18.8
3	RUMORES	114	4	0.5	0.0	49.4	3	SOLTURA	29	1	0.1	0.0	16.7
4	EDUCACIÓN	25	1	0.1	0.0	43.4	4	ENERGÍAS	33	2	0.1	0.0	9.5
5	TRANSPARENCIA	21	1	0.1	0.0	36.4	5	ORDENADORES	24	3	0.1	0.0	4.6
6	OBJECIONES	20	1	0.1	0.0	34.7	6	TELÉFONOS	21	3	0.0	0.0	4.0
7	CARRETERA	58	3	0.2	0.0	33.5	7	CLUB	26	4	0.1	0.0	3.7
8	DISCURSO	26	2	0.1	0.0	22.5	8	HILOS	239	42	0.6	0.2	3.3
9	AUTOMÓVIL	62	6	0.3	0.0	17.9	9	CÁMARA	79	14	0.2	0.1	3.3

3. Browsing and searching through the top 40,000 words (go to beginning)

In the Web/Dialects corpus, we have created a carefully reviewed list of the top 40,000 words in the two billion word corpus, and you can search through these words – by word form, part of speech, word frequency, and (basic) English translation. For example, the following are a few words at three different **frequency** bands: near 4200 (i.e. the 4,200th most frequent word (lemma) in the corpus), near 16,000, and near 35,000:

	RANK #	PoS	WORD	MORE DEFIN	IMAGE	VIDEO	WEB		ACAD	NEWS	FIC	SPOK
1	4206	J	HÚMEDO	RAE / L R / W C S O			26885		323	44	520	25
2	4207	R	ORIGINALMENTE	RAE / L R / W C S O			26864		131	47	7	30
3	4208	N	HEMBRA	RAE / L R / W C S O			26862		376	56	138	46
4	4209	V	CHOCAR	RAE / L R / W C S O			26848		79	109	182	88
5	4210	J	SINGULAR	RAE / L R / W C S O			26828		146	137	105	27
6	4211	N	CUERO	RAE / L R / W C S O			26800		136	31	276	23
7	4212	J	SOBERANO	RAE / L R / W C S O			26798		97	41	32	22

	RANK #	PoS	WORD	MORE DEFIN	IMAGE	VIDEO	WEB		ACAD	NEWS	FIC	SPOK
1	16028	N	COHECHO	RAE / L R / W C S O			2247		0	4	2	3
2	16029	V	PODAR	RAE / L R / W C S O			2247		7	2	22	6
3	16030	N	RADICALIZACIÓN	RAE / L R / W C S O			2247		11	3	0	7
4	16031	N	PEREGRINAJE	RAE / L R / W C S O			2247		8	14	10	1
5	16032	N	JOYITA	RAE / L R / W C S O			2247		0	3	0	0
6	16033	V	REPUGNAR	RAE / L R / W C S O			2246		1	6	27	10
7	16034	J	GNÓSTICO	RAE / L R / W C S O			2246		6	0	0	0

	RANK #	PoS	WORD	MORE DEFIN	IMAGE	VIDEO	WEB		ACAD	NEWS	FIC	SPOK
1	35005	N	GORRÓN	RAE / L R / W C S O			244		0	1	3	2
2	35006	N	CARGADORA	RAE / L R / W C S O			244		0	0	0	2
3	35007	N	CACHIPORRA	RAE / L R / W C S O			244		1	1	11	0
4	35008	J	PROGUBERNAMENTAL	RAE / L R / W C S O			244		1	2	0	0
5	35009	J	DESLIGADO	RAE / L R / W C S O			244		1	2	2	1
6	35010	J	CAPSULAR	RAE / L R / W C S O			244		0	0	2	0
7	35011	N	VULCANÓLOGO	RAE / L R / W C S O			244		0	6	1	2

The following are lower frequency **ista* words (lower than #20,000 in the list):

	RANK #	PoS	WORD	MORE DEFIN	IMAGE	VIDEO	WEB	ACAD	NEWS	FIC	SPOK
1	20035	J	DETERMINISTA	RAE / L R / W C S O			1308	3	4	4	2
2	20071	N	SEPARATISTA	RAE / L R / W C S O			1302	9	2	1	2
3	20099	N	RACIONALISTA	RAE / L R / W C S O			1298	96	18	16	8
4	20194	J	SINDICALISTA	RAE / L R / W C S O			1280	5	20	3	1
5	20215	N	DIETISTA	RAE / L R / W C S O			1277	0	0	3	26
6	20233	N	ARREGLISTA	RAE / L R / W C S O			1274	3	0	1	0
7	20293	N	MARCHISTA	RAE / L R / W C S O			1266	0	11	0	0

And the following are words that have *leather* in their basic English gloss:

