

UN CORPUS ORAL EN LÍNEA COMO RECURSO DIDÁCTICO

Eli-Marie Danbolt Drange¹
Universidad de Bergen

Introducción

La expresión oral es una de las destrezas más importantes para desarrollar la competencia E/LE, pero el hecho de aprender la lengua fuera de su contexto dificulta el proceso de aprendizaje de esta destreza. En esta comunicación quiero presentar el corpus COLA (Corpus Oral de Lenguaje Adolescente), un corpus oral de lenguaje adolescente en línea, desarrollado en el Departamento de Lenguas Extranjeras de la Universidad de Bergen, Noruega, bajo la dirección de la Dra. Annette Myre Jørgensen. Primero voy a presentar el corpus en sí, y después mostraré algunos ejemplos de cómo se puede emplear este corpus como un recurso didáctico para la enseñar la expresión oral en clase.

En primer lugar es necesario aclarar algunos conceptos, como por ejemplo: ¿Qué es un corpus? En este contexto se entiende un corpus como una selección de textos, un conjunto de textos. Los textos pueden ser de distintos tipos; hay corpus de textos periodísticos que recogen todo el material publicado en ciertos periódicos; hay corpus de textos literarios, e incluso hay corpus de textos literarios traducidos. Los corpus más conocidos en el ámbito español son quizás los dos corpus de la Real Academia Española: el Corpus de Referencia del Español Actual (CREA) y el Corpus diacrónico del español (CORDE). El corpus *CREA* contiene textos pertenecientes a los últimos veinticinco años de historia del español, mientras que el *CORDE* se ocupa de todo lo demás (Sánchez Sánchez 2005: 37).

El corpus COLA

El corpus COLA es un corpus de conversaciones informales orales de jóvenes, y el objetivo del corpus es reunir una base de datos de conversaciones de lengua hablada informal de jóvenes de las principales capitales hispanohablantes. Actualmente el corpus cuenta con conversaciones grabadas de Madrid, Santiago de

¹ Eli-Marie Danbolt Drange es becaria de doctorado del Departamento de Lenguas Extranjeras de la Universidad de Bergen, Noruega, donde trabaja con una tesis sobre el empleo de anglicismos en el lenguaje juvenil chileno y noruego, Eli-Marie.Drange@if.uib.no Es integrante del grupo de investigación COLA (Corpus Oral de Lenguaje Adolescente), véase <http://www.colam.org>, y sus áreas de investigación son el lenguaje juvenil, el análisis del discurso oral y la enseñanza de E/LE.

Chile, Buenos Aires y Ciudad de Panamá. Para seleccionar a los jóvenes que participaron en las grabaciones, se buscaron escuelas que cumplieran con ciertos requisitos predefinidos. Así se ha logrado encontrar informantes de entre 13 y 19 años, de ambos sexos y de diferentes niveles sociales (Jørgensen 2007: 226). El personal docente de las escuelas seleccionadas asistió en la búsqueda de jóvenes voluntarios para participar en el proyecto. Luego estos jóvenes recibieron un minidisco y un micrófono de solapa, y se les pidió grabar conversaciones con otros jóvenes durante más o menos una semana. Para obtener un lenguaje lo más informal y auténtico posible, se empleó este método usado en la elaboración de una parte del corpus nacional británico BNC, que recoge diferentes textos escritos y orales del inglés británico (Crowdy 1995).

El corpus COLA consiste entonces principalmente en grabaciones de conversaciones informales entre jóvenes. Una gran parte de las conversaciones se han grabado en la escuela, en clase y en el recreo, mientras también hay conversaciones grabadas en la calle, en parques u otros lugares de reunión de jóvenes, y también en las casas de los informantes. En muchas grabaciones hay bastante ruido de ambiente, y en algunas conversaciones participan los padres en casa y profesores en la escuela. Pese a estos inconvenientes, la mayor parte del corpus muestra cómo los jóvenes se comportan lingüísticamente en un contexto natural.

