LINGÜÍSTICA DE CORPUS Y GRAMÁTICA INDUCTIVA: CÓMO TRABAJAR CON LOS CORPUS EN LA CLASE DE ESPAÑOL DE LOS NEGOCIOS

Inmaculada Martínez Martínez
Centro Internacional de Estudios Superiores del Español (CIESE-Comillas)

Resumen

El presente artículo trata sobre cómo obtener un mayor rendimiento en la clase de español de los negocios si trabajamos la gramática de manera inductiva mediante la herramienta de los corpus textuales.

El objetivo es mostrar de forma concreta cómo una aplicación de la tecnología informática como es un corpus digitalizado, puede llevarse al aula de una clase de Español de los Negocios (ENE) bajo el enfoque conocido como Aprendizaje guiado por datos (*Data-driven Learning*).

La metodología que se emplea es descriptivo-cualitativa. Se aborda, en primer lugar, el necesario marco teórico enclavado en la Lingüística de corpus y el Cognitivismo. A continuación, se define el Aprendizaje guiado por datos y se ofrece una tipología de actividades para su inclusión en el aula de ENE. Posteriormente, se desglosan los principales resultados –en forma de implicaciones didácticas- que conlleva la implementación de este enfoque:

1. Trabajar la **gramática en contexto** a través de las líneas de concordancia, por ejemplo, nos permite realzar mejor para el estudiante la relación entre las formas gramaticales y su uso.
2. El acercamiento a la gramática mediante los corpus puede llevarse a cabo de manera inductiva (**gramática inductiva**). Con ello conseguiremos que el estudiante focalice su atención sobre la lengua del corpus y, al hacerlo, desarrolle una batería clave de mecanismos de procesamiento cognitivo.

Todo ello debería conducir a un incipiente aprendizaje autónomo en el estudiante de ENE a partir de muestras reales de lengua, muestras a las que ellos podrán acudir siempre que lo necesiten para confirmar o refutar sus hipótesis lingüísticas.

1. Introducción

Resulta obvio que la lengua es tan vasta y variada que nunca podremos proporcionarles a nuestros aprendientes una descripción abarcable de esta como un todo. Sí podremos, sin embargo, suministrarles pautas o directrices que los guíen en su camino de aprendizaje (Willis & Willis, 1996). En este sentido, una manera adecuada puede ser mediante actividades que los impulsen a reflexionar sobre muestras de lengua para trazar sus propias conclusiones sobre cómo funciona el sistema lingüístico. Este tipo de

actividades se conocen como actividades de despertar la consciencia (*consciousness raising activities*) y activan en la mente del alumno importantes destrezas cognitivas, clave para su aprendizaje en la segunda lengua (Schmidt, 1990; Fotos, 1994; Martínez, 2004).

Por otro lado, los usos de la lengua que se pasan por alto o resultan implícitos en un solo texto pueden destacarse y ser más explícitos si manejamos la información intertextual que nos proporcionan los corpus lingüísticos y su evidencia (Willis & Willis, 1996; Hunston, 2002; Tsui, 2005). Así, la técnica básica de análisis de un corpus, las líneas de concordancia, puede llegar a convertirse en una valiosa herramienta, pues permite trabajar con un gran número de ejemplos de un ítem, presentes en muchos textos y que estos ejemplos aparezcan juntos en su contexto original (Evison, 2010).

En consecuencia, ambos marcos teóricos –el Cognitivismo y la Lingüística de corpus- unidos parecen contribuir inicialmente a una mejora en la calidad de la enseñanza y el aprendizaje de la lengua, en nuestro caso, del español como lengua extranjera o el del español de los negocios. Si esto es así, se mostraría como una evidencia la necesidad de unir ambas perspectivas, porque consideramos, y esta es la idea central del presente artículo, que trabajar la inducción a través de los corpus produce un aprendizaje superior[i].

2. **Marco teórico**

2.1. La Lingüística de corpus

El profesor Parodi (2008) señala en su artículo sobre la Lingüística de corpus, el anclaje conceptual en el que esta disciplina debe situarse para comprender después mejor el concepto de corpus y sus aplicaciones:

> "La Lingüística de corpus constituye un conjunto o colección de principios metodológicos para estudiar cualquier dominio lingüístico y se caracteriza por brindar sustento a la investigación de la lengua en uso a partir de corpus lingüísticos con sustrato en tecnología computacional y programas informáticos *ad hoc*"
>
> (Parodi, 2008: 95).

No se trata, por tanto, de una rama más de la Lingüística, como la Sintaxis o la Semántica: sirve a estas y otras disciplinas como método de investigación que es. Las investigaciones en él basadas se llevan a cabo, por tanto, mediante la compilación electrónica de textos que contienen información lingüística original, los llamados corpus textuales.

