

Introducción al corpus DeReKo*

1. Introducción

Un corpus es una colección de expresiones escritas o habladas. Los datos del corpus suelen estar digitalizados, es decir, almacenados en ordenadores y son legibles por ellos. Los textos son los componentes fundamentales del corpus y estos son los propios datos, adicionalmente podemos encontrar metadatos que describen estos datos y anotaciones lingüísticas asociadas a estos datos. (Lemnitzer y Zinsmeister, 2006: 7).

Cuando hablamos de corpus lingüísticos, nos referimos a colecciones de textos completos o, al menos, con fragmentos de texto muy extensos. Además, los corpus lingüísticos suelen ser:

- representativos,
- documentados mediante metadatos y
- anotados lingüísticamente. (Lemnitzer y Zinsmeister, 2006: 40)

1.1. Relevancia del corpus DeReKo

Partimos del hecho de que cuanto más grande es un corpus, más confiables pueden ser las conclusiones sobre eventos raros y diversificados que podemos analizar o estudiar (Lüngen y Kupietz, 2014: 25).

De igual modo, cuanto mayor es la cantidad de datos y cuanto mejor es la forma en que se recopila la muestra, más probable es que las afirmaciones se apliquen al comportamiento transversal de la comunidad lingüística (Perkuhn et al. 2015, 182).

De este modo, el Deutsches Referenzkorpus (DeReKo) es una colección de múltiples corpora de textos escritos desde 1964. Mantenido por el Institut für Deutsche Sprache (IDS) de Mannheim y que continúa incorporando fondos a sus datos.

El tamaño del corpus se ha incrementado desde los 28 millones de palabras en 1992 a 55 mil millones en 2023. Asumiendo que un libro podría tener unas 100 mil palabras, estaríamos hablando de una biblioteca de 550000 volúmenes.

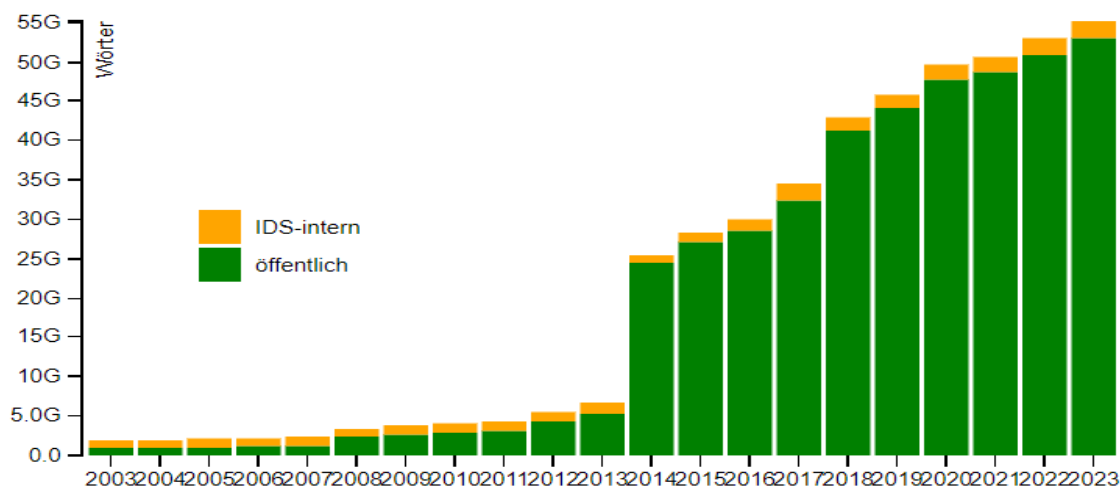


Ilustración 1: Evolución del tamaño del corpus entre 2003 y 2023. Fuente:

* El presente trabajo se realiza en el marco del PID 23/24 Historia de la Lengua Alemana 2.0.

1.2. Qué podemos hacer con un corpus.

Uno de los objetivos primordiales de un corpus es satisfacer las necesidades de cualquier estudio empírico en lingüística y campos relacionados. Algunos de los escenarios de aplicación en los que muestra algunas de sus ventajas podrían ser los siguientes (Kupietz, et al, 2010: 1851-1852):

- Investigación en torno al cambio lingüístico

Podemos investigar, por ejemplo, la evolución del significado de una palabra a lo largo del tiempo. Para ello, el mejor modo de abordar esta tarea es usar corpus de tamaño similar que cubran diversos periodos de tiempo (como por ejemplo hace Gondring, 2010). Para garantizar la objetividad de los datos es necesario que los corpus tengan tamaños similares y pertenezcan a los periodos delimitados en nuestra investigación para no provocar falseos estadísticos. Un aspecto que también puede alterar la investigación es que podemos conocer la fecha de publicación de un texto, pero esta no tiene por qué coincidir con la de la creación, por ello usar textos periodísticos puede ser una buena idea. También podemos restringir aún más el estudio limitándonos a una región determinada (por ejemplo, la RDA, el norte de Alemania) y usar textos de esa región. También podemos limitarnos a textos de un tipo determinado o de secciones concretas de un periódico (Kupietz, et al., 2010: 1851ss).