Después de grabar las conversaciones, éstas fueron trasladadas a un formato digital y transcritas con el programa *Transcriber*. La ventaja de este programa de transcripción es que se sincronizan los archivos sonoros con el texto transcrito, y así se puede leer la transcripción y escuchar el sonido al mismo tiempo. Las conversaciones están transcritas ortográficamente, lo que significa que en las transcripciones se respetan las reglas ortográficas independientemente de la pronunciación de las palabras (Drange en proceso: 90). Así las conversaciones se han convertido en textos, y de estos textos se han hecho archivos HTML con enlaces a los archivos de sonido. Estos archivos finalmente se han trasladado al formato de Corpus Workbench en Internet (Hofland 2005). Los archivos se pueden bajar desde las páginas web de proyecto COLA: <http://www.colam.org> como muestra la imagen a continuación:



Para tener acceso al corpus, hay que tener una clave que se obtiene rellenando una pequeña solicitud que está en las mismas páginas web. Como el material incluye referencias personales, estamos obligados a mantener el anonimato de los informantes y el material solamente se puede emplear en la investigación y en la enseñanza.

¿Para qué sirve el corpus?

Ya he repasado las características del corpus COLA, y es el momento de preguntar por la utilidad del mismo. En primer lugar el corpus sirve para comprobar cómo se habla en una situación informal. Como el corpus es de conversaciones orales juveniles, se puede estudiar las palabras y expresiones típicas del lenguaje juvenil, incluyendo palabras malsonantes. Además se puede ver y escuchar cómo hablan los jóvenes de Madrid, de Santiago de Chile, etc.

Uno de los objetivos del plan de estudios de lengua extranjera en Noruega, es obtener competencia comunicativa y poder comunicarse también en situaciones espontáneas.² El hecho de aprender una lengua fuera de su contexto, dificulta el

² Véase la página web:

http://www.udir.no/templates/udir/TM_Læreplan.aspx?id=2100&laereplanid=123914&visning=5

aprendizaje de una lengua espontánea. Así, este corpus de conversaciones informales es una manera de tener acceso a situaciones auténticas. Las conversaciones grabadas pueden servir para ilustrar cómo se habla en un contexto informal. El material se puede emplear para mejorar el nivel de comprensión auditiva y lograr una expresión oral más auténtica, o sea que el corpus se puede utilizar para perfeccionar dos de las cuatro destrezas lingüísticas.

