

Los corpus de nativos como recurso para aprender español dentro y fuera del aula

Mar Cruz Piñol



1. <https://dle.rae.es/?id=AwTBMcs> [¿qué es un corpus (1)?]
2. <http://users.ox.ac.uk/~martinw/dlc/chapter1.htm> [¿qué es un corpus (2)?]
3. https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/21/21_0277.pdf [Buyse, 2011]
4. <https://www.google.com/search?client=firefox-b-d&q=estudio> [Google como corpus]
5. https://cvc.cervantes.es/ensenanza/biblioteca_ele/carabela/pdf/56/56_121.pdf [Sala Caja, 2004]
6. <http://corpus.rae.es/creanet.html> [corpus CREA]
7. [http://corpus.leeds.ac.uk/cgi-bin/cqp.pl?q=estudio&c=INTERNET-ES&searchtype=conc&contextsize=60c&sort1=word&sort2=right&terminate=100&llstat=on&cleft=0&tright=1&cfilter= \(de http://corpus.leeds.ac.uk/internet.html\)](http://corpus.leeds.ac.uk/cgi-bin/cqp.pl?q=estudio&c=INTERNET-ES&searchtype=conc&contextsize=60c&sort1=word&sort2=right&terminate=100&llstat=on&cleft=0&tright=1&cfilter= (de http://corpus.leeds.ac.uk/internet.html)) [corpus Leeds]
8. <http://web.frl.es/CORPES/> (lema) [corpus CORPES]
9. <http://formespa.rediris.es/pdfs/asele37.pdf> [Alonso Pérez-Ávila, 2007]
10. https://books.google.com/ngrams/graph?content=ordenador%2Ccomputadora%2Ccomputador&year_start=1985&year_end=2019&corpus=21&smoothing=3&share=&direct_url=t1%3B%2Cordenador%3B%2Cc0%3B.t1%3B%2Ccomputadora%3B%2Cc0%3B.t1%3B%2Ccomputador%3B%2Cc0 [corpus GoogleBooks NGram]
11. <https://es.glosbe.com/es/ja/aunque> [Glosbe: memorias de traducción como corpus]
12. <https://context.reverso.net/traduccion/espanol-japones/darse+cuenta> [Reverso Context: memorias de traducción como corpus]
13. <http://valesco.es/?q=palabras> [corpus Valesco]
14. <http://opus.lingfil.uu.se/cwb/OpenSubtitles/frames-cqp.html> [corpus Opus]
15. <http://corpus.consumer.es/corpus/konsulta> [corpus Consumer]
16. http://corpora.informatik.uni-leipzig.de/en/res?word=paseo&corpusId=spa_newscrawl_2011 [corpus Wortschatz]
17. <https://www.corpusdelespanol.org/web-dial/> [Corpus del español, de M. Davis]
18. <http://multiwordnet.fbk.eu/online/multiwordnet.php> [MultiwordNet]
19. <https://eric.ed.gov/?id=EJ468954> [Honeyfield, 1989]
20. <https://dialnet.unirioja.es/servlet/articulo?codigo=2095396> [McCullough, 2001]
21. http://www.ict4lt.org/en/en_mod2-4.htm#sec4.2 [Lamy M-N. y Klarskov Mortensen H. J., 2009]
22. <http://tagpacker.com/user/mar.cruz.pinol?t=corpus> [más sobre corpus]

<http://sites.google.com/view/corpus-ele-univ-jap-2019/>

URL

¿Qué es un corpus?

corpus²

Del ingl. *corpus*, y este del lat. *corpus* 'cuerpo'.

1. *m.* Conjunto lo más extenso y ordenado posible de datos o textos científicos, literarios, etc., que pueden servir de base a una investigación.

Real Academia Española © Todos los derechos reservados

URL 1

“A corpus is a collection of pieces of **text** in **electronic form**, selected according to external criteria to **represent**, as far as possible, **a language** or language variety” ([Sinclair](#), 2004)

URL 2

Corpus
de nativos ~ **de aprendices**

Corpus y L2

- lengua en uso
 - **competencias lingüísticas**
 - carga cultural → **competencias interculturales**
- exposición a la lengua (**input**)
- aprendizaje autónomo (**dentro y fuera del aula**)

Muchos corpus textuales de nativos

¿Útiles? ¿Difíciles?

→ **Depende**

- Conceptos básicos sobre LC
- Tipos de corpus
- Tipos de interfaces
- Proyectos (beta)