	RANK #	PoS	WORD	DEFINITION	MORE DEFIN	IMAGE	VIDEO	WEB	ACAD	NEWS	FIC	SPOK
1	845	N	PIEL	skin, fur, peel, leather , hide, (calf)	RAE / L R / W C S O			207092	592	179	1245	128
2	4211	N	CUERO	leather , hide, skin, whip, washer, pelt, (ball)	RAE / L R / W C S O			26800	136	31	276	23
3	15091	J	CURTIDO	hardened, tanned, leathery , tanning, weather-beaten, (experienced)	RAE / L R / W C S O			2563	56	12	32	6
4	20014	N	CHUPA	leather jacket, (drink)	RAE / L R / W C S O			1311	0	2	1	3
5	21297	N	NAPA	leather , (tip)	RAE / L R / W C S O			1119	1	0	1	3
6	25721	N	GAMUZA	chamois, duster, (leather)	RAE / L R / W C S O			676	12	5	13	1
7	26250	N	CHAROL	patent leather , (tray)	RAE / L R / W C S O			639	2	3	40	3

For each of the words in the list, you can see the rank order (1-40,000), part of speech, basic English gloss, links to several online **dictionaries**, Google **images**, **video** from Youglish.com (Spanish), overall **frequency** and by **genre** in the Historical/Genres corpus. And of course, there is also a link to the "word sketch" for that word (see the following section). For example, the following are images, videos, and definitions for *charol*, one of the words in the list above:

charol

Del port. *charão*, y este del chino *chat liao*.

- m. Barniz muy lustroso que se adhiere perfectamente a la superficie.
- m. Cuero con **charol**. *Botas de charol*.
- m. *Col. y Ec. bandeja* (|| pieza para servir).

darse charol

- loc. verb. coloq. Alabarse, darse importancia.

El calzado de **charol** para señoras estaba disponible en los colores de moda del momento.

La gran mayoría del **charol** se produce a partir de grano en crust corregido.

Pero en Shanghai habrá un nuevo y magnífico efecto, **el charol blanco**.

A excepción del **charol** y otras pieles también muy brillantes en las que el objetivo es alcanzar un brillo lo más alto posible, [...]

[...] impermeables como en aquéllos con acabados más elaborados como el **charol**, donde la capa de superficie podría poseer mayor fuerza [...]

Ladies **Patent** leather footwear became available in the fashion colours of the day.

The vast majority of **Patent** leathers are produced on corrected grain crust.

But at Shanghai there will be a brilliant new effect, **Patent white**.

With the exception of **patents** and other very bright leathers where the gloss target is set as high as practicable, ordinary [...]

[...] on both waterproof and heavily finished leathers such as **patent**, where the surface coating may possess greater film strength [...]



en **charol**, la mitad en gamuza y tiene un

4. Word sketches (go to beginning)

For each of the top 40,000 words (lemmas) in the corpus, you can see a wide range of information, including the following: [1] word (lemma) [2] frequency information (including rank, #1-40,000) [3] links to images and videos (with pronunciation) [4] simple English glosses, and links to entries in detailed monolingual dictionaries (e.g. from the Real Academia Española, with etymology) [5] frequency by genre (from the Historical/Genres corpus) [6] add word to "favorites" list (where you can sort and group the words) and History (to review recent words) [7] synonyms (you can click on any synonym, collocate, or topic to see the page for that word) [8] collocates (nearby words; 4 left to 4 right) [9] topics (words that co-occur anywhere in the 2,000,000+ texts in the corpus) [10] frequency by dialect [11] frequency of forms of the word (perhaps most useful for verbs), and [12] 200 sample re-sortable concordance lines.

1

2

3

FAROL *n* (RANK 13505, FREQ 3287)

4

5

DEFINITIONS (click for more)

lamp, bluff, lantern, headlamp, street, lamp, (light)

RAE Ling WR Rev Coll SpD Oxf

6

7

SYNONYMS (click to see word)

fanal *n* **lantern** luz, lámpara, foco, reflector, proyector, faro, linterna,
faro *n* **lighthouse** faro,
farola *n* (**lamp**) foco, faro, (tulipa), linterna, luz, lámpara, fanal,
foco *n* **focus** proyector, faro, luz, fanal, linterna, reflector,
fuego *n* **light** candil, vela, bombilla, foco, faro, luz,
hornillo *n* **stove** hogar, cocina, fogón, jaula, (fogaril),
lámpara *n* **light** velador, candelero, (palmatoria), lamparilla, candelabro, velón,
lamparilla *n* (**aspen**) bombilla, foco, fanal, candil, linterna, lámpara,
lance *n* **incident** jugada, engaño, trampa, truco, argucia,
luz *n* **electricity** ascua, lámpara, (alcandora), candil, faro, luminaria,