- Estudios lingüísticos sincrónicos

Los estudios sincrónicos investigan en torno a algún aspecto del lenguaje contemporáneo. Belica et al. (2010) proponen una estrategia general de muestreo (a la que denominan FReD4) para construir un corpus sincrónico. La clave es otorgar mayor peso a los textos más recientes. Esto es algo que el sistema de consulta de DeReKo puede reflejar sin problemas (Kupietz, et al., 2010: 1852)

- Otros escenarios

Hay gran variedad de escenarios de investigación que se benefician del diseño de DEREKO. Por ejemplo, un estudio puede estar interesado en contrastar el uso del lenguaje entre las variedades regionales del alemán escrito (por ejemplo, Austria vs. Alemania vs. Suiza). Los corpus virtuales adecuados podrían definirse exigiendo que contengan solo datos de la respectiva región y que sean comparables en cuanto a tamaño, época, etc. (Kupietz, et al., 2010: 1852).

2. Corpus presentes en DeReKo

En la actualidad DeReKo está compuesto por 336 corpora de entre los cuales podemos destacar, por el número de tokens que lo integran, los siguientes:

Sigla	Nombre	desde	hasta	Tokens
wpd	Wikipedia	2002	2019	4101253320
wdd	Wikipedia Diskussionen	2002	2019	1621975730



rhz	Rhein-Zeitung	1996	2022	1456674919
rpo	Rheinische Post	2001	2022	1233199146
mze	Mitteldeutsche Zeitung	2000	2022	999243211
sze	Sächsische Zeitung	2000	2022	931297509
sku	Südkurier	2000	2022	926908005
pnp	Passauer Neue Presse	2001	2022	864927318
lvz	Leipziger-Volkszeitung	2000	2022	850409094
sbz	Saarbrücker Zeitung	2000	2022	844762399
u	Süddeutsche Zeitung	1992	2022	804381714
nwe	Neue Westfälische	2003	2022	795896420
r	Frankfurter Rundschau	1997	2022	795004150
dpa	Deutsche Presse-Agentur	2006	2022	769460987
fnp	Frankfurter Neue Presse	2000	2022	747984081
ksa	Kölner Stadt-Anzeiger	2000	2022	673134720
mpo	Main-Post	1997	2022	667120170
rhp	Die Rheinpfalz	2007	2022	663115918
o	Neue Kronen-Zeitung	1994	2022	622989991
tha	Thüringer Allgemeine	2000	2022	612902815
kru	Kölnische Rundschau	2002	2022	607628831
scz	Schwäbische Zeitung	2011	2022	606327818
mag	Märkische Allgemeine	2007	2022	600822841
lru	Lausitzer Rundschau	2000	2022	597541575
wud	Wikipedia Nutzer-Diskussionen	2002	2017	592386408
bdz	Badische Zeitung	2003	2022	588280965
t	die tageszeitung	1986	2022	554108392
m	Mannheimer Morgen	1989	2020	548991133
k	Kleine Zeitung	1996	2022	525062638

El corpus es de libre uso pero no puede ser ni copiado ni descargado y ha de ser usado a través del sistema COSMAS II (*Corpus Search, Management and Analysis System*) en el que es requerido el registro por parte de los usuarios finales y la aceptación del uso no comercial y académico del mismo.

El acceso se realiza a través del siguiente enlace: <https://cosmas2.ids-mannheim.de/cosmas2-web/faces/noAccess.xhtml>

Ilustración 2: Acceso al sistema de búsqueda COSMAS II



En la actualidad se está implementando el sistema KorAp que sustituirá al sistema de búsqueda de Cosmas II. Se trata de una nueva plataforma de análisis optimizada para búsquedas de corpus anotados complejos (<https://korap.ids-mannheim.de/>). En comparación con COSMAS II, KorAP permite búsquedas en corpus completo, sin embargo, todavía están en desarrollo algunas opciones relacionadas con la búsqueda en torno a colocaciones (Kupietz et al, 2018)