Continuamos con el andamiaje teórico para recordar que:

> "un corpus es un conjunto amplio de textos de los más diversos tipos, representativos del estado de una lengua, en formato electrónico, codificados de modo que sea posible obtener de ellos los casos de una determinada forma o expresión en la totalidad de su conjunto"
>
> (Sinclair, 2004)

Entre las herramientas de los corpus disponibles para trabajar la gramática y el léxico de manera inductiva en clase, nos centraremos en las listas de frecuencia y en las líneas de concordancia. Es importante señalar que, mediante este acercamiento, la lengua se convertirá no solo en objetivo sino también en medio de aprendizaje (Svalberg, 2016).

1) **Las listas de frecuencia** constituyen una relación de todas las palabras del corpus en orden descendente con respecto a su frecuencia de aparición (Gilquin y Granger, 2010). Se muestra en la Figura 1 la lista de frecuencia del *Corpus Comillas del Español de los Negocios* (CORPEN Comillas)[ii] en el estado actual en el que se encuentra:

File Edit View Compute Settings Windows Help

N	Word	Freq.	%	Texts	%	ispersion
1	#	320	9,85	15	100,00	0,95
2	DE	184	5,66	15	100,00	0,94
3	EL	155	4,77	15	100,00	0,95
4	EN	126	3,88	15	100,00	0,90
5	LOS	107	3,29	15	100,00	0,94
6	LA	105	3,23	15	100,00	0,91
7	UN	98	3,02	15	100,00	0,94
8	Y	78	2,40	15	100,00	0,89
9	LAS	66	2,03	15	100,00	0,93
10	DEL	60	1,85	15	100,00	0,89
11	A	58	1,79	15	100,00	0,89
12	Ã	51	1,57	15	100,00	0,93
13	IBEX	46	1,42	15	100,00	0,94
14	POR	41	1,26	15	100,00	0,82
15	SE	39	1,20	15	100,00	0,89
16	QUE	37	1,14	15	100,00	0,91
17	MERCADOS	30	0,92	15	100,00	0,90
18	CON	29	0,89	15	100,00	0,88
19	MAYORES	28	0,86	15	100,00	0,87
20	REGISTRAN	28	0,86	15	100,00	0,89
21	PUNTOS	26	0,80	15	100,00	0,89
22	LO	23	0,71	15	100,00	0,86
23	BOLSA	20	0,62	15	100,00	0,83
24	CIERRE	20	0,62	15	100,00	0,87
25	SUBIDAS	20	0,62	15	100,00	0,85

Figura 1. Lista de frecuencia de CORPEN Comillas

O´Keeffe *et al.* (2007) destacan la importancia de este recurso al señalar que una lista de este tipo será un importante índice del léxico que debería incluir un *syllabus* básico, un manual o una gramática. Para los profesores parece esencial, en definitiva, conocer qué palabras o expresiones se repiten más, de manera que estén seguros de que enfatizan en sus clases las formas frecuentes (Chambers, 2010).

Entre las diversas aplicaciones didácticas de esta herramienta, resulta útil:

- comparar, por ejemplo, dos listas de frecuencia, la de un corpus general y la de otro específico, con el objetivo de sensibilizar al aprendiente sobre la especificidad léxica de los discursos profesionales;

- comparar, así mismo, pequeños corpus especializados de, al menos, dos lenguas. Con ello, el estudiante desarrollará una mirada más profunda al uso intercultural de la lengua en contextos específicos (O´Keeffe *et al*, 2007: 183);
- tomar estas listas como punto de partida para los análisis de concordancias (Gilquin & Granger, 2010) que trataremos a continuación.

2) **Las líneas de concordancia** constituyen la primera técnica analítica básica del trabajo con corpus y la más conocida. Es valiosa porque permite, por un lado, que se visualice un gran número de ejemplos de una forma y, por otro lado, que tales ejemplos aparezcan juntos en su contexto original (Evison, 2010). Cada forma es indexada y, algo muy importante, se incluye la referencia al lugar de ocurrencia en el texto correspondiente (Sinclair, 1991).