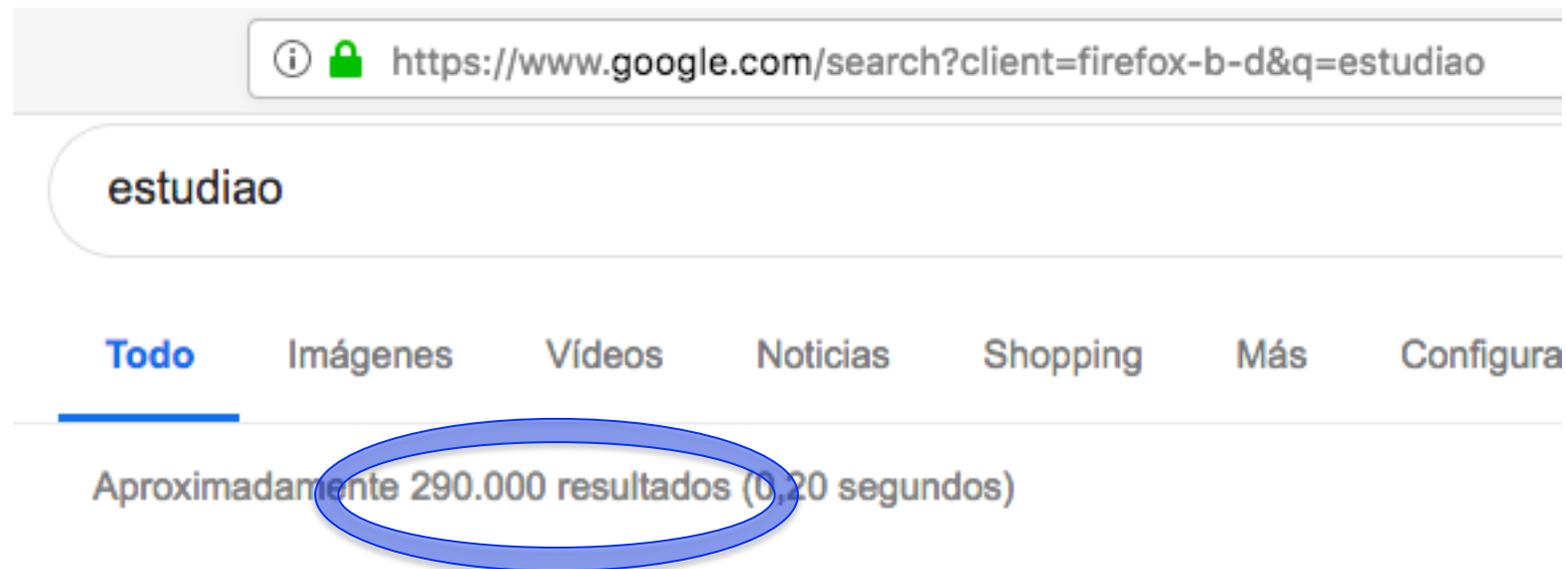
→ **Conocer para escoger**



La palabra X, ¿se dice?

“estudio”, ¿se dice?

→ Google



The screenshot shows a Google search interface. The address bar contains the URL <https://www.google.com/search?client=firefox-b-d&q=estudio>. The search bar contains the text "estudio". Below the search bar, there are navigation tabs: "Todo" (selected), "Imágenes", "Vídeos", "Noticias", "Shopping", "Más", and "Configura". Below the tabs, the text "Aproximadamente 290.000 resultados (0,20 segundos)" is displayed, with "290.000 resultados" circled in blue.

“(...) quien sabe extraer información lingüística de Google tendrá el camino más fácil cuando intente manejar un verdadero programa de concordancias como el CREA (Corpus de Referencia del Español Actual)”



[Sala Caja, 2004, p. 141]

Real Academia Española - Corpus de Referencia del Español Actual (CREA)

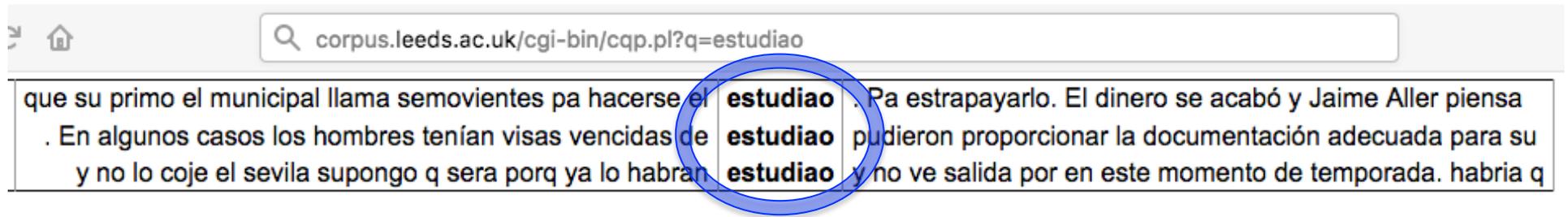
| | | | |
|--------------------------------|--|---|--|
| Consulta: | estudioa | | |
| Criterios de selección: | | | |
| Autor: | | Obra: | |
| Cronológico: | | Medio: | Geográfico: |
| | | (Todos) Libros Periódicos Revistas Miscelánea Oral | (Todos) Argentina Bolivia Chile Colombia Costa Rica |
| Tema: | (Todos) 1.- Ciencias y Tecnología. 101.- Biología. 102.- Veterinaria. 103.- Ecología. 104.- Tecnología. | | |

Buscar Limpiar

Nº CONCORDANCIA

1 libros!... Eso te sale o no te sale y, como no es estudioa, tú ni te enteras bien y muchas veces coge d **
2 bro y algunos hasta de poesías, yo creo que había estudioa y todo. Bueno, pues esa calcumanía de mujer, **