3. Datos presentes en DeReKo

Una vez nos hemos identificado en Cosmas, podemos realizar la búsqueda en cualquiera de los 18 archivos predefinidos que nos ofrece (ver Ilustración 3):

vordefinierte Archive	
W - Archiv der geschriebenen Sprache	DeReKo (Teil 1)
W2 - Archiv der geschriebenen Sprache	DeReKo (Teil 2)
W3 - Archiv der geschriebenen Sprache	DeReKo (Teil 3)
W4 - Archiv der geschriebenen Sprache	DeReKo (Teil 4)
W-ÜBRIG - Archiv der aussortierten geschriebenen Korpora	DeReKo (partiell)
PARONYM - Archiv des Paronym-Projekts des IDS	DeReKo (partiell)
WP - Wikipedia Artikel u. Artikel- & Benutzerdiskussionen 2013/15/17/19	DeReKo (Teil 5)
WPE - englischsprachige Wikipedia Artikel u. Artikeldiskussion & Benutzerdiskussion 2015	
WP_FS - fremdsprachige Wikipedia Artikel u. Artikeldiskussion & Benutzerdiskussion 2015	
TAGGED-C - Archiv morphosyntakt. annotierter Korpora (CONNEXOR)	DeReKo (partiell)
TAGGED-C2 - Archiv morphosyntakt. annotierter Korpora (CONNEXOR)	DeReKo (partiell)
TAGGED-T - Archiv morphosyntakt. annotierter Korpora (TreeTagger)	DeReKo (partiell)
TAGGED-T2 - Archiv morphosyntakt. annotierter Korpora (TreeTagger)	DeReKo (partiell)
TAGGED-M - Archiv morphosyntakt. annotierter Korpora (MECOLB; ehem. TAGGED)	DeReKo (partiell)
HIST - Archiv der historischen Korpora	Historisches Textkorpus
UMB - Archiv des Projekts "Sprachliche Umbrüche des 20. Jahrhunderts"	
GFDS - Kartei der Gesellschaft für deutsche Sprache	
WK-PH - Archiv der phasengegliederten Wendekorpora	DeReKo (partiell)

Ilustración 3: Archivos presentes en DeReKo

Para esta introducción nos centraremos en el corpus: W – Archiv der geschriebene Sprache. Con ello estaremos utilizando la mayor cantidad de datos disponible públicamente.

The screenshot shows the COSMAS II search interface. At the top, it displays 'IDS UNIVERSITÄT DER DEUTSCHEN SPRACHE' and 'COSMAS II' with navigation options like 'Anmeldung', 'Suche', and 'Optionen'. The main search area shows 'Aktuelles Archiv: W - Archiv der geschriebenen Sprache' and 'Aktuelle Suchanfrage: W - Archiv der geschriebenen Sprache'. Below this, there are tabs for 'Archiv', 'Korpus', 'Such', 'Workform', 'Ergebnisse', 'Kook', 'KWIC', 'Volltext', and 'Export'. The 'Archiv' tab is active, showing a list of predefined archives with their corresponding DeReKo links. The 'W - Archiv der geschriebenen Sprache' is highlighted.

Ilustración 4: Selección del corpus W

4. Consulta en Cosmas II: conceptos básicos

A continuación, describiremos brevemente el lenguaje de consulta utilizado en la interfaz web de COSMAS II y los conceptos fundamentales.

4.1. Búsquedas

Podemos realizar búsquedas de diversos tipos:

- Palabras
- Signos de puntuación u otros caracteres especiales
- Números
- Patrones de búsqueda

Las palabras y los números pueden ingresarse directamente sin caracteres adicionales. Por ejemplo, la palabra:

> **Universität**

Los signos de puntuación y otros caracteres especiales deben ingresarse entre comillas (excepción: para buscar signos de interrogación, se coloca una barra invertida antes: \?)

> "!"

Existen también operadores comodines que permiten patrones de búsqueda:

4.2. Operadores comodín:

* = múltiples caracteres arbitrarios

? = 1 carácter arbitrario

> ***tät** ofrece palabras como Universität, Realität, Aktivität, etc.

> **?uch** ofrece palabras como Buch, Tuch, etc.

4.3. Operador de formas básicas &:

Al colocar & antes de una palabra en su forma básica, se encuentran todas las formas flexivas, compuestas u otras formas de formación de palabras, según la configuración.

> **&denken** encuentra todas las formas flexivas del verbo denken: denke, denkst, dachte, gedacht, etc.