Se inserta en la Figura 2 las líneas de concordancia pertenecientes a la búsqueda de la forma "bolsas":

File Edit View Compute Settings Windows Help

N	Concordance	Set	Tag	Word #	Sen	Sen	Para	Para	Hea	Hea	Sec	Sec	File
1	, respectivamente. Por otro lado, Bolsas y Mercados (-0,37%) registra			137	8	16%	0	84%			0	84%	PE2014_0003.x
2	de estrÃ©s a la banca europea, las bolsas cierran con signo mixto pero			25	2	69%	0	10%			0	10%	PE2014_0009.x
3	las alzas registradas en su inicio, las bolsas europeas cerraban en negativo.			82	4	73%	0	40%			0	40%	PE2014_0010.x
4	de cierre de mercados Cautela en las bolsas			11	2	53%	0	5%			0	5%	PE2014_0008.x
5	Pese a la escalada registrada por las bolsas europeas el pasado viernes, al			28	3	28%	0	12%			0	12%	PE2014_0008.x
6	centran la atenciÃ³n en las bolsas europeas. Los mercados			30	2	97%	0	12%			0	12%	PE2014_0012.x
7	crecimiento en EE.UU. impulsa a las bolsas			15	3	17%	0	9%			0	9%	PE2014_0013.x
8	en EE.UU. daban un giro a las bolsas que se recuperaban de las caÃ			52	5	15%	0	31%			0	31%	PE2014_0013.x
9	de mercados Cierre mixto en las bolsas			12	2	37%	0	5%			0	5%	PE2014_0012.x
10	Las bolsas europeas cerraban al alza tras			14	2	33%	0	7%			0	7%	PE2014_0002.x
11	de cierre de mercados Subidas en las bolsas			11	2	37%	0	7%			0	7%	PE2014_0003.x
12	favorables desde Asia y las bolsas europeas no dudaban en			29	2	52%	0	13%			0	13%	PE2014_0001.x
13	ContinÃºan las subidas en las bolsas			13	2	31%	0	6%			0	6%	PE2014_0002.x
14	cierre de mercados Ascensos en las bolsas			11	2	48%	0	7%			0	7%	PE2014_0006.x
15	35 encabezaba las subidas en las bolsas gracias a la banca y a			31	3	34%	0	21%			0	21%	PE2014_0006.x
16	cierre de mercados Descensos en las bolsas			11	2	43%	0	6%			0	6%	PE2014_0004.x
17	del Ibex 35, las compaÃ±Ã-as IAG y Bolsas y Mercados registran las			161	8	63%	0	78%			0	78%	PE2014_0014.x

Figura 2. Líneas de concordancia de la forma "bolsas"

Nos recuerda O´Keeffe *et al.* (2007) que, en el ámbito de los negocios, tanto como aprender a usar una lengua de manera apropiada, deberían ser primordiales la observación detallada y profunda, así como el despertar la consciencia de las formas lingüísticas. Estos procesos cognitivos, que abordaremos a continuación, podemos reforzarlos en el contexto reflexivo de la clase si ponemos la lupa sobre dichas formas realzadas en las líneas de concordancia de los corpus (O´Keeffe *et al*. 2007).

2.2. El Cognitivismo

Entramos de lleno, pues, en el mundo de la reflexión y, con él, en el segundo de los marcos teóricos que sustenta la presente investigación: el cognitivismo. Desarrollaremos, también aquí, los conceptos fundamentales. Analizaremos, en primer lugar, la concienciación formal, término que integra los conceptos de realce del input, el

advertir, la consciencia y la focalización en la forma. Todas estas operaciones se desarrollarán cuando el profesor permita que la forma lingüística sobresalga en el texto y todas ellas serán posibles a través de las muestras de corpus.

La **concienciación formal** parece convertirse en un proceso esencial para la adquisición de la lengua. Centrar la atención sobre un componente gramatical previamente realzado se convierte en concienciación formal en el sentido de que los aprendientes desarrollan conocimiento sobre ese rasgo y, lo más importante, tienden a ser más conscientes de él en el input comunicativo posterior (Fotos, 1994).

Se trataría, en suma, de despertar la consciencia del estudiante sobre los elementos lingüísticos problemáticos en la L2 y hacerlo a través de la interacción y del realce que nos permiten las líneas de concordancia (Svalberg, 2016), es decir, a partir de actividades que permitan el intercambio comunicativo dentro del aula. En definitiva, integrar gramática y comunicación.

Parece claro, por tanto, que todas las operaciones que realice el aprendiente cuando lo expongamos a muestras reales de lengua a través del corpus pasarán necesariamente por la concienciación formal.

Es preciso señalar que esta consciencia sobre el caudal lingüístico de los corpus es solo posible porque previamente el profesor ha resaltado esas formas concretas que quiere que sean advertidas, focalizadas y, finalmente, adquiridas por el aprendiente. Esta operación llevada a cabo por el docente se denomina **realce del *input***.

Ciertos autores como Svalberg, que trabajan en estos días sobre el enfoque, nos hablan de "realce visible del input" (*visual input enhancement*) o de "realce textual del input" (*textual enhancement of input*) (Svalberg, 2016: 403). El propio *software* de las líneas de concordancia permite destacar en contexto determinadas formas al aprendiente de la L2. El profesor puede, a partir de este material ya destacado, agrupar, rodear con un círculo, dar color o subrayar aquellas formas sobre las que desea focalizar la atención de sus estudiantes (Sripicharn, 2010). Todo ello lo mostraremos más adelante a través de modelos concretos de actividades que permiten trabajar estas técnicas en el aula.