Corpus Leeds



A screenshot of a web browser search results page. The address bar shows the URL `corpus.leeds.ac.uk/cgi-bin/cqp.pl?q=estudio`. The search results are displayed in a table with three columns. The first column contains text snippets, the second column contains the word **estudio**, and the third column contains text snippets. The word **estudio** in the second column is circled in blue. The text snippets in the first and third columns are: "que su primo el municipal llama semovientes pa hacerse el", ". Pa estrapayarlo. El dinero se acabó y Jaime Aller piensa", ". En algunos casos los hombres tenían visas vencidas de", "y no lo coje el sevila supongo q sera porq ya lo habran", and "y no ve salida por en este momento de temporada. habria q".

| que su primo el municipal llama semovientes pa hacerse el | estudio | . Pa estrapayarlo. El dinero se acabó y Jaime Aller piensa |
|---|----------------|--|
| . En algunos casos los hombres tenían visas vencidas de | estudio | pudieron proporcionar la documentación adecuada para su |
| y no lo coje el sevila supongo q sera porq ya lo habran | estudio | y no ve salida por en este momento de temporada. habria q |

La palabra X, ¿dónde se dice?

La palabra “altoparlante”, ¿dónde se dice?

REAL ACADEMIA ESPAÑOLA

Corpus del Español del Siglo XXI
(CORPES)

Concordancias | Comparaciones | Configuración | Ayuda | Modo de cita | Sugerencias

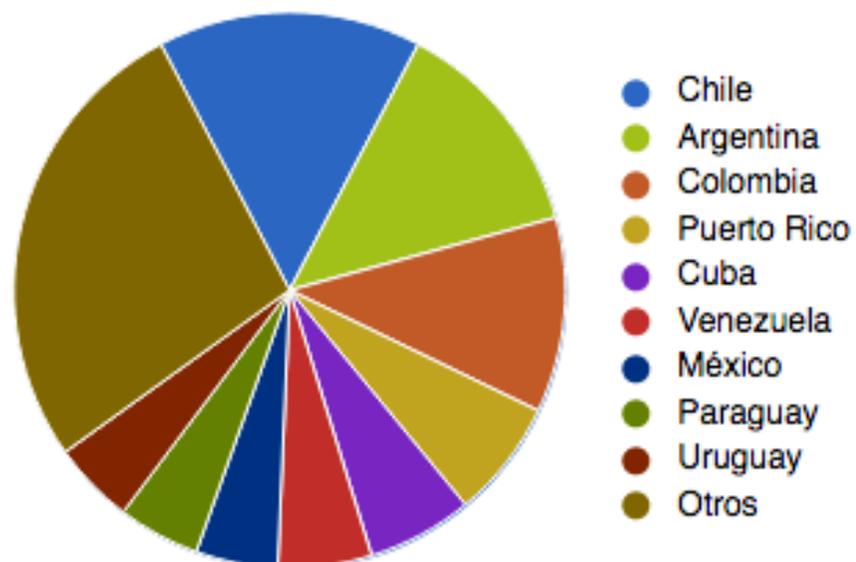
Lema: **altoparlante** Forma: Clase de palabra: (Todos) Grafía original + Subcorpus

Concordancia Estadística Nueva consulta

| REF. (Clasificación, país) | CONCORDANCIA | Ordenar |
|----------------------------|---|---------|
| 1 2001 Col. | sumido en desesperada soledad. Tal es su ensi | |
| 2 2001 Col. | Semejante regocijo, desbordado y volcánico, acalla las campanas, acalla los altoparlantes de las torres, aca | |
| 3 2001 Chile | una estación, el tren se detendrá más de lo necesario. El conductor hablará por los altoparlantes excusándose sin | |
| 4 2001 Arg. | empecé a chorrear transpiración. Tuvieron que interrumpir la escena por mí. Desde un altoparlante el flaquito de los | |
| 5 2001 Arg. | "Todos los extras de Ta condenada" al estudio 3", dijeron por el altoparlante en el momento er | |
| 6 2001 Arg. | El altoparlante anunció nuestra f | |
| 7 2001 Ur. | Los tres toman café en silencio. Por los altoparlantes anuncian el emba | |
| 8 2001 Guat. | a la gente, y se acordó del entrenamiento en sus años de soldado, cuando por el altoparlante ordenaban en tod | |
| 9 2001 Pan. | con el fin de evitar el Mitin, ya que el Alcalde mantuvo la prohibición. Desde nos altoparlantes se escucha La Ma | |
| 10 2001 Col. | Por los altoparlantes anunciaron el reir | |
| 11 2001 Col. | silencio en la alcoba contrastaba con el repicar de campanas y villancicos. El altoparlante de la iglesia; con l | |
| 12 2001 Col. | de la conveniencia de no acudir al llamado, que en ese momento le hacían por el altoparlante . Preferible era pe | |
| 13 2001 Chile | a hablar con el "futre" (dueño del basural), de quien sólo aparece la voz por un altoparlante . Nada consiguen, | |
| 14 2001 R.Dom. | , un puente aéreo, un flujo de gente que viene y va. Allí esperábamos la voz del altoparlante el aviso, la orden | |
| 15 2001 R.Dom. | El viaje termina, la voz de altoparlante dice que se abroc | |
| 16 2001 Perú | de la acera, en su lugar se han instalado ambulantes que vociferan con ayuda de altoparlantes a a voz en cuell | |
| 17 2001 EE.UU. | Carlos Kunst era casi como ir a otro planeta. La música rock, proveniente de los altoparlantes | |
| 18 2001 EE.UU. | primera vez desde que conocía a los fichas, la casa no retumbaba con la música de los altoparlantes | |
| 19 2001 EE.UU. | gruesas y arcos de magnesio cuyos vidrios ahumados vibran debido al retumbar de sus altoparlantes | |