4.4. Operadores de distancia

Uno de los operadores más útiles es el de distancia. Si se desean buscar varias palabras, es necesario definir cómo deben estar relacionadas. Para esto, se utilizan operadores de distancia u operadores lógicos como AND, OR y NOT.

Si no se utilizan tales operadores, se utilizará la configuración que se puede hacer directamente debajo de la ventana de búsqueda. Allí se puede especificar el "Operador de Conexión Estándar". Las palabras pueden ser conectadas con el operador de distancia de palabras "/w1" o con el "ODER" lógico.

Con el operador de distancia "/w", se especifica lo lejos que pueden o deben estar las dos palabras entre sí. Después de "/w", se especifica en un número la distancia máxima, por ejemplo:

> **Universität /w2 Valladolid**



Encuentra *Universität Valladolid*, *Universität von Valladolid*, *Universität in Valladolid*
El operador `/w2` especifica la "ventana de palabras" permitida entre las dos palabras clave. En este caso, el 2 sugiere que no debe haber más de dos palabras entre "Universität" y "Valladolid". Esto significa que podrían estar adyacentes (sin palabras intermedias) o tener hasta dos palabras entre ellas.

The screenshot shows the DeReko search interface. At the top, it displays the current archive ('W - Archiv der geschriebenen Sprache') and corpus ('W-öffentlich - alle öffentlichen Korpora des Archivs W (mit Neuaquisitionen) [1]'). The search query is 'Universität /w2 Valladolid' and the reference is 'Deutsches Referenzkorpus DeReKo-2023-1'. The number of hits is 40. Below the search bar, there are tabs for 'Archive', 'Korpus', 'Suchanfragen', 'Wortform.', 'Ergebnisse', 'Kook.', 'KWIC', 'Volltext', and 'Export'. The 'Suchanfragen' tab is active, showing the input field with 'Universität /w2 Valladolid' and a search button. Below the input field, there are options for the search operator: 'Wortabstand /+w1' (selected) and 'logisches 'ODER''. There is also a 'Suchen' button and a 'Wortformlisten:' link.

O la búsqueda:

> **gedacht /w1 haben**

Encontraría: "gedacht haben" y también lógicamente "haben gedacht", pero no "haben sie gedacht" dado que existe una distancia mayor entre los dos conceptos solicitados. Con los símbolos + y – se muestra, en qué dirección es posible la distancia: + hacia la derecha y el operador – en el sentido contrario.

> **gehabt /+w1 haben**

También podemos definir la distancia mínima. La consulta nos muestra la combinación de las dos palabras con una distancia de 3 tokens:

> **Vater /+w3 Sohn**

Encuentra "Vater und Sohn", "Vater mit seinem Sohn", "Vater zu seinem Sohn" etc. Sin embargo:

> **Vater /+w2:2 Sohn**

Encuentra solo la combinación de Tokens con una distancia mínima y máxima. La primera cifra indica la distancia mínima y la segunda cifra la máxima.

También es posible especificar una distancia de 0, por ejemplo, `/w0`". Por ejemplo. Para buscar signos de puntuación inmediatamente después de una palabra, ya que estos tienen una distancia 0 con la palabra anterior.

Este tipo de operador de búsqueda es útil para investigaciones lingüísticas que requieren un análisis de la colocación de palabras en un orden específico, especialmente útil para idiomas donde el orden de las palabras puede ser más flexible y tener implicaciones semánticas o sintácticas significativas.

4.5. Operadores Lógicos "AND", "OR" y "NOT"

Estos operadores pueden usarse para conectar palabras de búsqueda entre sí. Sin embargo, es importante tener en cuenta que se aplican al texto completo:

> **Vater AND Sohn**

muestra en los resultados solo una palabra, ya que la otra puede estar bastante lejos. En este caso, suele ser más útil trabajar con el operador de distancia de palabras.

4.6. Paréntesis

Si se utiliza más de un operador (operador de distancia de palabras u operador lógico), se deben colocar entre paréntesis. Esto agrupa siempre dos palabras de búsqueda con su operador de conexión en una expresión. Por ejemplo:

Supongamos que quieres investigar cómo se utilizan los verbos "gehen" y "laufen" en relación con adverbios de dirección como "nach Hause" o "zur Schule". Puedes estructurar la búsqueda de la siguiente manera:

> (gehen OR laufen) /+w1:3 (nach Hause OR zur Schule)

Esta búsqueda busca las combinaciones:

gehen nach Hause

gehen zur Schule

laufen nach Hause

laufen zur Schule

El operador /+w1:3 significa que el verbo (gehen/laufen) puede estar hasta tres palabras de distancia del adverbio de dirección (nach Hause/zur Schule), lo que permite cierta flexibilidad para palabras intermedias que podrían aparecer en un uso natural del lenguaje.