El input realzado permitirá que el estudiante lleve a cabo la operación del **advertir** (*Noticing*). Schmidt (1990) sostiene que el aprendizaje subliminal –el llevado a cabo sin ninguna colaboración de la consciencia- es imposible en el caso de los adultos y que advertir es condición necesaria para convertir el *input* en aprendizaje efectivo. Así, el aprendizaje incidental –en el que sí incide o interviene la consciencia- es tanto posible como efectivo, cuando las necesidades de una tarea centran la atención en lo que va a ser aprendido por parte del estudiante de la L2.

Cabe señalar aquí la relación que se establece entre advertir determinadas formas en el *input* y producirlas posteriormente en la comunicación. El estudio de Schmidt y Frota (1986), que analiza la adquisición del portugués-L2 por uno de los autores, confirma que estamos ante una conexión muy estrecha entre **advertir** y **producir** (Schmidt, 1990). Sólo cuando el autor advierte la forma lingüística, comienza a usarla.

Ahora bien, solo el **advertir** no parece suficiente. Para que el procesamiento de la forma sea lo bastante firme –de manera que asegure la retención- deben darse otras

condiciones, condiciones que quizá hayan de pasar por la consciencia y la atención a la forma, conceptos que se analizan a continuación.

El concepto de **consciencia** (*awareness*) es un concepto de plena vigencia que integra, entre otras dimensiones, el conocimiento *sobre* la lengua y no solo la habilidad para usarla (Svalberg, 2016). De ahí que también se emplee con frecuencia el término *consciencia lingüística*. Podría entenderse, así, el aprendizaje como un proceso que comienza en el aprendiente cuando este se da cuenta de que determinado material lingüístico va a ser aprendido, evoluciona a través de una serie de etapas que establecen un contexto para la comprensión de nuevo material, y concluye con la desaparición del nuevo material del campo de la consciencia para convertirse él mismo en parte del contexto subconsciente. Una vez desarrollado, este proceso será el que fije la interpretación de futuros acontecimientos lingüísticos (Martínez, 2004).

En este componente cognitivo de la consciencia, que incluye el conocimiento explícito, la percepción consciente y la sensibilidad lingüística, nos posicionamos a la hora de defender el trabajo con corpus en el aula.

Continuamos ahora con la **focalización en la forma** (*Focus on Form)*. Schmidt (1990) desarrolla la tesis de que la noción de consciencia es útil porque se vincula a otros conceptos relacionados como el de la atención y la atención selectiva y esta última es, precisamente, la que posibilita la focalización en la forma. Este foco en la gramática emerge de la lengua usada en la clase, es decir, en torno a rasgos lingüísticos que surgen de manera natural dentro de las actividades de clase (Hall, 2011) en las que los aprendientes atienden brevemente a la forma, el significado y el uso.

Se pretende capitalizar, así, las propias tendencias analógicas del aprendiz y forzar con ello un cambio en su competencia: si el alumno está atendiendo a la forma, algunas de esas señales se adaptarán a sus actuales creencias metalingüísticas y algunas otras, no (Sharwood Smith, 1993). Esta táctica que se adopta asume el modelo de comprobación de hipótesis y conecta de forma estrecha con las ideas más recientes en torno a los corpus y al estudiante como investigador (Johansson, 2009).

3. El Aprendizaje guiado por datos

3.1. Definición y objetivos

Todo el marco teórico hasta aquí expuesto encuentra su concreción en la clase de español a través del Aprendizaje guiado por datos (*Data-Driven Learning, DDL*). El Aprendizaje guiado por datos es una aplicación del Aprendizaje de la Lengua Asistido por Ordenador (ALAO), e incluye el uso de los corpus textuales para la adquisición de la lengua extranjera. En dicho aprendizaje los estudiantes experimentan el uso de un corpus a través de tareas guiadas o a través de materiales basados en evidencias del corpus, como las líneas de concordancia.

En definitiva, consiste en usar las herramientas y técnicas de los corpus lingüísticos con objetivos pedagógicos (Gilquin & Granger, 2010). La finalidad de este enfoque no es otra que la de poner a disposición de los aprendientes muestras de

lengua auténtica para el aprendizaje de la gramática de manera inductiva, sin reglas y a través de la reflexión (Larsen-Freeman, 1994).

Este acercamiento inductivo estimula, como ya se ha señalado, la habilidad para descubrir patrones en la lengua objeto. Se trata de un proceso que enfrenta al aprendiente, tan directamente como sea posible, con los datos reales de la lengua y hace de él un investigador lingüístico (Johns, 2002), un segundo *Sherlock Holmes*, en palabras de Johns (1997:101). O un viajero, en palabras de Sripicharn, en el sentido de que los aprendientes actúan como viajeros que exploran un corpus para encontrar posibles respuestas a sus búsquedas lingüísticas (Sripicharn, 2010: 372).