URL 8

Distribución País



Distribución Zona



“(Las concordancias) no solo garantizan el **uso** y **frecuencia** de una unidad léxica determinada, sino que esta aparece **contextualizada y clasificada** en géneros textuales, períodos históricos, **zonas geográficas**, etc.”

[Alonso Pérez-Ávila, 2007]

Qué palabra se usa más, ¿X o Y?

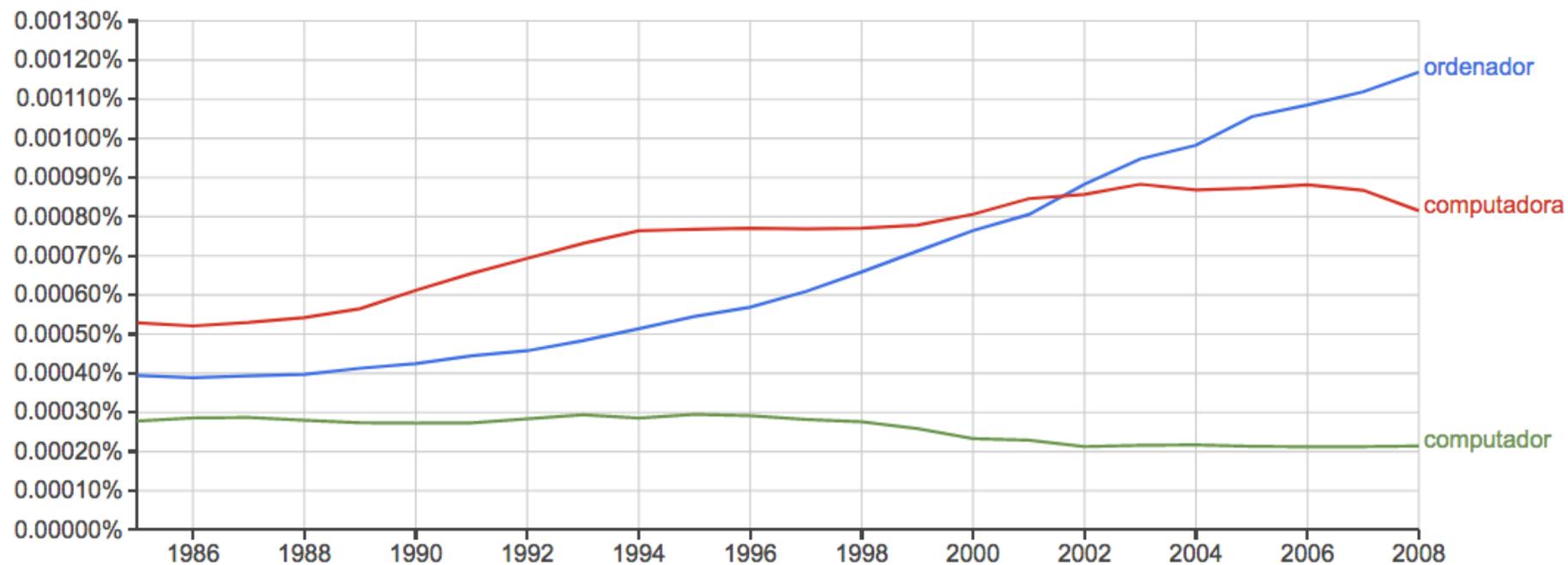
Qué palabra se usa más:

→ “ordenador”, “computadora” o “computador”

Google Books Ngram Viewer

Graph these comma-separated phrases: case-insensitive

between and from the corpus with smoothing of [Search lots of books](#)



La palabra X, ¿cómo se usa?

La palabra “aunque”, ¿cómo se usa?

Oraciones de ejemplo con "aunque", memoria de traducción

stemming

[add example](#)

Aunque en muchas viviendas no hay extintores, la ley suele exigir que los edificios públicos, fábricas y oficinas los tengan.

Aunque aplicar shuba es una tarea muy laboriosa, el mantenimiento que requiere es mínimo.

Tuve que luchar mucho para no amoldarme a lo que otros esperaban de mí, **aunque** me alegro de haberlo hecho.