Este tipo de búsqueda es útil para investigar patrones de colocación y preferencias de uso en diferentes contextos, ayudando a comprender mejor cómo se combinan ciertos verbos y expresiones.

BÚSQUEDA	RESULTADO
Universität	Muestra solo los resultados hallados con la palabra introducida
&denken	Muestra todas las formas verbales, no solamente el infinitivo
*tät	Búsqueda de múltiples caracteres
?uch	Búsqueda de un único carácter
Universität /+w2 Valladolid	Búsqueda de palabras con una distancia máxima de dos entre ellas
Universität /+w2:2 Valladolid	Búsqueda de las dos palabras introducidas con una distancia exacta de 2 entre ambas
Universität /-w2:7 Valladolid	Búsqueda de Valladolid antepuesto a Universidad con una distancia de entre 2 y 7 palabras
&d- /+w1 ?oghurt	Búsqueda de las distintas formas flexivas del artículo determinado (masculino, femenino y neutro) a una distancia máxima de 1 palabra del sustantivo Yoghurt (escrito con J o Y

5. Resultados

DeReKo nos permite una clara visualización de los resultados obtenidos.



The screenshot shows the DeReKo search interface. At the top, there are navigation tabs: Archive, Korpus, Such., Wortform., Ergebnisse (selected), Kook., KWIC, Volltext, and Export. Below the tabs, there is a search bar with 'Aktuelle Suchanfrage: Valladolid' and 'Aktuelles Korpus: W-öffentlich - alle öffentlichen Korpora des Archivs W (mit Neuausgaben) [1]'. The search results are displayed in a table with columns: Treffer, Texte, von, bis, Korpus. The table lists various newspaper and magazine corpora, such as 'A07 St. Galler Tagblatt 2007', 'A14 St. Galler Tagblatt 2014', 'A15 St. Galler Tagblatt 2015', etc. The total number of results is 6.665.

Ilustración 5: Pestaña de resultados

Estos podemos ordenarlos de múltiples formas (ver Ilustración 6): en orden alfabético, cronológico, etc.

The screenshot shows the DeReKo search interface with the 'KWIC' tab selected. Below the tabs, there is a 'Sortierung:' dropdown menu with the following options:

- KWIC (unsortiert)
- KWIC (alphabet. sortiert)
- KWIC (chron. sortiert)
- KWIC (zufällig sortiert)
- KWIC (nach Kookk.stärke sortiert)

The 'KWIC (chron. sortiert)' option is highlighted. To the right of the dropdown, there is a 'zu Treffer:' field with the value '1' and a 'springen' button. Below the dropdown, there is a table with columns: Treffer, Texte, von, bis, Korpus. The table lists search results, such as 'WPD11/A43.87806:', 'WPD11/B23.59552:', 'WPD11/D42.66175:', 'WPD11/F05.72056:', 'WPD11/G45.29835:', and 'WPD11/H47.47481:'. The text in the table is partially obscured by a watermark.

Ilustración 6: clasificación de los resultados

La pestaña KWIC (Key Word in Context) permite consultar los resultados en contexto. En ella, los términos de búsqueda aparecen centrados y resaltados en color (Ilustración 7)