Por todo ello, debemos señalar que la inclusión –y explotación- de los corpus en la clase de ENE, no solo realza el papel de la gramática en la adquisición de una segunda lengua. Nos permite, así mismo, acercarnos a su comprensión y uso a través del importante mecanismo cognitivo de la inducción por parte del estudiante. Para ello debemos enfocarles antes la lupa sobre los componentes gramaticales específicos que queramos trabajar en clase. Y lo haremos, como ya se ha señalado, realzando el input al "colocar banderas sobre él" (Sharwood Smith, 1993: 176).

Este realce lo llevaremos a cabo sobre los componentes que la Lingüística de corpus nos haya revelado previamente como fundamentales para nuestra didáctica específica. En nuestro caso concreto, una buena pedagogía de la lengua de los negocios debe incluir áreas como la deixis personal, la modalidad, la protección o la vaguedad (O´Keeffe *et al,* 2007). Este acercamiento inductivo, en el cual el estudiante juega un papel activo en el proceso de aprendizaje, es la esencia del Aprendizaje guiado por datos. Y tiene mucho que ver con las corrientes actuales de la pedagogía, que hacen a los estudiantes más activos, reflexivos y autónomos en su aprendizaje (Chambers, 2010).

3.2. Tipología de ejercicios

Las actividades en las que se concreta lo analizado hasta aquí, pueden localizarse en un rango que va desde las más dirigidas por el profesor hasta las controladas por el propio aprendiente. Las primeras serán adaptaciones a partir del corpus, por lo tanto acceso indirecto a este, y las segundas serán ya consultas personales al corpus, por lo tanto acceso directo (McCullough, 2001; Chambers, 2010). Los docentes pueden usar una técnica de andamiaje en la que los estudiantes vayan progresivamente familiarizándose con las herramientas y la metodología de este enfoque (Granath, 2009). Las maneras de mostrar la información a los estudiantes para su trabajo en clase, pueden ser variadas.

A. En primer lugar, y como punto de partida, para evitar incluir un elemento de distracción en los estudiantes, puede ofrecérseles las líneas de concordancia, pero editadas de forma que contengan frases completas. En ellas, la palabra buscada o *nodo* aparecerá resaltada en negrita o en un color diferente, lo cual les ayudará a buscar el comportamiento de dicha palabra en contexto, guiados en todo momento por el profesor.

B. Otra manera de visualizar la información puede ser idéntica a la que aparece tras la búsqueda en el programa de concordancias, pero pidiendo al estudiante que focalice de manera consciente su investigación. En esta segunda fase se le puede pedir que indague sobre las colocaciones o combinaciones que se repiten con mayor insistencia en torno a la forma realzada. Así, por ejemplo, a través de las líneas de concordancia el aprendiente descubrirá la frecuencia de aparición de la combinación "ascensos/descensos en las bolsas", tal y como se muestra en la Figura 3:



File Edit View Compute Settings Windows Help

N	Concordance	Set	Tag	Word #	Sen	Sen	Para	Para	Hea	Hea	Sec	Sec	File
1	, respectivamente. Por otro lado, Bolsas y Mercados (-0,37%) registra			137	8	16%	0	84%			0	84%	PE2014_0003.x
2	de estrÃ©s a la banca europea, las bolsas cierran con signo mixto pero			25	2	69%	0	10%			0	10%	PE2014_0009.x
3	las alzas registradas en su inicio, las bolsas europeas cerraban en negativo.			82	4	73%	0	40%			0	40%	PE2014_0010.x
4	de cierre de mercados Cautela en las bolsas			11	2	53%	0	5%			0	5%	PE2014_0008.x
5	Pese a la escalada registrada por las bolsas europeas el pasado viernes, al			28	3	28%	0	12%			0	12%	PE2014_0008.x
6	centran la atenciÃ³n en las bolsas europeas. Los mercados			30	2	97%	0	12%			0	12%	PE2014_0012.x
7	crecimiento en EE.UU. impulsa a las bolsas			15	3	17%	0	9%			0	9%	PE2014_0013.x
8	en EE.UU. daban un giro a las bolsas que se recuperaban de las caÃ			52	5	15%	0	31%			0	31%	PE2014_0013.x
9	de mercados Cierre mixto en las bolsas			12	2	37%	0	5%			0	5%	PE2014_0012.x
10	Las bolsas europeas cerraban al alza tras			14	2	33%	0	7%			0	7%	PE2014_0002.x
11	de cierre de mercados Subidas en las bolsas			11	2	37%	0	7%			0	7%	PE2014_0003.x
12	favorables desde Asia y las bolsas europeas no dudaban en			29	2	52%	0	13%			0	13%	PE2014_0001.x
13	ContinÃºan las subidas en las bolsas			13	2	31%	0	6%			0	6%	PE2014_0002.x
14	cierre de mercados Ascensos en las bolsas			11	2	48%	0	7%			0	7%	PE2014_0006.x
15	35 encabezaba las subidas en las bolsas gracias a la banca y a			31	3	34%	0	21%			0	21%	PE2014_0006.x
16	cierre de mercados Descensos en las bolsas			11	2	43%	0	6%			0	6%	PE2014_0004.x
17	del Ibex 35, las compaÃ±Ã-as IAG y Bolsas y Mercados registran las			161	8	63%	0	78%			0	78%	PE2014_0014.x