Uno de los principios más eficaces del Evangelio, **aunque** a veces difícil de aplicar, es la humildad y la sumisión a la voluntad de Dios.

Aunque el 39 por ciento de rusos fuma [en], la cifra llega al 60 por ciento entre hombres rusos (en comparación, la cifra es de 21 por ciento para ambos sexos en el Reino Unido; [en]).

Aunque fracasas, no abandones.

La expresión X, ¿cómo se dice en otras lenguas?

La expresión “darse cuenta de...”,
¿cómo se dice en otras lenguas?

darse cuenta

Español

Japonés

Conjugar este verbo

Traducción de "darse cuenta" en japonés Ver también: darse cuenta de

Verbo

わかる 自覚 認識 気づく 気づか みれ 気付く

Es difícil no darse cuenta.

簡単にわかるね

Kip está creciendo, llegando a la edad en que empieza a darse cuenta... de que es parte de un mundo más grande.

キップは 成長して行きます 彼が自覚する 年齢に達するまで 大きな世界の 一員だと

Poco pudo darse cuenta que viajó al pasado de un universo paralelo.

並行宇宙に旅行した自覚は 彼にはなかった

Es muy joven para poder darse cuenta.

君はまだ若いからユダヤ人虐殺を 知らないのだ

Se dejan influenciar por cosas sin darse cuenta.

彼らは自らが気づかないことに 影響されるのです

La palabra X, ¿cómo se dice en otras lenguas
parecidas?

La palabra “hijo”,
¿cómo se dice en otras lenguas románicas?

"hij.*"

sort:

Display: context =

Alignments: bg cs da de el en et fr he hr hu it ja lt lv nl no pl pt ro ru sk sl sv tr zh

1 - 11 [11 matches] context =

| | fr | it | pt | ro |
|---|--|---|--|---|
| 1. heat, movie 190, sub 154315 context Tengo una hijastra bien jodida porque su papá es un desgraciado monumental . | J' ai une belle- fille totalement paumée parce que son vrai père est un gros enulé . | Ho una figliastra fuori di testa ... perche suo padre è un grande stronzo . | Tenho uma enteada completamente desorientada porque o pai dela é uma besta . | Am o fiica vitrega care are probleme emotionale deoarece tatal ei e o jigodie . |
| 2. heat, movie 190, sub 154315 context Mi hija me la dio . | - C' est un cadeau de ma fille . | Me l' ha regalata mia figlia . | Deu- ma a minha filha . | - E cadou de la fiica mea . |
| 3. heat, movie 190, sub 154315 context ¿ Quién te dijo , soplón hijo de perra ? | Qui t' a dit , salope ? | Chi te l' ha detto , spione di merda ? | Quem te disse , sacana ? | Cine ti- a zis , javra ? |
| 4. heat, movie 190, sub 154315 context Puedes hacerlo por tu hijo , para que lo | Vous pouvez le faire , pour pouvoir élever | Può farlo ... per suo figlio , per poterlo | Pode fazê- lo pelo seu filho , para que possa | Poti s- o faci , ca sa- ti poti creste coopilul . |

Corpus Lingüístico Consumer

La herramienta muestra cómo se utiliza una palabra o expresión en su contexto real, dentro de los contenidos publicados en la revista CONSUMER EROSKI.

Pregunta

Idioma: Castellano | Buscar por: Lema | Modo: Es | Palabra: helado | Categoría: [dropdown]

Idioma: Castellano | Buscar por: Lema | Modo: Es | Palabra: [dropdown] | Categoría: [dropdown]

Resultado

Castellano Euskera Gallego Catalán

Opciones del gráfico

[dropdown]

Buscar

Resultados: 83 artículo(s), 478 caso(s)

2018-02: TEMA DE PORTADA

Lácteos: se consumen sin conocer su valor

También tomaron 42 litros de yogures y leches fermentadas, 6 kilos de queso y requesón y 8 kilos de otros productos elaborados con leche, como nata, natillas, arroz con leche, batidos, helados, etc.

Por su parte, los derivados lácteos (yogur, batido, cuajada, helado, kefir o queso) son alimentos de fácil digestión y consumo.

2018-02: TEMA DE PORTADA

Lácteos: consómense sen coñecer o seu valor

Tamén tomaron 42 litros de iogures e leites fermentados, 16 quilos de queixo e queixo e 18 quilos doutros produtos elaborados con leite, como nata, cremas, arroz con leite, batidos, xeados etc.

Pola súa banda, os derivados lácteos (iogur, batido, callada, xeados, quefir ou queixo) son alimentos de fácil dixestión e consumo.

2018-02: TEMA DE PORTADA

Lactis: es consumeixen sense coñèixer el seu valor

També van prendre 42 litres de iogurts i llets fermentades, 16 quilos de formatge i mató i 18 quilos daltres productes elaborats amb llet, com ara nata, natilles, arròs amb llet, batuts, gelats, etc.

Daltra banda, els derivats lactis (iogurt, batut, quallada, gelat, quefir o formatge) són aliments fàcils de digerir i de consumir.

