

Ejercicios con Cosmas II¹

1. Variantes ortográficas de *Yoghurt/Joghurt/Jogurt*
 - a) Analice las diferentes grafías de la palabra *Yogurt* e identifique la más frecuente.
 - b) ¿Qué género tiene la citada palabra?
 - c) ¿Podría identificar en qué documento aparece por primera vez la palabra y con qué género?
2. Cada vez es más frecuente la presencia de anglicismos en la lengua alemana. El lenguaje específico de la informática es uno de los más afectados por dicha tendencia.
 - a) ¿cómo forman el Partizip II verbos como *downloaded* o *scannen*?
 - b) Es evidente el incremento de la presencia de anglicismos a lo largo del tiempo, pero justifique ese aumento con datos más precisos.
3. ¿Cómo podría realizar una búsqueda de un verbo con prefijo separable en Cosmas II?
4. Expresión de causalidad con las conjunciones *weil* y *da*.
 - a) La búsqueda en Cosmas II de *weil* no conlleva excesivas dificultades, sin embargo, ¿cómo podría determinar el uso de la conjunción causal *da* y diferenciarlo del resto de usos?
5. ¿Cómo realizaría la búsqueda de dos sinónimos en un mismo párrafo?
6. Búsqueda de un sintagma nominal con adjetivo en todas sus flexiones. (*Das gute Beispiel, die besten Beispiel...*)
7. Cómo realizaría una búsqueda del pronombre relativo *der*?
8. ¿Cómo buscaría las citas a los partidos políticos Sumar y Podemos en el corpus de la *Süddeutsche Zeitung*?
9. Describa el uso del término *Schiri* en lugar de *Schiedsrichter* en la prensa alemana (tipo de prensa, cuándo comienza a usarse, etc)
10. Comparar el uso de la palabra *Klima* en dos décadas diferentes.

¹ Los materiales y ejercicios propuestos se hacen en el marco del PID Historia de la Lengua Alemana 2.0.

Soluciones

Ejercicio 1

Buscamos todas las ocurrencias de las variantes ortográficas "Yoghurt", "Joghurt" y "Jogurt" en el corpus completo de Dereko.

Para cada una tendremos un número de resultados:

Yoghurt	686
Joghurt	25823
Jogurt	1710

Podemos realizar la siguiente búsqueda:

&d- /+w1 ?oghurt

El operador & busca todas las formas del artículo.

/+w1 Indicaría que la palabra siguiente ha de ser Joghurt

?oghurt El operador ? Indica que ha de buscarse cualquier grafía

Ordenamos cronológicamente los resultados (KWIC (chron. sortiert) y comprobamos que la primera mención de Joghurt se realiza en Der Spiegel (23-01-1963, pág. 56) y con el género masculino.

Ejercicio 2

Podemos usar el operador & junto a la forma verbal para determinar las formas conjugadas de cada anglicismo

&downloaden

Para buscar ambas formas de perfecto podemos hacerlo con el comando:

gedownloadet /w8 &haben

Con ello buscamos la combinación de cualquier forma del verbo *haben* (&) con el participio a una distancia de 8 palabras (podríamos definir una mayor).

Podemos buscar la forma alternativa *gedownloaded* y *downgeladen* para comprobar la frecuencia de las mismas.

&scannen

Este comando busca todas las formas del verbo "scannen" en el corpus.

Para determinar la frecuencia temporal elegimos la pestaña Volltext y realizamos la organización de los datos KWIC por orden cronológico. De este modo comprobamos que la primera aparición del verbo es el 15 de junio de 1999 en el St. Galler Tagblatt.

Ejercicio 3

Si usamos el operador & junto al verbo:

&aufstehen

obtendremos todas las formas posibles del verbo unido al prefijo.

Si deseamos obtener los resultados con el prefijo separado, debemos usar el operador /s para obtener verbo y prefijo en la misma frase:

&stehen /+s0 auf

Ejercicio 4

Esto puede implicar buscar patrones donde "da" es seguido por un sujeto y un verbo, o por un sujeto y una forma de "sein" (ser/estar) o "haben" (tener). Para ello cambiamos de corpus y buscamos en: TAGGED-C - Archivo de corpus morfosintácticamente anotados (CONNEXOR) La búsqueda podría ser algo así:

da [pos="NN"] [pos="VVFİN"]

Aquí, [pos="NN"] representa un sustantivo y [pos="VVFİN"] un verbo conjugado

En COSMAS II, además del nivel de palabra, tenemos a disposición las clases de palabras. Sin embargo, las clases de palabras no pueden consultarse directamente en los archivos DeReKo W - W4, sino en los archivos más pequeños anotados TAGGED-C, TAGGED-C2, TAGGED-T y TAGGED-T2.

Si quisiéramos buscar en estas últimas categorías como SUBJ, tendríamos que reescribirlas indirectamente utilizando patrones de clases de palabras.