8

9

COLLOCATES

(ADJ) encendido, rojo, tenue, amarillento, colonial, apagado, delantero, trasero, mortecino, iluminado
(NOUN) luz, calle, vela, gas, farol, esquina, desfile, mantilla, lámpara, alumbrado
(MISC) encender, iluminar, alumbrar, apagar, tirar, colgar, prender, envidar, adornar, besar

10

9

TOPICS (click to see)

encender *v* desfile *n* festival *n* oscuro *j* tembloroso *j* iluminar *v* marinero *n* apagar *v*
calle *n* balcón *n* chimenea *n* farolero *n* luz *n* vela *n* vapor *n* patio *n* setiembre *n* reloj *n*
sombra *n* planeta *n* murmurar *v* labios *n* inmóvil *j* entretanto *r* don *n* castillo *n*
alumbrado *n*

11

FORMS


1,698 farol ☐ 1,381 faroles ☐

CONCORDANCE LINES PROBLEMS?


GENRE		12			SORT		
33	B AR	chachara.net	y ella había reaccionado mal y había roto una cámara y	faroles	de un auto	y me añadió que eran fotografías míos , contó	
34	B VE	mujerxmujer.com	de tanto amor . Que si caminamos juntas , nos volvemos	farolas	de una sola	calle en una sola dirección Que no hay escaleras	
35	B GT	guatemaladeayer.blogspot.com	una de estas pasarelas en el pasado estuvieron iluminadas con	24 faroles	de vidrio grandes	y pequeños de diferentes colores , con formas	
36	B ES	cronicasdeunmundofeliz.com	de el mal . A mi el alcalde me pone una	farola	delante de casa	y no me pregunto si la farola se la	
37	G AR	jfaltasescobar.gov.ar	, cada cuatro (4) metros a contar desde el	farol	delantero hasta el	largo de el vehículo . Las luces serán de	
38	G UY	kombiuruguay.wordpress.com	hacer si son profundas . 5 - Los Plásticos de los	faroles	descoloridas o las	ópticas oxidadas por dentro , son piezas que	
39	G US	missionlocal.org	atroz cerca de dicha intersección y muchos vecinos culparon a el	farol	descompuesto y la	lenta respuesta de la compañía municipal de	

For [3] the images and videos, it will take you to external resources:


Lighthouse




File:Farol de Galinhos, Rio Gra...
commons.wikimedia.org




Lighthouse of Pr...
en.wikipedia.org



Itapua Lighthouse (S...
tripadvisor.com





240 ideias de Fa...
pinterest.com





Calcanhar Li...
en.wikipedia.o


Lantern












Francisca Serrano en "La Aventura del Saber" de RTVE (Vivir del trading)



9:17 / 11:48

Speed: normal

trolera y que me estoy marcando un farol muy gordo y entonces le digo vale

You can also find "random" words (via the "dice" icon at the top of the corpus) and can even set the range of the random words, such as words from 15,000-20,000 in the frequency list (which may be helpful for language learners).

Corpus del Español: Web/Dialects

LIST OF WORDS WORD CONTEXT OVERVIEW

BOLLO *n* (RANK 15304, FREQ 2490)

0k 1k 3k 6k 10k 15k 20k

DEFINITIONS (click for more)

bun, bread, roll, bump, dent, scone, mix-up, (dish)

RAE Ling WR Rev Coll SpD Oxf

SPOK FIC NEWS

bollo, confederación, diptongo, nirvana, resucitación, encarcelamiento, carbonifero, fluorescencia, totalitario, coralino

SYNONYMS (click to see word)

arimez *n* (-----) (bocel) botón, ménsula, (mediacaña), moldura, resalto,
bocadillo *n* balloon canapé, panecillo, medianoche, bocado, tentempié, (piscolabis), emparedado,
bollo *n* bun trena,
bulto *n* bulge abolladura, prominencia, resalto, bombeo,
dulce *n* candy pasta, masa, torta, hojaldre, tarta, pastel,

And you can also see "Words of the Day" from different frequency bands:

Click below to see detailed information on word

Frequency band	Rank #	Word	PoS
1 1-999	199	escribir	v
	644	campana	n
2 1000-2999	1572	extranjero	n
	2733	arquitectura	n
3 3000-5999	4395	colar	v
	5302	paramilitar	j
4 6000-9999	6111	acogida	n
	8842	comisionado	n
5 10000-14999	12016	rosca	n
	13740	desborde	n
6 15000-19999	15484	diseminación	n
	16382	rizar	v
7 20000-29999	21974	anilla	n
	22742	capitel	n

22-05-09 ▶ 22-05-11

"Words of the Day" for [22-05-10]

The 20 words to the left were selected at random by the computer from the top 40,000 words in SPAN, with two words in each of the seven frequency ranges.