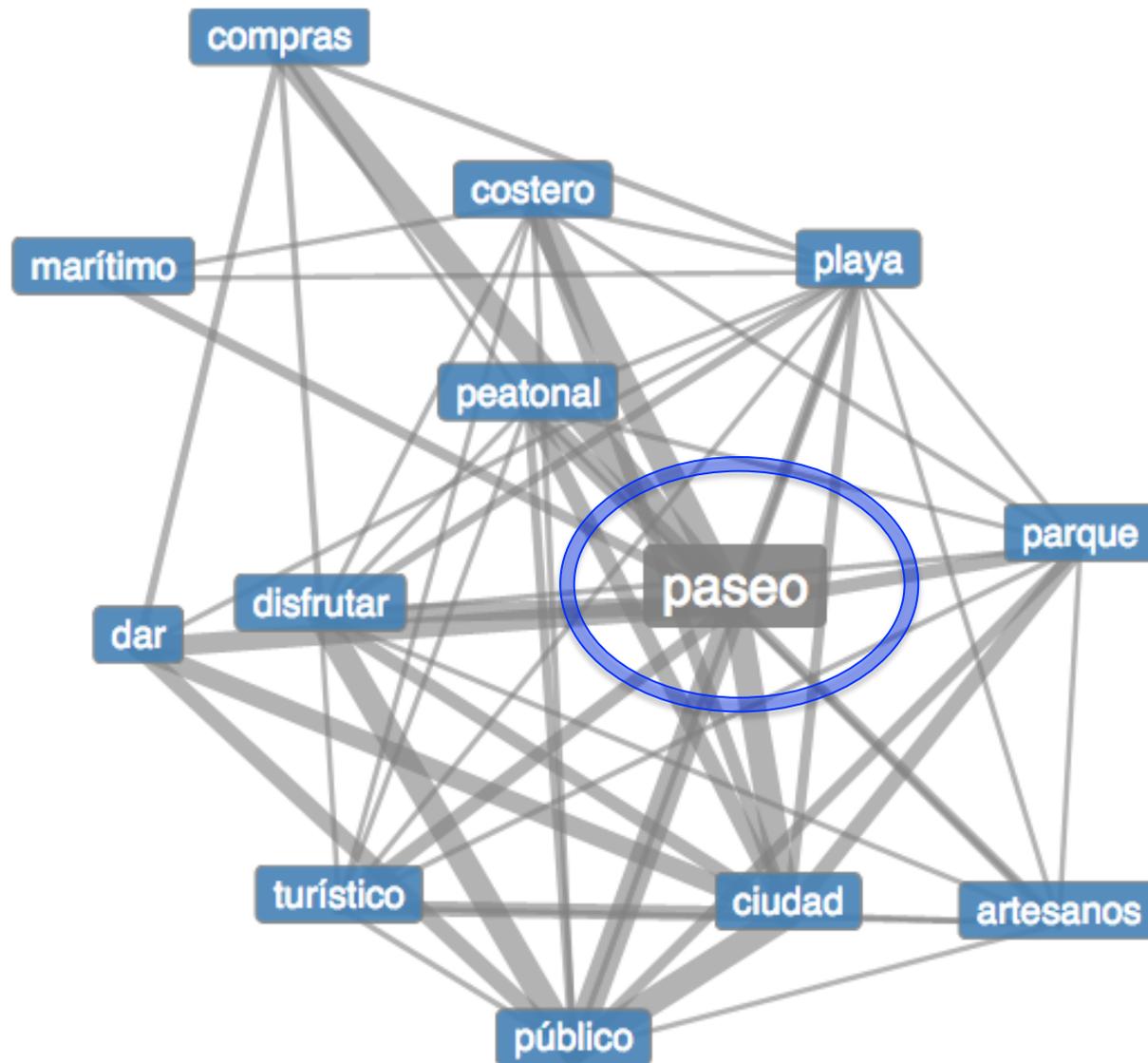
La palabra X, ¿cómo se usa en la lengua oral?

La palabra “mola”, ¿cómo se usa en la lengua oral?

| * Palabra (campo obligatorio) | Alargamiento vocálico | Orden palabra en intervención | Orden de la intervención | Texto en la intervención |
|----------------------------------|-----------------------|----------------------------------|-----------------------------|--|
| mola | no | 002 | 0009 | C: cómo mola nano (()) |
| mola | no | 005 | 0010 | B: a mí no me mola NAda |
| mola | no | 003 | 0011 | C: no te mola nada ¿el qué?§ |
| mola | no | 002 | 0015 | C: §no mola nada hoy sucesos† 'tío' mola más†= |
| mola | no | 007 | 0015 | C: §no mola nada hoy sucesos† 'tío' mola más†= |
| mola | no | 002 | 0017 | C: =Levante mola más§ |
| mola | no | 006 | 0019 | A: §((en el levante sí que mola))§ |
| mola | no | 007 | 0067 | B: [tomás ese/ tomás a] mí me mola porque es un tio [que/ aunque no tenga muy buen niv se lo curra nano] |
| | | | | C: § v. to(do) el r: |

¿Qué palabras se suelen combinar con la palabra X?