Figura 3. Búsqueda de las colocaciones en torno a la forma "bolsas"

C. También puede pedírsele al estudiante que observe con atención las formas más alejadas de la palabra clave para llegar, así, inductivamente, a descubrir la combinatoria de los verbos que acompañan frecuentemente al término "bolsas": "cerrar, recuperar, registrar, etc.".

D. El profesor puede seleccionar inicialmente estas líneas de concordancia y agregarles un conjunto de preguntas para guiar al alumno a la hora de advertir la información relevante y despertar así su consciencia sobre los modelos presentes en la muestra seleccionada (Hunston, 2002; Sripicharn, 2010). Cuestiones como las formuladas en la Figura 4 permitirán que el estudiante se familiarice con la metodología y estimularán en él la generación y comprobación de hipótesis.

1. ¿Cuál es la forma que queremos analizar?
2. ¿A qué número pertenece, singular o plural?
3. ¿Qué sustantivos aparecen con mayor frecuencia antes de esta forma?
4. ¿Qué formas verbales se combinan con la forma que estamos analizando? Reúnelas y busca similitudes.
5. ¿…?

Figura 4. Preguntas para provocar la inducción en el estudiante

E. A continuación, con otro modelo de ejercicio como el que refleja la Figura 5, se puede emplear la evidencia negativa –o errores del propio estudiante- para focalizar sobre ellos la búsqueda en el corpus y, por tanto, la consciencia sobre la forma focalizada.

Tarea: las palabras o frases con asterisco contienen errores o usos inapropiados. Busca algunas líneas de concordancia y usa la información obtenida para sugerir una versión revisada de cada frase. Te proponemos las palabras de búsqueda:

Versión original	Palabras de búsqueda	Versión revisada
* "La mejor cantidad de ventas se produce en el primer trimestre"	**"mayor"**	
* "El volumen de ventas ha subido y por eso ahora los datos son mayores"	**"mejor"**	
* "La cualidad es la principal característica de nuestros productos"	**"calidad"**	
* "Nuestros principales calidades son la puntualidad y los buenos precios"	**"cualidad"**	

Figura 5. Trabajo en el corpus con los errores del propio estudiante

F. Por último, y en un estadio superior de adaptación y aprendizaje, se puede conducir al estudiante a que formule por sí mismo cuestiones lingüísticas que le preocupen y que intente darles respuesta mediante la consulta directa a muestras de corpus. Así puede indagar, en función del nivel en el que nos encontremos, diferentes cuestiones, como, por ejemplo, el uso de la determinación con el que se ha encontrado en búsquedas anteriores (el porqué en determinados contextos decimos "las bolsas", "los mercados", "las subidas" … y no *unas bolsas, *unas subidas, *unos mercados…).

Estos son solo algunos ejemplos de cómo los estudiantes, mediante ejercicios prácticos de consulta al corpus, y en un proceso gradual de dificultad, pueden trazar hipótesis, confirmarlas o rechazarlas y llegar a conclusiones basadas en las líneas de concordancia (Granath, 2009). Como hemos podido comprobar, los corpus son una manera de realzar (*enhance*) y focalizar (*focus*) el input para el estudiante (Johansson,

2009), de forma que este lleve a cabo toda clase de procesos cognitivos que estimulen su aprendizaje.

4. Conclusión y discusión

Los beneficios didácticos que se derivarían de la implementación de los corpus con un tratamiento inductivo en la clase de español han recorrido transversalmente esta exposición. Procede ahora su sistematización y la incorporación de otros que aún no se han mencionado, en aras de una mayor claridad conceptual.