You might want to come back to this page every day and see how many of the words you know. You can also expand your vocabulary by focusing on words that are unfamiliar, and then "save" these words for later review via ☆ on the "word" page .

You can also click below to see the words for past days, from 21-04-01 through 22-06-09 (one month from today).

Search by date

To get to this page in the future, click on 📅 and then 📅 .

You can also find many random words for each of the ten frequency ranges (e.g. words 3000-6000) via the "word page".

Or browse to focus on an even more narrow frequency band, or to search by word form, pronunciation, meaning, synonyms, etc.

5. Links to external resources (including translations) (go to beginning)

As is shown in the "browsing" and "word sketches" sections above, there are many links to external resources. This is true for other pages at the corpus as well. For example, in the basic "results" page after a search, the default action is to see concordance lines for the word, phrase, or collocate. But you can also choose to see external resources for the entry – translation (100+ languages from Google Translate), Google search, video/pronunciation, or Google Books search

ON CLICK:

CONTEXT

TRANSLATE (EN)

ENTIRE PAGE

GOOGLE

IMAGE

PRON/VIDEO

BOOK

HELP	1	★	ALL FORMS (SAMPLE): 100 200 500 WORDS	FREQ	TO
1	1	★	PIEL SUAVE	636	
2	2	★	VOZ SUAVE	616	
3	3	★	MÚSICA SUAVE	569	
4	4	★	FORMA SUAVE	451	
5	5	★	MANERA SUAVE	346	
6	6	★	TEXTURA SUAVE	277	
7	7	★	MOVIMIENTOS SUAVES	270	
8	8	★	FUEGO SUAVE	259	

Images

Video

How to pronounce **textura suave** in Spanish (1 out of 48):

Como hacer Arepas Venezolanas PASO A PASO (con RELLENO MEXI...

make a dough of soft texture or that do not feel

se haga una masa de **textura suave** o que no se sientan los

Book

books.google.com · books

Cocina Betty Crocker: Recetas Americanas Favoritas en Español ...

Betty Crocker · 2005

FOUND INSIDE – PAGE 107

Relleno de Pan Bread Stuffing PREPARACIÓN : 15 min COCCIÓN : 5 min

HORNAR : 40 min • RINDE 10 PORCIONES, 1/2 TAZA CADA UNA Para obtener un relleno más suave y húmedo , utilice pan de **textura suave** ; si desea un relleno más firme y ...

Preview

books.google.com · books

El arte de la comida sencilla: Notas, lecciones y recetas de ... - Page 193

Alice Waters · 2013

FOUND INSIDE – PAGE 193

El escarchado tiene una agradable textura granulada, mientras que un sorbete se congela en una máquina de hacer hielo, dándole una **textura suave** y aterciopelada. Las frutas y el azúcar son los ingredientes básicos de los sorbetes y ...

You can also see a translation of all of the entries on the page:

ON CLICK:

CONTEXT

TRANSLATE (EN)

ENTIRE PAGE

GOOGLE

HELP	1	★	ALL FORMS (SAMPLE): 100 200 500 WORDS
1	1	★	PIEL SUAVE
2	2	★	VOZ SUAVE
3	3	★	MÚSICA SUAVE
4	4	★	FORMA SUAVE
5	5	★	MANERA SUAVE
6	6	★	TEXTURA SUAVE
7	7	★	MOVIMIENTOS SUAVES
8	8	★	FUEGO SUAVE

Choose a language to translate to:

English

Submit

Danish

Dutch

English

Esperanto

Estonian

Filipino

Finnish

Google Translate

Spanish → English

636 piel suave	soft skin
616 voz suave	soft voice
569 música suave	soft music
451 forma suave	smooth shape
346 manera suave	soft way
277 textura suave	smooth texture
270 movimientos suaves	smooth movements
259 fuego suave	soft fire
250 ejercicio suave	gentle exercise
218 colores suaves	Soft colors
217 sabor suave	Smooth flavor

In the context / concordance / KWIC page, you can see a translation of any line of text [1], as well as hear the pronunciation of the Spanish text [2]:

3	B MX	serpapas.com	se pone de rodillas y gatea largas distancias es posi
4	G AR	alimentacion-sana.com.ar	ser alérgicas, por lo que siempre es mejor frotar un
5	B AR	site.informadorpublico.com	era lindo. tiene 21 años. un morochito con lindos ra

SPANISH - DETECTED SPANISH PORTUGUESE ENGLISH ↔ SPANISH ENGLISH GERMAN ↕

ser alérgicas, por lo que siempre es mejor frotar un poco en la piel suave de el antebrazo, y si eres alérgico a los pocos minutos se siente un

be allergic, so it is always better to rub a little on the soft skin of the forearm, and if you are allergic to a few minutes you feel a

