



Consulta del Corpus PRESEEA

La muestra

El Corpus PRESEEA que se ofrece en esta página web está integrado por una muestra de 18 entrevistas de cada comunidad investigada. Estas entrevistas tienen la forma de conversación semidirigida y proceden de un hombre y una mujer de tres grupos de edad (20-34 años / 35-54 años / 55 años o más) y de tres niveles educativos diferentes (sin estudios o con estudios primarios / con estudios medios / con estudios superiores) de cada comunidad.

Los resultados de las búsquedas en el Corpus PRESEEA van acompañados de un código que permite identificar a los hablantes según su comunidad de origen, sexo, edad y nivel educativo, con las siguientes claves:

Códigos identificativos de cada comunidad: por ejemplo, ALCA (Alcalá de Henares), MADR (Madrid), BOGO (Bogotá)...

Código de sexo/género: H (Hombre) / M (Mujer)

Código de edad: 1: 20-34 años / 2: 35-54 años / 3: 55 años o más

Código de nivel educativo: 1: sin estudios o con estudios primarios / 2: con estudios medios / 3: con estudios superiores

Muestra de hablantes por comunidad en la página web del Corpus PRESEEA

		NIVEL EDUCATIVO		
		1	2	3
EDAD	1	H / M	H / M	H / M
	2	H / M	H / M	H / M
	3	H / M	H / M	H / M

Por ejemplo, ALCA_H11_037 se interpreta del siguiente modo: entrevista realizada en Alcalá de Henares, con un hombre del primer grupo de edad y con estudios primarios; el número final indica que se trata del hablante 37 de los 54 entrevistados para el corpus PRESEEA-Alcalá de Henares.

Las entrevistas que aquí se ofrecen no constituyen la totalidad de los materiales del Corpus PRESEEA, sino una muestra representativa del español hablado en cada comunidad extraída con fines de investigación y educativos. Para aquellas investigaciones que requieran más materiales de lengua hablada, los interesados deben solicitarlos a los coordinadores de cada equipo, dado que todos los materiales son propiedad de los equipos que los recolectaron.

La búsqueda

En CONSULTA BÁSICA, se introduce la secuencia de caracteres que se desea buscar dentro del corpus, con la posibilidad de incluir las etiquetas que se consideren adecuadas.

En CONSULTA AVANZADA, se puede especificar la forma ortográfica que se desea buscar o su lema. En este último caso, el lema incorpora todas las variantes morfológicas de una palabra: por ejemplo, el lema *perro*, permitirá recuperar las formas *perro*, *perra*, *perros* y *perras*. Asimismo, se puede especificar la categoría gramatical de las formas que se desean buscar.

En FILTROS DISPONIBLES, pueden seleccionarse las comunidades cuyos materiales quieren ser consultados y los perfiles sociolingüísticos de los hablantes con los que se realizaron las entrevistas. El sistema de búsqueda permite recuperar materiales de una, de varias o de todas las comunidades incluidas en el corpus.

Al realizar las consultas, se obtendrá el listado de ocurrencias numeradas, con cada ocurrencia en su concordancia y alineada con el audio correspondiente. Junto a ellas, se ofrece la posibilidad de descargar la transcripción de cada entrevista y una muestra de audio de 10 minutos. Asimismo, es posible descargar un fichero CSV con el resultado de cada búsqueda.