Archiv	Korpus	Such.	Wortform.	Ergebnisse	Kook.	KWIC	Volltext	Export
Sortierung: KWIC (chron. sortiert) zu Treffer: 1 springen								
Seite 1 von 1								
1	Z96/FEB.00137					Verfasser ist der spanische Kunsthistoriker Pedro de Paloi von der Universität Valladolid.		
2	S75/MAR.00149					Wiedereröffnung der vom Erziehungsminister für das ganze Semester geschlossenen Universität Valladolid zu erzwingen, zu einem Sit-in in die Kathedrale von San Isidro in Madrid		
3	WPD11/U45.41051					WEITERLEITUNG Universität Valladolid		
4	WPD11/M53.37283					studierte er zunächst Philosophie und Literaturwissenschaften an der Universität Valladolid und schloss dieses Studium mit einem Lizientat 1942 ab.		
5	WPD11/A43.87806					Sie schrieb und gab Kurse an der Universität Valladolid.		
6	WPD11/F05.72056					1600 alle Sprachen in Alcalá de Henares und von 1601 bis 1604 Theologie an der Universität Valladolid.		
7	WPD11/U29.06093					Von 1973 bis 1977 lehrte sie Deutsch an der Universität Valladolid.		
8	WPD11/T31.72836					am 3. Januar 2007.) Sie belegte außerdem Kurse für Kino-Ästhetik an der Universität Valladolid (1990) und kreatives Schreiben an der Escuela de Letras de Madrid (1998).		
9	WPD11/U28.19409					Hochschule, die 1994 aus der Abspaltung des Campustells in Burgos der Universität Valladolid hervorging.		
10	T11/MAI.04271					Arranz, der für das agrarwissenschaftliche Institut der zentralspanischen Universität Valladolid (inea) einen Biohof leitet, sagte: „In ganz Spanien ist es verboten, mi		
11	WPD11/L32.43696					er an den Lehrstuhl für Rechtswissenschaften, Philosophie und Rhetorik der Universität Valladolid berufen, deren Rektor er später auch wurde.		
12	WPD11/B23.59552					Er wird heute in der Bibliothek der Universität Valladolid aufbewahrt (MS 433) und gehört der Familie Ila an.		
13	WPD11/L05.57636					Universität Valladolid in Valladolid		
14	WPD11/L05.57636					Europäische Universität Miguel de Cervantes in Valladolid Universität Vc		
15	WPD11/H47.47441					an der Universität Madrid, 1995 an der Universität Barcelona und 1996 an der Universität Valladolid, 1997 am King Mongkut's Institute of Technology in Bangkok, 1999 an de		
16	WPD11/M13.04723					1997 erhielt er den Ehrendokortitel der Universität Alicante, der Universität Valladolid und der Universität Havanna.		
17	WPD11/U51.57667					Er studierte Zivilrecht an der Universität Valladolid und lehrte das Fachgebiet an der Universität Deusto.		
18	WPD11/U63.76581					WEITERLEITUNG Universität Valladolid		
19	WPD11/M32.46641					Nach dem Schulbesuch absolvierte er zunächst ein Studium der Philosophie an der Universität Valladolid.		
20	WPD11/M05.75282					Anschließend gehörte er bis 1960 zum Lehrkörper der Universität Valladolid und arbeitete in verschiedenen Diözesen Spaniens als Seelsorger.		
21	WPD11/P06.39379					Von 1912 bis 1916 gehörte er der katholisch-theologischen Fakultät der Universität Valladolid an und arbeitete in der Administration des Erzbistums Valladolid.		
22	WPD11/D42.66175					Diego Valverde Villena studierte Hispanistik, Anglistik und Germanistik an der Universität Valladolid.		
23	WPD11/G45.29835					Universität Valladolid Universität Vigo		
24	WPD11/M29.13739					1952 verbrachte er zwei Semester an der Universität Valladolid.		
25	SOL17/MAI.02626					es so zu einer Verjüngung der Bevölkerung“, sagt die Migrationsoziologin der Universität Valladolid. Foto: manso/afp		
26	U20/NOV.01193					Anderorts wird weiter geforscht, wie hier an der spanischen Universität Valladolid. Foto: manso/afp		
27	NUN22/AUG.00013					mit Kindern übergreift“, sagte der Immunologe Alfredo Corell von der Universität Valladolid der Zeitung El Mundo.		
28	NUZ22/AUG.00013					mit Kindern übergreift“, sagte der Immunologe Alfredo Corell von der Universität Valladolid der Zeitung El Mundo.		

Ilustración 7: Resultados en visualización KWIC ordenados cronológicamente

Por último, la pestaña Export (Ilustraciones 8 y9) nos permitirá exportar nuestros resultados en diversos formatos.

Aktuelles Archiv:	W - Archiv der geschriebenen Sprache	Aktuelles Korpus:	W-öffentlich - alle öffentlichen Korpora des Archivs W (mit Neuaquisitionen) [1]
Aktuelle Suchanfrage:	Universität /w1 Valladolid	Referenz:	Deutsches Referenzkorpus DeReKo-2023-1
Treffer:	28	Aktive Treffer:	<input type="text"/>