143 / 5,000

And finally, you can see the translation from the "Expanded Context" page:

Google Translate Spanish - detected → English

Translation

diariamente el aloe como un ingrediente esencial en su cuidado diario. Uso interno Hay gente que toma una o dos cucharadas de zumo de aloe varias veces a la semana. Las propiedades que se le atribuyen son, entre otras: Actúa como anti-inflamatorio, analgésico, antiviral, antitoxico Previene y reduce la inflamación artrítica. En úlceras gastrointestinales, tras la atención médica precisa, puede tomar se jugo de aloe cuatro veces a el día antes de las comidas y a el acostar se. Trata los problemas intestinales y facilita la digestión por su alto contenido en encimas. Reduce el azúcar en la sangre por diabetes Precauciones El aloe carece de efectos secundarios y no es tóxico, pero menos de un 1 % de las personas pueden ser alérgicas, por lo que siempre es mejor frotar un poco en la piel suave de el antebrazo, y si eres alérgico a los pocos minutos se siente un pequeño picor y algo de erupción. También se pueden derivar efectos indeseables por su mal uso. Siempre es necesario consultar antes de dejar el tratamiento médico. Así mismo, ante cualquier afectación que se pueda beneficiar de el aloe, primero es aconsejable que sea atendida y valorada por el médico. El uso de el aloe, tanto por vía interna como externa, es prácticamente siempre compatible con cualquier otro tratamiento. No emplear durante el embarazo ya que los glúcidos de la antraquinona son fuertemente purgantes. Las hojas en grandes dosis pueden producir vómitos. FOSFATO DE MANOSA: Actúa como agente de crecimiento de los tejidos con efecto cicatrizante MINERALES: calcio, magnesio, fósforo, potasio, zinc, cobre MUCILAGO: actividad emoliente sobre la piel SAPONINAS: antiséptic Existen varias formas de usar el Aloe Vera (éstas dependen de su finalidad)

daily aloe as an essential ingredient in your daily care. Internal use There are people who take one or two tablespoons of aloe juice several times a week. The properties attributed to it are, among others: It acts as an anti-inflammatory, analgesic, antiviral, antitoxic Prevents and reduces arthritic inflammation. In gastrointestinal ulcers, after proper medical attention, you can take aloe juice four times a day before meals and at bedtime. It treats intestinal problems and facilitates digestion due to its high content of enzymes. Lowers blood sugar due to diabetes Cautions Aloe has no side effects and is non-toxic, but less than 1% of people may be allergic, so it's always best to rub some into smooth skin of the forearm, and if you are allergic, after a few minutes you feel a little itching and some rash. Undesirable effects can also be derived from its misuse. It is always necessary to consult before stopping medical treatment. Likewise, in the event of any affectation that can benefit from aloe, it is first advisable that it be attended to and evaluated by the doctor. The use of aloe, both internally and externally, is practically always compatible with any other treatment. Do not use during pregnancy since anthraquinone carbohydrates are strongly purgative. The leaves in large doses can cause vomiting. MANNNOSE PHOSPHATE: Acts as a tissue growth agent with a healing effect MINERALS: calcium, magnesium, phosphorous, potassium, zinc, copper MUCILAGE: emollient activity on the skin SAPONINS:

6. Virtual Corpora (go to beginning)

You can quickly and easily create "Virtual Corpora", which is like a "corpus within a corpus". For example, you can find articles with the phrase *medio ambiente*, or articles from Spain in the "General" (non-blog) part of the Web/Dialects corpus, which have *nuclear* in the title:

List Chart Word Browse +

medio ambiente [POS] ?

Find matching strings Reset

☐ Sections **Texts/Virtual** Sort/Limit Options

FIND TEXTS

MY CORPORA

inversiones

medioambiente

nuclear

Create corpus

Edit corpora

Find keywords

Refresh list

Web domain	<input type="text"/>	Find sources
	(can use substring, e.g. *futbol*, *.mx)	
Article title	<input type="text" value="nuclear"/>	
Country	<div> Costa Rica Cuba Rep Dom Ecuador España </div>	
Genre	<input type="text" value="General"/>	
Words in text	<input type="text"/>	
<input type="button" value="Submit"/> <input type="button" value="Reset"/>		

One or two seconds later, it shows a list of the best matching texts. Just name the corpus, and you've created the Virtual Corpus.