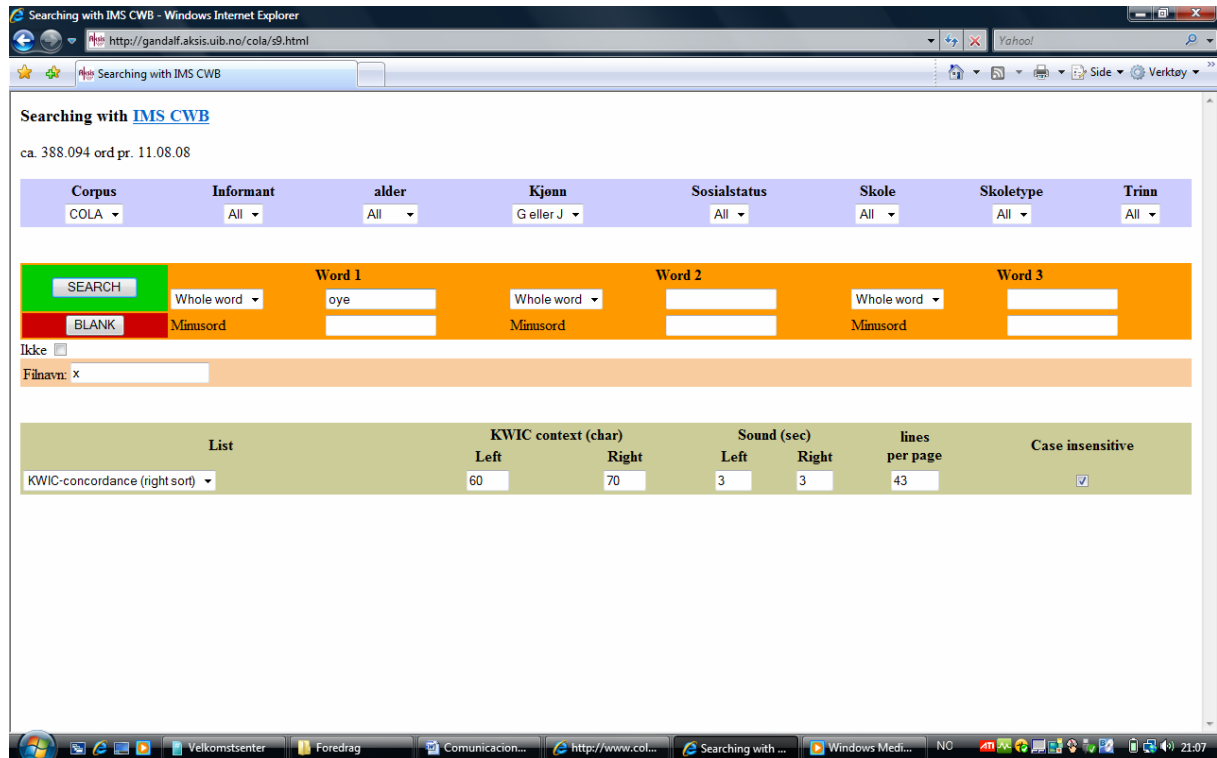
Cómo trabajar con el corpus

Se puede proceder de dos maneras para tener acceso al material; se puede hacer una búsqueda de palabras o expresiones definidas, y se puede estudiar una conversación o un extracto de una conversación como una unidad definida. En lo que sigue voy a mostrar ejemplos de lo que se puede hacer con cada una de estas entradas.

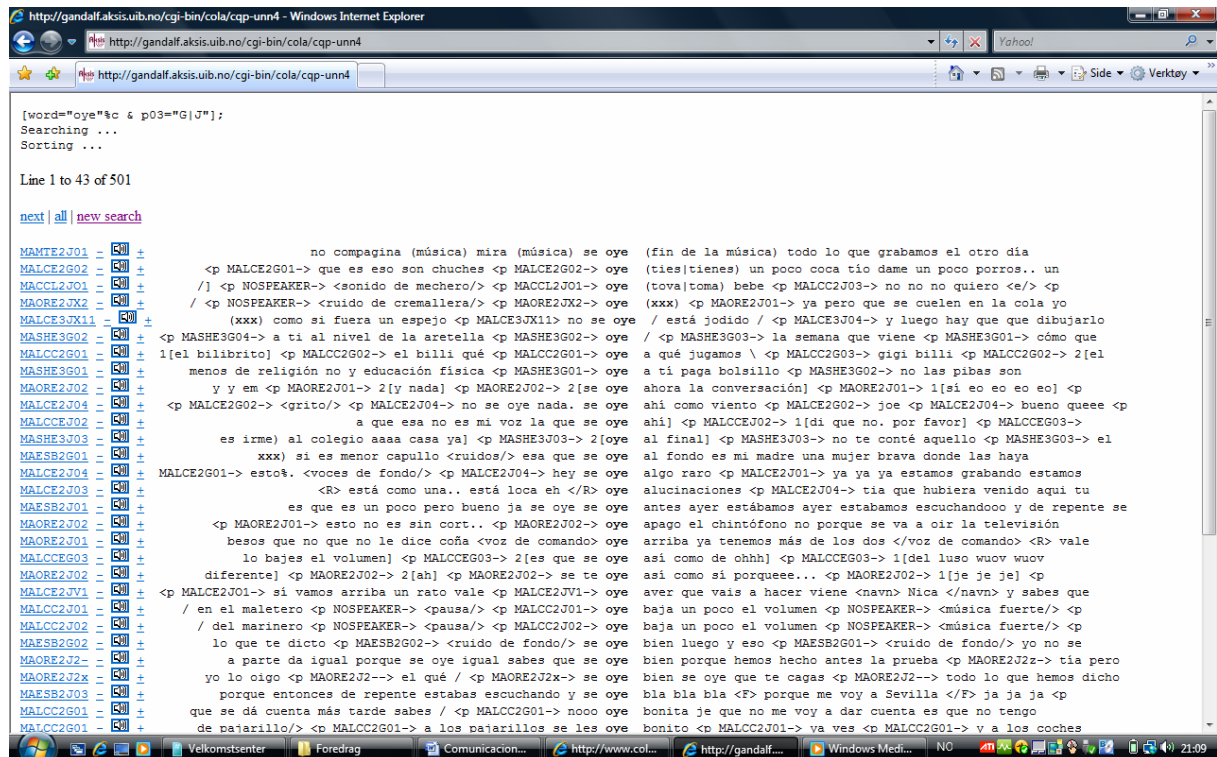
Un desafío en la enseñanza de E/LE es explicar a los estudiantes o alumnos las expresiones típicas del habla informal. Dos expresiones que siempre aparecen en los libros de texto de E/LE, pero que no son fáciles de explicar son: el uso de verbos de percepción con función apelativa y los apelativos. Un ejemplo de lo primero es el uso de *oye* para llamar la atención, mientras un ejemplo de lo segundo es el empleo de diferentes apelativos para reforzar la comunión fática entre los interlocutores (Stenström y Jørgensen 2008).