¿Qué palabras se suelen combinar con la
palabra “paseo”?



Buscar colocaciones

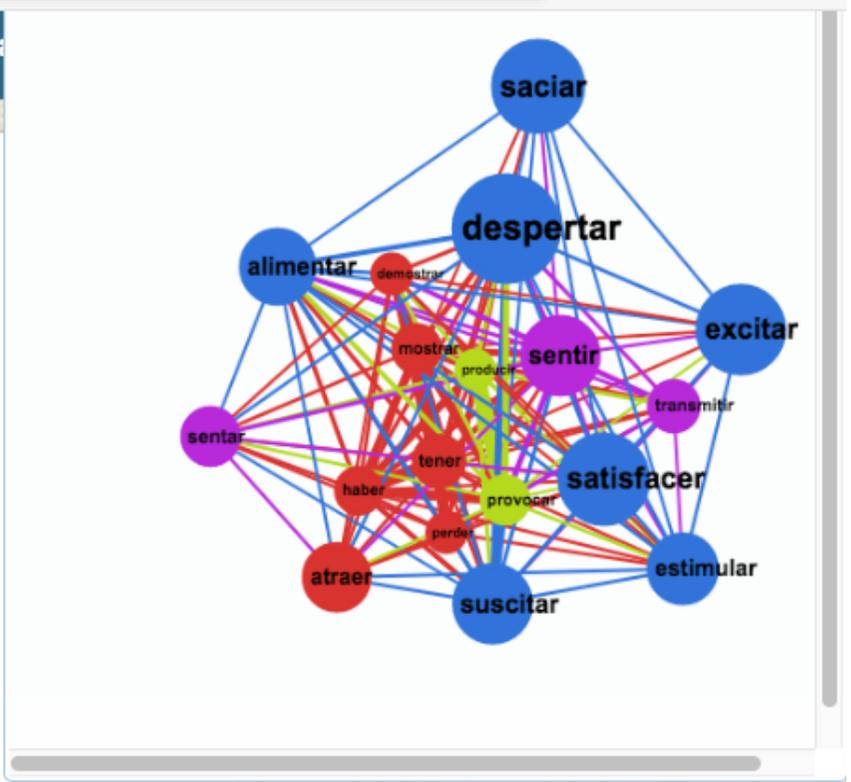
Lemma:
curiosidad

Categoría:
Nombre

Combinado con...
Adjetivo

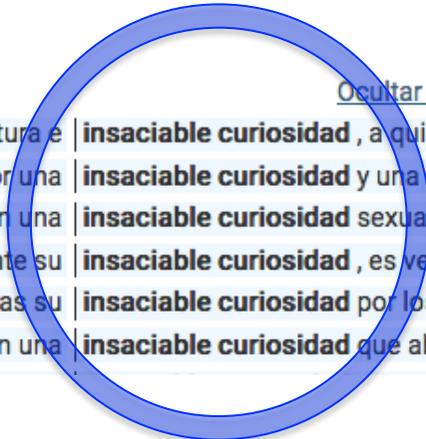
Mostrar grafo

Buscar



Results

- curiosidad insaciable (s.a) [Ocultar ejemplos](#) + ✓ ✗
- ... , el emprendedor de vastísima cultura e **insaciable curiosidad** , a quien Mitterra... +
 - ...ó ser periodista , quizás movido por una **insaciable curiosidad** y una temprana ... +
 - ...mo , se combinan últimamente con una **insaciable curiosidad** sexual . +
 - ... no lleven más allá de lo conveniente su **insaciable curiosidad** , es ver que al m... +
 - ...osidad por encima de todas las otras su **insaciable curiosidad** por los hábitos s... +
 - ...de los mayores y de esos niños con una **insaciable curiosidad** que abarca desd... +



CORPUS DEL ESPAÑOL

M CRUZ PIÑOL

INTRODUCCIÓN

Ayuda / información / contactar

historia | listas | salir

BUSCAR

PALABRA(S) un paseo

COLOCADOS [v*] 2 0

CAT GRAM

AL AZAR

BUSCAR

BORRAR

SECCIONES

MOSTRAR

1

- IGNORAR-
- s20
- s19
- s18
- s17
- s16
- s15

2

- IGNORAR-
- s20
- s19
- s18
- s17
- s16
- s15

ORDENAR Y LIMITAR

ORDENAR FRECUENCIA

MÍNIMO FRECUENCIA 4

ESCONDER OPCIONES

OCURRENCIAS FREC 100 PCEC 100

AGRUPAR POR LEMAS

| | <input type="checkbox"/> | CONTEXTO | FREC | |
|-------|--------------------------|-------------|------|--|
| 1 | <input type="checkbox"/> | [DAR] | 58 | |
| 2 | <input type="checkbox"/> | [HACER] | | |
| 3 | <input type="checkbox"/> | [SER] | 9 | |
| 4 | <input type="checkbox"/> | [REALIZAR] | 4 | |
| 5 | <input type="checkbox"/> | [TENER] | 2 | |
| 6 | <input type="checkbox"/> | [TOMAR] | 1 | |
| 7 | <input type="checkbox"/> | [PERDER] | 1 | |
| 8 | <input type="checkbox"/> | [ORGANIZAR] | 1 | |
| 9 | <input type="checkbox"/> | [OFRECER] | 1 | |
| 10 | <input type="checkbox"/> | [HABER] | 1 | |
| 11 | <input type="checkbox"/> | [APARECER] | 1 | |
| TOTAL | | | 88 | |