(Optional) SAVE AS: medio_ambiente		OR ADD TO: --SELECT--	SUBMIT	RESET	HELP			
HELP	<input type="checkbox"/>	100	TEXT	SEE PAGE	# WORDS	# HITS ↓	RELEVANCE ↑	PER MILLION WORDS
1	<input checked="" type="checkbox"/>		G ES: PROTECCIONCIVIL.ORG		455421	424	931.0	<div></div>
2	<input checked="" type="checkbox"/>		G DO: SUPREMA.GOV.DO		23395	263	11,241.7	<div></div>
3	<input checked="" type="checkbox"/>		G US: SOCIALWATCH.ORG		132018	239	1,810.4	<div></div>
4	<input checked="" type="checkbox"/>		G CU: GACETAOFICIAL.CU		15686	184	11,730.2	<div></div>
5	<input checked="" type="checkbox"/>		G ES: JUNTADEANDALUCIA.ES		45148	175	3,876.1	<div></div>
6	<input checked="" type="checkbox"/>		G MX: GLOOBAL.NET		89810	162	1,803.8	<div></div>
7	<input checked="" type="checkbox"/>		G CO: SECRETARIASENADO.GOV.CO		30657	145	4,729.8	<div></div>
8	<input checked="" type="checkbox"/>		G CU: UN.ORG		47072	114	2,421.8	<div></div>
9	<input checked="" type="checkbox"/>		G ES: HABITAT.AQ.UPM.ES		51261	91	1,775.2	<div></div>
10	<input checked="" type="checkbox"/>		G GT: TESIS.UFM.EDU.GT		19105	85	4,449.1	<div></div>
11	<input checked="" type="checkbox"/>		G ES: CMA.GVA.ES		16695	84	5,031.4	<div></div>
12	<input checked="" type="checkbox"/>		G AR: CICODI.ORG		24461	73	2,984.3	<div></div>
13	<input checked="" type="checkbox"/>		G CO: DEFENSORIA.ORG.CO		11458	65	5,672.9	<div></div>
14	<input checked="" type="checkbox"/>		G CO: UN.ORG		20225	64	3,164.4	<div></div>

Once the Virtual Corpus is created, you can limit future searches to that corpus (for example, to find matching phrases or collocations), and organize the corpora in several ways. And in one or two seconds, you can also find "keywords" from the Virtual Corpus, as shown for nouns in the [medio ambiente] and [nuclear] corpus that we've just created, and you can also see a translation of all of the keywords into whatever language you choose.

MEDIO_AMBIENTE [4,266,600 WORDS, 100 TEXTS] (0.4% OF TOTAL)		NOUN	VERB	ADJ	ADV	N+N	ADJ+N	[ALL CORPORA]	SAVE LIST
HELP	SAVE	WORD (CLICK FOR CONTEXT) TRANSLATE ALL ENTRIES	FREQ	# TEXTS	SPECIFIC FREQ 90 10 TEXTS		ENTIRE CORPUS	EXPECTED	
1	★	RESIDUO	2495	34	129.1		4,528	19.3	
2	★	ATRIBUCIÓN	575	12	21.2		6,366	27.2	
3	★	ASENTAMIENTO	508	14	15.0		7,953	33.9	
4	★	ECOSISTEMA	1010	37	12.8		18,463	78.8	
5	★	HABITANTE	556	31	12.5		10,423	44.5	
6	★	HECTÁREA	260	11	10.4		5,865	25.0	
7	★	DESECHOS	775	25	9.3		19,487	83.1	
8	★	TONELADA	251	16	9.3		6,347	27.1	
9	★	CONTAMINACIÓN	1998	51	8.3		56,338	240.4	
10	★	SUSTANCIA	1611	23	8.2		46,322	197.6	

NUCLEAR [274,708 WORDS, 167 TEXTS] (0.0% OF TOTAL)		NOUN	VERB	ADJ	ADV	N+N	ADJ+N	[ALL CORPORA]	SAVE LIST
HELP	SAVE	WORD (CLICK FOR CONTEXT)	FREQ	# TEXTS	SPECIFIC FREQ 20 7 TEXTS		ENTIRE CORPUS	EXPECTED	
1	★	NEUTRÓN	101	10	477.5		770	0.2	
2	★	FISIÓN	130	14	322.1		1,469	0.4	
3	★	REACTOR	404	34	305.5		4,814	1.3	
4	★	RESIDUO	215	18	172.8		4,528	1.2	
5	★	URANIO	297	16	160.0		6,757	1.9	
6	★	CENTRAL	937	51	105.8		32,228	8.9	
7	★	COSTE	298	19	28.0		38,803	10.7	
8	★	FUSIÓN	180	13	25.7		25,501	7.0	
9	★	COMBUSTIBLE	262	23	23.8		39,998	11.0	
10	★	ENERGIA	58	7	21.8		9,667	2.7	
11	★	ELECTRICIDAD	182	21	15.8		42,034	11.5	