Por ejemplo, en el archivo TAGGED-T:

da /+w1,s0 MORPH(N)

o también

da /+w1,s0 (MORPH(ADJ) /+w1,s0 MORPH(N))

Ejercicio 5

Para buscar sinónimos en un mismo párrafo usaremos el operador /s. Por ejemplo:

&Palabra /s2 &sinonimo

Asumimos una distancia de dos frases en la búsqueda del sinónimo.

&programmieren /s2 &coden

Ejercicio 6

(&d- /+w1 &gut) /+w1 &Beispiel

Los paréntesis marcan el ámbito de referencia cuando hay más de un operador de distancia.

Ejercicio 7

Se busca la palabra *der* como pronombre relativo en dativo, por ejemplo.

En este ejemplo, la palabra *der* está vinculada a varias condiciones, que pueden formularse a través de una expresión de búsqueda de anotación. Por lo tanto, se utiliza el operador de distancia entre palabras como operador de enlace.

der /w0

A continuación, se debe generar la expresión de búsqueda de anotación que falta.

En el asistente morfosintáctico, seleccione la clase de palabra Pronombre y, luego, en la característica Tipo, seleccionar el valor relativo, y en la característica Caso, seleccionar el valor Dativo.

El asistente morfosintáctico generará a partir de esto la expresión de búsqueda

MORPH(PRN rel dat)

y la copiará a la consulta de búsqueda al salir del asistente.

der /w0 MORPH(PRN rel dat)

Ejercicio 8

Primero, se debe asegurar de que la búsqueda se limita al corpus de la *Süddeutsche Zeitung* (sz - Süddeutsche Zeitung, 1992 - 2023)

Luego, se realiza la búsqueda de los términos "Sumar" y "Podemos".

Podemos usar el operador /s y darle una distancia de 2 ó 3 frases para comprobar el uso en el mismo párrafo.

Ejercicio 9

Realizamos una búsqueda en el corpus para encontrar todas las instancias de ambos términos en el corpus general (con el operador UND)

Schiri UND Schiedsrichter

En la pestaña *Ergebnisse* podemos seleccionar la vista según el tipo de texto (*Ansicht nach Textsorten*) y determinar cual de ellos es más frecuente en la prensa deportiva.

Ejercicio 10

Para este ejercicio tenemos múltiples opciones. Una podría ser comprobar los distintos compuestos de sustantivos existentes con la palabra *Klima*.

*Klima**

El resultado son más de 25000 formas compuestas con la citada palabra. Podemos delimitar los resultados escogiendo compuestos concretos (*Klimawandel, Klimaschutz, Klimakrise, etc.*) o buscando la combinación del término con otras palabras (*Reduktion, Nachhaltigkeit, Emissionen, Katastrophe, Erwärmung, Klimagipfel, etc.*)

Posteriormente podemos ordenar cronológicamente los resultados y ver el diferente tratamiento.

Bibliografía y materiales de ayuda

- Andresen, M y Zinsmeister, H. Bearbeitungs- und Lösungshinweise für die Aufgaben in Korpuslinguistik (Narr Starter) von Melanie Andresen und Heike Zinsmeister (https://meta.narr.de/9783823382263/Starter_Korpuslinguistik_Loesungshinweise.pdf)
- Bodmer Mory, F. (2014). Mit COSMAS II "in den Weiten der IDS-Korpora unterwegs". Steinle M. Y Berens F.J. (eds.). Ansichten und Einsichten. 50 Jahre Institut für Deutsche Sprache. Mannheim. IDS. 376-385.
- Bubenhofer, N. (2006-2024). Einführung in die Korpuslinguistik: Praktische Grundlagen und Werkzeuge. Elektronische Ressource: <http://www.bubenhofer.com/korpuslinguistik/>.
- Duffner, R. (2007). Cosmas II - Wegleitung. Université de Neuchâtel. (http://unine.ch/files/live/sites/allemand/files/shared/documents/COSMAS_II_Wegleitung.pdf)
- Muñoz Acebes, F. J. (2024). Introducción al Deutsches Referenz Korpus (DeReKo). <https://uvadoc.uva.es/handle/10324/65452>
- Muñoz Acebes, F. J. (2024). Operadores de búsqueda frecuentes en Cosmas II. <https://uvadoc.uva.es/handle/10324/70353>
- Schiller, A., Teufel, S., Stöckert, C. y Thielen, C. Guidelines für das Tagging deutscher Textcorpora mit STTS. Technical report, Institut für maschinelle Sprachverarbeitung, Stuttgart, 1999. URL <http://www.ims.unistuttgart.de/projekte/corplex/TagSets/stts-1999.pdf>.
- Zinsmeister, H. (2017). Cosmas II Hilfe - Übersicht. (<https://www.korpuslab.uni-hamburg.de/lehrrmaterialien/diverses/cosmas2-hilfe.pdf>)