Búsqueda de “oye”

Para estudiar el uso de “oye”, es interesante hacer una búsqueda en el material. En el buscador se especifican los criterios de la búsqueda:



Mientras el resultado de la búsqueda se aprecia en la siguiente imagen:




Primero vemos que hay 501 ejemplos de esta palabra en un corpus de unas 400 000 palabras, por lo que se puede considerar una palabra con cierta frecuencia. El primer ejemplo de esta búsqueda coincide con el significado léxico del verbo "oír",

pero también se aprecia que muchos de los ejemplos de oye cumplen la función de llamar la atención. Para facilitar la comprensión de este verbo y su uso en clase, podemos seleccionar algunos ejemplos que indican cómo se usa esta expresión:

Oye

- Para llamar la atención:
 - Luisa: **oye** habla claro porque me he dado cuenta las otras grabaciones que no se te oía tía
 - Pedro: **oye** ricardo dime algo
 - Julia: **oiga** richi mañana en mi casa estúpido
 - Juan: **oye** chiquillas ¿no chatean no hacen nada?
 - Jorge: **oye** loco <xxx/> escúchameeeee




www.uib.no

Aquí también es posible escuchar las expresiones, si se mantiene el enlace con el archivo de sonido.

Una muestra como ésta, también nos da pie para presentar otro recurso frecuente para llamar la atención: los apelativos. En la imagen vemos que los apelativos pueden ser nombres propios, hipocorísticos y vocativos. El corpus COLA muestra que en el lenguaje juvenil español los vocativos más frecuentes son “tío” y “tronco”, mientras en el lenguaje juvenil chileno sobresale “loco”.

El corpus también se puede usar para ilustrar la creatividad para crear apodos y vocativos en español, tal como vemos en la siguiente imagen que muestra ejemplos de anglicismos empleados como apodos y vocativos en el lenguaje juvenil chileno.

Vocativos y apodos	
VOCATIVOS	APODOS
batman	barbi
companion	barni
couboy	eagle
gordineitor	goofy
highlander	jeidi
madriman	mofasa
taxiboy	revolution
toplero	robinson
pokemon	tom sawyer
	spinak
	supermac



www.uib.no

Este empleo de nombres y sobrenombres para llamarse es un elemento fundamental en las conversaciones entre jóvenes hispanohablantes. Al comparar este uso con un corpus de lenguaje juvenil noruego en el trabajo con mi tesis doctoral encontré muy pocos ejemplos de este tipo de apelativos en noruego (Drange en proceso). Este hecho indica una diferencia fundamental entre el español y el noruego en este sentido, que es interesante abordar en un contexto E/LE. Stenström y Jørgensen (2008) encontraron la misma diferencia en las estrategias comunicativas de los jóvenes ingleses y españoles, lo cual muestra una comunión estrecha entre los jóvenes españoles; una manera de involucrar al interlocutor que no está presente en las conversaciones entre jóvenes noruegos e ingleses. Este punto es interesante abordar al comparar las estrategias lingüísticas de la lengua extranjera con la lengua materna, otro de los objetivos del plan de estudio antes mencionado.