A. **Desde la perspectiva del aprendiente**, el alumno adopta bajo este enfoque un papel más activo e independiente, más autónomo (O´Keeffe *et al.* 2007), al dirigir sus búsquedas en el corpus hacia las carencias encontradas en su propia interlengua y controlar, así, su propio aprendizaje. En este sentido, las ideas más interesantes sobre el uso de los corpus en la enseñanza de la lengua tienen que ver con la noción del estudiante como investigador (Johns, 2002). El aprendizaje efectivo de la lengua es, en sí mismo, una forma de investigación lingüística y las concordancias facilitan dicho proceso mediante estrategias de aprendizaje inductivas. En particular, las estrategias para percibir similitudes y diferencias y las de formación y comprobación de hipótesis (Johansson, 2009).

B. **Desde la perspectiva del docente**. El profesor se convierte aquí en un activo diseñador de sus propios materiales que adapta en todo momento al considerar el perfil y el ritmo de aprendizaje de sus estudiantes. Será también él, por tanto, más autónomo (Svalberg, 2016). Además, su papel cambia, pues ya no es él solo la única fuente de conocimiento en la lengua objeto: es un facilitador que manipula el input, ayuda a interpretar los datos y da consejos sobre cómo buscar mejor en el corpus. Todo ello obligará al docente a adoptar un acercamiento descriptivo, más que prescriptivo, a la lengua, todo un reto para él (Chambers, 2010).

C. **Desde la perspectiva del aprendizaje**. El aprendizaje así concebido resulta más motivador, al incluir un elemento de descubrimiento, lo cual redunda en una mayor adquisición (Johansson, 2009). Algunos autores como Svalberg (2016: 403) hablan en este sentido de compromiso con el aprendizaje (*Engagement With Learning, EWL*). Con él destacan el acuerdo tácito en el que el estudiante debe involucrarse: se asume en este sentido que la interacción con otros –compañeros y profesores- facilitará esta construcción consciente del conocimiento lingüístico.

D. **La perspectiva de la metodología**.

   a. Los materiales que se emplean representan muestras de lengua auténtica y no lengua artificial o inventada, con el estímulo que ello supone para los estudiantes (Chambers, 2010). Estamos, además, ante recursos explícitamente empíricos (Dash 2010) que pueden desencadenar el que los creadores de materiales para la

enseñanza del español de los negocios comiencen a incorporar los datos de los corpus lingüísticos en sus textos, como ya están llevando a cabo en el caso del inglés en contextos específicos (O´Keeffe *et al.* 2007).

b. **Las ventajas que entraña la técnica se conforman a medio-largo plazo**. Más allá de la adquisición de un ítem, sobre el que muchos autores reconocen que el impacto es escaso, puede tener efectos a largo plazo sobre el desarrollo de destrezas cognitivas reutilizables en su proceso de aprendizaje (Gilquin & Granger, 2010).

c. **Todas las directrices del enfoque están orientadas a partir de un acercamiento de abajo hacia arriba** (*bottom-up*), es decir, extraído de la propia realidad de la clase para, desde ella, llegar a las conclusiones pedagógicas pertinentes: entre ellas, que el lenguaje observado es un objeto de estudio más válido que la intuición; que es muy importante la frecuencia de aparición de un ítem lingüístico y que un corpus puede producir resultados que pueden extrapolarse a categorías más generales de la lengua (Hunston, 2002).

d. Bajo esta perspectiva, la **creatividad** se introduce también en la clase de español mediante el uso de los corpus lingüísticos. El objetivo es desarrollar en los aprendientes la consciencia de las propiedades y funciones de los patrones lingüísticos que trabajan creativamente en la comunicación diaria (O´Keeffe *et al.* 2007), de manera que ellos mismos puedan experimentar esta dimensión creativa de la lengua.

e. **Dimensión creativa y constructora de la lengua**. Bajo esta nueva perspectiva, las reglas gramaticales, así como los significados de las palabras, no solo se aprenden sino que se construyen. En la sociedad del conocimiento en la que estamos, el aprendiente ha de situarse como activo constructor del conocimiento (Dash 2010). Y somos nosotros, como docentes, los que podemos ayudarles en la búsqueda de este perfil.

La Lingüística de corpus es una metodología que está en plena ebullición (Parodi 2008) y que reporta importantes ventajas como las analizadas aquí. Dichas ventajas nos incitan a incluirlo como un complemento en la clase del español de los negocios.

Es cierto también que muchos detractores invocan obstáculos importantes como, por ejemplo, el tiempo que le lleva al profesor preparar los materiales y para el estudiante familiarizarse con la técnica. Cabe señalar, en este sentido, que hoy en día este software ha alcanzado tal grado de sofisticación que su uso es sumamente sencillo, de forma que en muchos países su acceso a él dentro del aula se ha convertido ya en una rutina (Dash, 2010). Los propios programas incorporan plantillas listas para ser usadas por el docente con el objetivo de crear ejercicios de factura similar a los aquí presentados.