0.500 seconds

www.corpusdelespanol.org/x.asp?r1=&w=1093&h=614

Cerca

CORPUS DEL ESPAÑOL

M CRUZ PIÑOL

[RETURN TO SEARCH FORM](#) Ayuda / información / contactar

historia | listas | salir

IC PARA MÁS CONTEXTO [?] GUARDAR LISTA SELECCIONAR LISTA ----- CREAR NUEVA LISTA [?]

| | | | | |
|---|---|---|---|---|
| Habla Culta: Bogotá: M25 | A | B | C | - Sí, sí, sí realmente. Tú sabes que cuando uno hace un paseo , pues hay que aprovechar hasta el máximo, y más cuando hay esta p |
| Habla Culta: Bogotá: M27 | A | B | C | ..]. Inf.a. - De todas maneras es motivo para hacer un paseo e irlo a buscar. Inf.b. - Por una parte, y luego |
| Habla Culta: Bogotá: M27 | A | B | C | . - Una linda casa. Inf.a. - Pues hay que hacer un paseo a ver y llevar uno la lista de lo que se tiene... |
| Habla Culta: Bogotá: M38 | A | B | C | dijo: vamos a... a las Salinas de Zipaquirá, hacemos un paseo y allí hablaremos más ampliamente. Yo convidé a una amiga y estuvim |
| Habla Culta: Bogotá: M45 | A | B | C | bosques de Viena también los recorrimos, y por el Canal del Danubio hicimos un paseo , pero eso sí fue más bien monótono, porque |
| Habla Culta: Santiago: M21 | A | B | C | estaban sufriendo las odiosidades del régimen, no. Entonces, de allí hicimos un paseo muy lindo en barco por el río, por el Danubio; |
| Habla Culta: San José (CR): ... | A | B | C | - Sí, a mí me encantan. Inf. - Una vez hice un paseo con dos muchachos que iban a coger pájaros ahí, al Bajo de la Hondura |
| Función patronal : novela | A | B | C | agregó este párrafo: No sé dónde vine a parar. Anoche pretendí hacer un paseo de reconocimiento por el parque pero enseguida me |
| Cien años de soledad | A | B | C | y luego para la cena, y en toda la semana sólo pudieron hacer un paseo a las plantaciones. Al anochecer, las monjas estaban agotac |

La palabra X, ¿qué relaciones semánticas mantiene (hiperónimos/hipónimos)?

La palabra “interrogar”,
¿qué relaciones semánticas mantiene?

MultiWordNet On-line

multiwordnet.fbk.eu/online/multiwordnet.php

Search Special Options Setting Login

Spanish Word interrogar Search

Word statistics | Database report | Bug report | Credits

Verb

Hypernyms

Hypernyms of "interrogar":

Verb

- 1. **interrogar**, preguntar -- (Factotum) [pose a question]
 - > preguntar -- (Factotum) [address a question to and expect an answer from; "Ask your teacher about trigonometry"]
 - > dirigirse -- (Factotum) [speak to; "He addressed the crowd outside the window"]
 - > comunicar -- (Factotum) [transmit thoughts or feelings; "He communicated his anxieties to the psychiatrist"]
 - > interactuar -- (Factotum) [act together or towards others or with others; "He should interact more with his colleagues"]
 - > actuar, hacer, llevar a cabo -- (Factotum) [perform an action; "think before you act"; "We must move quickly"]

Sense: **interrogar/1**

Pos: Verb

Domain: Factotum

Affective:

UK

Synset: **question, query**

Phrasal:

Gloss: pose a question

IT

Synset: **domandare, interrogare**

Phrasal:

Gloss:

ES

Synset: **interrogar, preguntar**

Phrasal:

Gloss:

Materiales y actividades basados en corpus de nativos



A Typology of Exercises Based on Computer-Generated Concordance Material.

Honeyfield, John

Guidelines: A Periodical for Classroom Language Teachers, v11 n1 p42-50 Jun 1989

URL 20

Los usos de los corpóra de textos en la enseñanza de lenguas

Autores: James L. McCullough

Localización: Nuevas tecnologías para el autoaprendizaje y la didáctica de lenguas / coord. por Mireia Trenchs Parera, 2001, ISBN 84-89790-60-4, págs. 125-140

URL 21

v.ict4lt.org/en/en_mod2-4.htm#sec4.2

ICT4LT Module 2.4

Using concordance programs in the Modern Foreign Languages classroom

[Lamy M-N. y Klarskov Mortensen H. J. , 2009]

URL 22

Más ejemplos de corpus

Muchas gracias.

ありがとうございます。

Mar Cruz Piñol