Otro de los verbos de percepción que tienen diferentes significados en español es “venga”. En la siguiente imagen vemos algunos ejemplos de distintos usos de este verbo, que no siempre significa precisamente ‘venir’:

Venga

- Estímulo a la acción:
 - Neus: ya da igual **venga** corre corre
 - Juli: **ah venga** cántate un rap cántate un rap
- Rechazo o protesta:
 - Eva: no te has enterado de que estaba embarazada
 - Sandra: **venga** ya
 - Celia: jodeeeeeer
- Complicidad o acuerdo:
 - Juana: ya está
 - Frida: vale **venga**



www.uib.ne

Los ejemplos que he mostrado aquí han sido seleccionados por medio de una búsqueda de palabras o expresiones específicas. El objetivo de una búsqueda como ésta es mostrar ejemplos del uso de las palabras seleccionadas para así visualizar las situaciones y el contexto donde aparecen.

Pero el corpus también se puede usar para ilustrar situaciones comunicativas específicas como por ejemplo conversaciones telefónicas, tal como en la próxima imagen:

Conversación telefónica

Teléfono que suena

- Mariana: 📞
- sí dígame..
 - eeh..no..o sea ha salido ya
 - quién eres/
 - hola qué tal...
 - venga...
 - hasta luego..



www.uib.ne

Esta conversación telefónica auténtica sirve para enseñar la estructura de una conversación por teléfono introducida por “dígame”.


Igualmente es posible sacar extractos más largos de conversaciones, como vemos en la siguiente imagen:

¿Qué hiciste ayer?

Ana: **una polla** yo no dormía desde el sábado **tío**
ayer dormí ocho horas **tío** llegué **tronco** con los
ojos hinchados

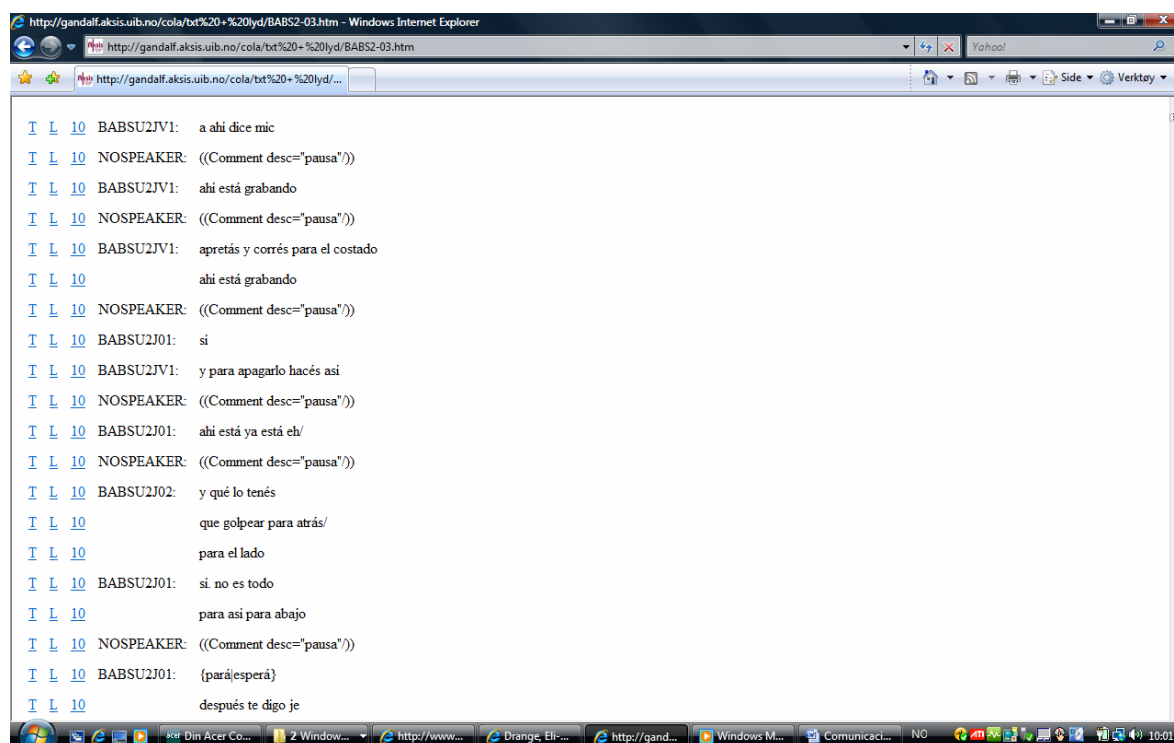
Roberto: a saber {**carías**|que harías} y con quién