Archiv	Korpus	Such.	Wortform.	Ergebnisse	Kook.	KWIC	Volltext	Export
Allgemeines								
Dateiname:	<input type="text" value="nachhaltigkeit"/>	Schriftgröße:	<input type="text" value="8"/>					
Dateiformat:	<input type="text" value="RTF"/>	Orientierung:	<input type="radio"/> Querformat <input checked="" type="radio"/> Hochformat					
<input type="checkbox"/> Ergebnisansicht	<input type="text" value="Korpusansicht"/>							
<input checked="" type="checkbox"/> KWIC	gemeinsame Optionen für KWIC und Volltext							
vorher:	<input type="text" value="0"/>	Sätze	Kontext umschließt:	<input checked="" type="radio"/> gesamten Treffer				
nachher:	<input type="text" value="0"/>			<input type="radio"/> 1. Wort des Treffers				
<input checked="" type="checkbox"/> Volltext								
vorher:	<input type="text" value="0"/>	Absätze	Sortierung:	<input type="text" value="KWIC (unsortiert)"/>				
nachher:	<input type="text" value="0"/>			<input type="checkbox"/> Nur ausgewählte Treffer				
				<input type="checkbox"/> Einschränken auf die ersten: <input type="text" value="500"/>	Treffer			
Annotationen:	<input type="radio"/> einblenden <input checked="" type="radio"/> ausblenden							

Ilustración 8: Pestaña de exportación de resultados

© Leibniz-Institut für Deutsche Sprache, Mannheim
 COSMAS II-Server, C2API-Version 4.23.10 - 08.01.2024
 glemm - a German Lemmatizer, version R-2.4.4, 04.07.2022 by Cyril Belica

Datum : Freitag, den 26. Januar 2024, 10:45:58
 Archiv : W - Archiv der geschriebenen Sprache
 Korpus : W-öffentlich - alle öffentlichen Korpora des Archivs W (mit Neuaquisitionen)
 Archiv-Release: Deutsches Referenzkorpus DeReKo-2023-I
 Suchanfrage : (Universität /2w,Max Valladolid)
 Suchoptionen : Ei+Ri+D1, Flex
 Ergebnis : 40 Treffer

KWIC (unsortiert)

Anz. Treffer : 40
 Anz. exportierte Zeilen: 40
 Angezeigter Kontext : 0 Sätze links, 0 Sätze rechts
 Kontext umschließt : gesamten Treffer

T11 Institut der zentralspanischen universität Valladolid (Inea) einen Bihof
 Z66 Kunsthistoriker Pedro de Paloi von der universität Valladolid.
 WP011 und an der Päpstliche Päpstlichen universität von Valladolid, ehe er in Astorga
 WP011 Sie schrieb und gab Kurse an der universität Valladolid.
 WP011 María Cristina, Geschichte an der universität von Valladolid und klassische
 WP011 Er wird heute in der Bibliothek der universität Valladolid aufbewahrt (MS 423) und
 WP011 Die universität von Valladolid wurde im 13.
 WP011 Anglistik und Germanistik an der universität Valladolid.
 WP011 situierten Familie und studierte an der universität von Valladolid.
 WP011 und von 1601 bis 1604 Theologie an der universität Valladolid.
 WP011 universität Valladolid universität Vigo
 WP011 universität Barcelona und 1996 an der universität Valladolid und lehrte das
 WP011 Er studierte Zivilrecht an der universität Valladolid in Valladolid
 WP011 universität Valladolid in Valladolid Valencia universität Valencia
 WP011 universität Miguel de Cervantes in Valladolid universität Vic
 WP011 Philosophie und Rhetorik der universität Valladolid berufen, deren Rektor er
 WP011 Ehrendoktorat der universität von Valladolid (1983)
 WP011 gehörte er bis 1968 zum Lehrkörper der universität Valladolid und arbeitete in
 WP011 der universität Alicante, der universität Valladolid und der universität
 WP011 1952 verbrachte er zwei Semester an der universität Valladolid.
 WP011 ein Studium der Philosophie an der universität Valladolid.
 WP011 und Literaturwissenschaften an der universität Valladolid und schloss dieses
 WP011 katholisch-theologischen Fakultät der universität Valladolid an und arbeitete in der
 WP011 Anklagen wegen Sodomie erst an die universität von Valladolid und bald danach in
 WP011 beherbergt er heute das Rektorat der universität von Valladolid, das
 WP011 Ehrendoktorat der universität von Valladolid (1969)
 WP011 außerdem Kurse für Kino-Ästhetik an der universität Valladolid (1990) und kreatives
 WP011 des Campustells in Burgos der universität Valladolid hervorging.
 WP011 1973 bis 1977 lehrte sie Deutsch an der universität Valladolid.
 WP011 Die universität von Valladolid ist eine öffentliche
 WP011 von Valladolid ist eine öffentliche universität in Valladolid, in der Comunidad
 WP011 Die Geschichte der universität von Valladolid ist mit der
 WP011 WEITERLEITUNG universität Valladolid
 WP011 S75 für das ganze Semester geschlossenen universität Valladolid zu erzwingen, zu einem
 SOL17 sagt die Migrationssoziologin der universität Valladolid.
 U20 geforscht, wie hier an der spanischen universität Valladolid. foto: manso/afp
 NUN22 der Immunologe Alfredo Corell von der universität Valladolid der Zeitung El Mundo.
 NUZZ2 der Immunologe Alfredo Corell von der universität Valladolid der Zeitung El Mundo.