7. Analyze texts (go to beginning)

You can also analyze entire texts, using the corpus data. For example, you can copy an article from an online newspaper and then paste it into the "Analyze Text" form (via the "Analyze" icon at the top of the corpus)

Incidente de aviones Volaris destapa riesgos en torres de control de aeropuertos de México

Señala Sinacta que el gremio atraviesa una crisis que se origina en dos situaciones: las malas condiciones laborales y el rediseño de la navegación



Corpus del Español: Web/Dialects

LIST OF WORDS DAY CONT

SAMPLES MY TEXTS

El sindicato de controladores aéreos reveló que el gremio enfrenta una crisis, tras el incidente del sábado pasado en el Aeropuerto Internacional de la Ciudad de México (AICM) entre dos aviones de Volaris, el cual derivó en la renuncia del director de Servicios a la Navegación en el Espacio Aéreo Mexicano (Seneam), Víctor Manuel Hernández. De acuerdo con el Sindicato Nacional de Controladores de Tránsito Aéreo (Sinacta), de diciembre de 2021 hasta el mes pasado se han reportado 30 incidentes aéreos graves en el país, de los cuales 10 ocurrieron en el AICM, a consecuencia del rediseño del espacio aéreo y de las condiciones laborales actuales.

From this page, you can find keyword word in the text synonyms, collocates that word. You can phrases in the Corp

There are several w

1. Just copy and paste (Word file) into the t

SEARCH CLEAR HELP

After clicking Submit, the following page shows you the [1] the number of words in high / medium / low frequency bands (to see the "difficulty" of the text) [2] the top words in each of these frequency bands, and [3] the text, with less frequent words (in the entire corpus) highlighted in green and especially yellow.

EDIT TEXT SAVE TEXT WORD PHRASE

1

FREQ RANGE 1-500 501-3000 > 3000

435 WORDS 54 % 21 % 25 %

3

El **sindicato** de **controladores aéreos** **reveló** que el **gremio** **enfrenta** una **crisis**, tras el **incidente** del **sábado** **pasado** en el Aeropuerto Internacional de la Ciudad de México (AICM) entre dos **aviones** de Volaris, el cual **derivó** en la **renuncia** del **director** de Servicios a la Navegación en el Espacio Aéreo Mexicano (Seneam), Víctor Manuel Hernández. De acuerdo con el Sindicato Nacional de Controladores de Tránsito Aéreo (Sinacta), de **diciembre** de 2021 hasta el mes **pasado** se han **reportado** 30 **incidentes aéreos** **graves** en el país, de los cuales 10 **ocurrieron** en el AICM, a **consecuencia** del **rediseño** del **espacio aéreo** y de las **condiciones laborales** **actuales**.

Los **incidentes** **graves** son aquellos donde los **pilotos** tienen que realizar una **maniobra**

(CLICK TO SEE WORD SKETCH)

LOW FREQ	MID FREQ	HIGH FREQ
5: aéreos, controladores 4: incidentes 2: aéreo, aterrizar, aviones, incidente, rediseño, sindicato 1: abortar, aerolínea, alertas, aterrizaje, auxiliar, bloques, colisión, congestionamientos, controlador, covarrubias, déficit, desconocían, despidos, exdirector, firmaran, fortuna, gremio, ida, injustificados, maniobra, ocasionado, pico, pilotos, pista, renuncia, reportado, restituir, supervisor	5: una 3: es 2: avión, condiciones, director, espacio, evitar, graves, laboral, laborales, pasado, posición, puesto, técnica 1: accidente, actuales, aire, atención, autoridades, cargo, cartas, colocar, consecuencia, crisis, cuanto, diciembre, enfrenta, entrara, entrevista, foto,	28: de 24: el 18: en 15: que 14: la 9: y 8: a, del, los 5: las, un 4: con 3: para, se 2: además, al, donde, dos, este, no, país, por, pues, quien, tienen 1: acuerdo, ahora, amigos, aquellos, bajo, cada,

You can then click on any of the words in [2] the frequency bands or [3] the text itself to see the "word sketch":

GREMIO *n* (RANK 3879, FREQ 30510)



DEFINITIONS (click for more)

guild, trade, profession, trade, union, union, labor, union, (body)

RAE Ling WR Rev Coll SpD Oxf



SYNONYMS (click to see word)

congregación *n* (meeting) cofradía, comunidad, compañía, regla, junta, conferencia, gremio, hermandad, **entidad** *n* company institución, corporación, asociación, mancomunidad, gremio, colectividad, organismo, **sindicato** *n* (union) colectividad, agrupación, asociación, cofradía, gremio,