Ana: eso es lo que no quiero saber
<R> eso eso lo quiero saber yo </R>
y nada y fuimos ayer al cine



www.uib.ne

En este extracto nos podemos fijar en los elementos marcados con rojo. Estas palabras o expresiones no añaden nada estrictamente informativo a la conversación, y por eso se pueden clasificar como innecesarias. Debido a esto no se suelen incluir en la enseñanza de E/LE, pero al sacarlas cambia la esencia de la conversación. Así nos podemos preguntar hasta qué punto es necesario que nuestros alumnos aprendan esta clase de expresiones para decir que tienen competencia comunicativa en español. También es posible trabajar con extractos más largos de conversaciones, pero por la brevedad de la comunicación lo he dejado fuera. Los extractos pueden servir como pautas para los alumnos, y se les puede incentivar a crear conversaciones similares. Si entramos en los archivos en la página web, encontramos conversaciones enteras, como en el ejemplo a continuación del corpus de Buenos Aires. Aquí también es posible escuchar al mismo tiempo que uno va leyendo la conversación:



Conclusión

En esta comunicación he mostrado el corpus COLA como un recurso para aprender cómo hablan los jóvenes hispanohablantes sin viajar a un país hispanohablante. El corpus BABS puede servir para explicar y mostrar el uso de palabras y expresiones específicas, como el empleo de vocativos o apelativos y los verbos de percepción, o para analizar conversaciones sobre distintos temas. Así el material se puede emplear para aumentar el nivel de comprensión de lengua informal y para incentivar a la conversación informal en clase. También se pueden usar los textos para estudiar diferentes fenómenos gramaticales; los estudiantes pueden buscar los verbos, los adjetivos u otros fenómenos de interés.

Finalmente se pueden emplear las conversaciones para aprender más sobre la cultura de los países hispanohablantes. Se puede estudiar las referencias culturales en las conversaciones y se puede discutir los temas que les interesan a los jóvenes. ¿De qué hablan los jóvenes? ¿Se preocupan de los mismos temas que los jóvenes noruegos? Así el habla de los jóvenes hispanohablantes también puede ser un recurso para aprender más sobre la cultura y las costumbres de estos jóvenes.

Bibliografía

- Crowdy, Steve (1995): "The BNC spoken corpus." En Leech, G.N./ Myers, Greg / Thomas, Jenny (eds): *Spoken English on Computer*. New York: Longman, pp. 224-234.
- Drange, Eli-Marie Danbolt (en proceso): *Anglicismos en el lenguaje juvenil chileno y noruego. Un análisis comparativo*. Tesis doctoral inédita, Universidad de Bergen.
- Hofland, Knut/ Jørgensen, Annette Myre/ Drange, Eli-Marie/ Stenström, Anna-Brita (2005): "COLA: A Spanish spoken corpus of youth language." *The Corpus Linguistics Conference Series (1)*, <http://www.corpus.bham.ac.uk/PCLC/>
- Jørgensen, Annette Myre (2007): "COLA: un corpus oral de lenguaje adolescente". En *Anejos de Oralía*, 3/1, pp. 225-234.
- Sánchez Sánchez, Mercedes (2005): "El corpus de referencia del español actual (CREA). El CREA oral". En *Oralía*, 8, pp. 37-56.
- Stenström, Anna-Brita/ Jørgensen, Annette Myre (2008): "La función fática de los vocativos en la conversación juvenil de Madrid y Londres". En: Briz, Antonio et. al. (eds): *Cortesía y conversación: de lo escrito a lo oral*" *Actas del III Coloquio internacional de EDICE*. Valencia: Universidad de Valencia, pp. 355-365.