Ilustración 9: Resultados exportados en formato PDF

6. Bibliografía

General

Belica, C., Kupietz, M., Lungen, H. y Witt, A. (2010). The morphosyntactic annotation of DEREKO: Interpretation, opportunities and pitfalls. M. Konopka, J. Kubczak, C. Mair, F. Šticha, and U. Wassner (eds.), *Selected contributions from the conference Grammar and Corpora 2009*, Tübingen. Gunter Narr Verlag.

Bopp, S. (2010). Einführung in die Korpuslinguistik mit DeReKo und COSMAS II (https://www.academia.edu/34491540/Einf%C3%BChrung_in_die_Korpuslinguistik_mit_DeReKo_und_COSMAS_II)

Kupietz, M., Belica, C., Keibel, H. y Witt, A. (2010). The German Reference Corpus DEREKO: A Primordial Sample for Linguistic Research. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Kupietz, M., Lungen, H., Kamocki, P. y Witt, A. (2018). The German Reference Corpus DeReKo: New Developments – New Opportunities. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Lemnitzer, L. y Zinsmeister, H. (2006). *Korpuslinguistik. Eine Einführung*. Tübingen. Narr.

Lungen, H. y Kupietz, M. (2014). Das Deutsche Referenzkorpus DEREKO im Jubiläumsjahr 2014. *Sprachreport*, 24-28.

Perkuhn, R., Belica, C., Keibel, H., Kupietz, M., Lungen, H. (2015). Valenz und Kookurrenz. Domínguez Vázquez, J. M. y Eichinger, L. M. (Eds.), *Valenz im Fokus: grammatische und lexikografische Studien; Festschrift für Jacqueline Kubczak*. Mannheim: Institut für Deutsche Sprache, 175-196.

Algunas Investigaciones realizadas con DeReKo

Abel, A. y Zanin, R. (Eds) (2011). *Korpora in Lehre und Forschung*. Bozen : Bozen-Bolzano University Press. - URN: urn:nbn:de:0111-pedocs-253204 - doi: 10.25656/01:25320

Gondring, O. (2010). Diachroner Wandel von Kol- lokationsprofilen: Zur Emergenz sprachlicher Regelhaftigkeit am Beispiel von Neologismen. Gastvortrag am 14.01.2010, Gesprächsrunde, Institut für Deutsche Sprache, Mannheim.

Mansilla, A. (2017). Lingüística de corpus y fraseología contrastiva (alemán-español): Las combinaciones usuales de estructura [PREP + S]. El caso de entre lágrimas y unter Tränen". *Quaderns de Filologia: Estudis Lingüístics* 22: 147- 164. doi: 10.7203/qf.22.11306

Moroni, M. (2021). Zum Einsatz von Daten aus FOLK und DeReKo für die Untersuchung und Didaktisierung der Formen nach dem Motto und von wegen. *KorDaF – Korpora Deutsch als Fremdsprache*, 2021, 1 (1). doi: 10.26083/tuprints-00019483

Tutoriales

Duffner, R. (2007). COSMAS II – Wegleitung (http://unine.ch/files/live/sites/allemand/files/shared/documents/COSMAS_II_Wegleitung.pdf)

Flinz, Carolina (2021). Titel KORPORA in DaF und DaZ: Theorie und Praxis. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 26: 1, 1–43. <http://tujournals.ulb.tu-darmstadt.de/index.php/zif>

Hartmann, S. (2018), Einsteiger-Tutorial zu Korpusrecherchen mit COSMAS II (https://hartmast.github.io/sprachgeschichte/begleitmaterial/04-CosmasII_Tutorial.pdf)

Vinckel-Roisin, Hélène (2019). Das *Deutsche Referenzkorpus*: Präsentation und Verwendungsmöglichkeiten (<https://lettres.sorbonne-universite.fr/sites/default/files/media/2021-04/Vinckel-Roisin%20H%C3%A9l%C3%A8ne%202021%20cv%20avril%202021.pdf>)

Zinsmeister, H. (2017). Korpuslinguistik. COSMAS II Hilfe – Übersicht (<https://www.korpuslab.uni-hamburg.de/lehrmaterialien/diverses/cosmas2-hilfe.pdf>)