SPOK	FIC	NEWS	ACAD

COLLOCATES

(ADJ) empresarial, docente, periodístico, magisterial, médico, sindical, seccional, paritario, lucrativo, cafetero
(NOUN) representante, trabajador, sindicato, dirigente, asociación, secretario, titular, empresario, camionero, paro

TOPICS (click to see)

paro *n* trabajador *n* sindicato *n* salarial *j* docente *n* gremial *j* sindical *j* huelga *n* aumento *n* docente *j* salario *n* sueldo *n* paritario *j* dirigente *n* provincial *j* movilización *n* periodista *n* acuerdo *n* camionero *n* secretario *n* negociación *n* protesta *n* profesión *n* reunión *n* obrero *j* reclamo *n* asociación *n* afiliado *n* productor *n* obrero

You can also find related phrases from the two billion word corpus. For example, you could highlight the phrase *incidentes graves* in the text [1], to find similar phrases in the corpus. It inserts the words into [2], and you can then click the buttons below that [3] to find all forms of a word, words with the same part of speech, synonyms of a word in the phrase, etc. Finally, it will then find matching strings from the corpus [4]:

EDIT TEXT

SAVE TEXT

WORD

PHRASE

FREQ RANGE

1-500

501-3000

> 3000

435 WORDS

54 %

21 %

25 %

FIND MATCHING PHRASES IN THE CORPUS (HELP)

NOUN

GRAVE

+ ADD

- DELETE

*

FORMS

SYNONYMS

POS:

NOUN

VERB

ADJ

ADV

PREP

Other:

SUBMIT

RESET

HELP

1

El **sindicato de controladores aéreos** **reveló** que el **gremio** **enfrenta una crisis**, tras el **incidente** del **sábado pasado** en el Aeropuerto Internacional de la Ciudad de México (AICM) entre dos **aviones** de Volaris, el cual **derivó** en la **renuncia** del **director** de Servicios a la Navegación en el Espacio Aéreo Mexicano (Seneam), Víctor Manuel Hernández . De acuerdo con el Sindicato Nacional de Controladores de Tránsito Aéreo (Sinacta), de **diciembre** de 2021 hasta el mes **pasado** se han **reportado 30 incidentes aéreos graves** en el país, de los cuales 10 **ocurrieron** en el AICM, a **consecuencia** del **rediseño** del espacio y de las **condiciones laborales actuales** .

Los **incidentes graves** son aquellos donde los **pilotos** tienen que realizar una **maniobra** para **evitar una colisión**, **explicó** en **entrevista** con EL UNIVERSAL el **secretario** general del

ON CLICK: [CONTEXT](#) [TRANSLATE \(EN\)](#) [ENTIRE PAGE](#) [GOOGLE](#) [IMAGE](#) [PRON/VIDEO](#) [BOOK](#) (HELP)

HELP	①	★	ALL FORMS (SAMPLE): 100 200 500 WORDS	FREQ	TOTAL 43,593 UNIQUE 2,755 +
1	①	★	PROBLEMA GRAVE	2041	
2	①	★	ENFERMEDAD GRAVE	1917	
3	①	★	PROBLEMAS GRAVES	1640	
4	①	★	LESIONES GRAVES	1505	
5	①	★	ENFERMEDADES GRAVES	1300	
6	①	★	CASOS GRAVES	1271	
7	①	★	FALTA GRAVE	1218	
8	①	★	CONSECUENCIAS GRAVES	970	
9	①	★	DELITOS GRAVES	955	
10	①	★	DELITO GRAVE	856	

Summary (go to beginning)

The Corpus del Español contains the largest, carefully annotated corpora of Spanish available from any source, including the Real Academia Española and Sketch Engine. The corpus allows you to search by word, phrase, lemma (all forms of a word), part of speech, synonym, and customized word list, and to see the matching words, strings, or collocates in just one or two seconds. You can use the corpora to see variation by genre, historical period, and dialect. You can also very quickly and easily create "Virtual Corpora" related to any topic (or based on metadata for the texts), and then limit searches to that Virtual Corpus, or see keywords from the Virtual Corpus.

The corpora allow you to perform "word-oriented" searches in ways that are not possible with any other corpus. You can browse and search the top 40,000 words (lemmas) in the corpus, by word form, part of speech, frequency range, and basic English gloss. For each of the top 40,000 words, you can see a "word sketch" that provides an incredible wealth of information, including frequency information, definition, synonyms, collocates, related topics, concordance lines, and links to external images and videos. You can also analyze entire texts, and then click on any word in the text to see the "word sketch", or find related phrases in the corpus. And for all of these searches, there are a wide range of links to very useful external data (translations, definitions, images, videos, etc).

In summary, the corpora from the Corpus del Español allow researchers, teachers, and students to gain insight into Spanish, which is not available from any other resource.