Otra desventaja invocada es el hecho de que los profesores pueden no tener la formación suficiente para manejar los programas de búsqueda (Frankenberg-Garcia, 2016), o el hecho de el estudiante pueda realizar inferencias equivocadas (Gilquin & Granger, 2010). En el primer caso, por supuesto el docente debe conocer la técnica pero esta no es complicada en absoluto. En el segundo caso, si el estudiante llega a reflexiones no acertadas, siempre estará el profesor como facilitador o el propio input como recurso de contraste y verificación.

Pese a ello, estamos ante un enfoque prometedor que no debería abandonarse (Chambers, 2010). Entre otras razones, porque lleva a la clase una lengua auténtica, motiva a los estudiantes por el elemento de descubrimiento que encierra, desarrolla importantes destrezas cognitivas como las analizadas líneas más arriba y les proporciona beneficios que van más allá del conocimiento de un ítem.

Sin embargo, es preciso admitir que existe todavía una brecha entre la Lingüística de corpus y la Didáctica de la lengua (McCarthy & O´Keeffe, 2010) y que serían deseables estudios empíricos sobre la efectividad de los corpus en la enseñanza de la lengua (Johansson, 2009). Debemos reclamar prudencia en este sentido. También en el hecho de que el uso de los corpus con finalidad didáctica son un apoyo, pues ni sustituyen a la comunicación –aunque sobre ellos se pueda trabajar comunicativamente- ni pueden reemplazar al profesor (Johansson, 2009; Svalberg, 2016).

Su inclusión en la clase de español, sin embargo, fortalecerá los procesos naturales de adquisición de la lengua en la medida en que hará más sencillo a los aprendientes advertir y experimentar secuencias repetidas de uso (Johansson, 2009).

No debemos olvidar, por último, que las mejores condiciones para el aprendizaje de una lengua tienen lugar cuando se produce un acercamiento combinado, que incluya lo implícito y lo explícito, lo inductivo y lo deductivo en el propio estudiante.

Notas

[1] Para Schmidt (2011:13) no hay duda de que el aprendizaje con atención es muy superior.

[1] El Corpus Comillas del Español de los Negocios (CORPEN Comillas) es un corpus específico que tiene como fin servir para obtener las características del español de los negocios en España y América en el periodo comprendido entre los años 2010 y 2020. Contará con 12 millones de formas y lo está llevando a cabo un equipo de investigación del Centro Internacional de Estudios Superiores del Español (CIESE-Comillas).

Bibliografía

Chambers, A. (2010). What is data-driven learning? En A. O´Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 345-358). Nueva York: Routledge.

Dash, Niladri Sekhar (2010) "Corpus Based English Language Teaching: A New Method". *Proceedings of the Orientation-Cum-Workshop Programme on Recent Trends in Language Teaching*, 4-6 June 2010, Institute of Advanced Studies in Education, David Hare Training College, Kolkata, pp. 19-36.

Evison, J. (2010). What are the basis of analysing a corpus? En A. O´Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 345-358). Nueva York: Routledge.

Fotos, S. (1994). Integrating Grammar Instruction and Communicative Language Use Through Grammar Consciousness-Raising Tasks. *TESOL Quarterly* 28 (2), 323-351.

Frankenberg-Garcia, A. (2016). Corpora en ELT. En G. Hall (Ed.) *The Routledge Handbook of English Language Teaching* (pp. 383-398). Nueva York: Routledge.

Gilquin, G. & Granger, S. (2010). How can data-driven learning be used in language teaching? En A. O´Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 359-370). Nueva York: Routledge.

Granath, S. (2009). Who benefits from learning how to use corpora? En K. Aijmer (Ed.), *Corpora and Language Teaching* (pp.47-65). Amsterdam: John Benjamins Publishing Company.

Hall, G. (2011). *Exploring English Language Teaching. Language in Action*. Nueva York: Routledge.

Hunston, S. (2002), *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Johansson, S. (2009). Some thoughts on corpora and second language acquisition. En K. Aijmer (Ed.), *Corpora and Language Teaching* (pp.33-44). Amsterdam: John Benjamins Publishing Company.

Johns, T. (1997). Contexts: The background, development and trialling of a Concordance-based CALL Program. En A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles, (Eds.), *Teaching and language corpora* (pp. 100-115). London: Longman.

Johns, T. (2002). Data-driven learning: the perpetual challenge. En B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Linguistics* (pp.107-117). Amsterdam: Rodopi.

Larsen-Freeman, D. & Long, M. H. (1994). *Introducción al Estudio de la Adquisición de Segundas Lenguas*. Madrid: Gredos.

Martínez, I. (2004). *Nuevas perspectivas en la Enseñanza-Aprendizaje de ELE para japoneses: la concienciación formal*. Tesis doctoral [en línea]. Disponible en: http://eprints.ucm.es/4338/. Madrid: Universidad Complutense de Madrid, Servicio de Publicaciones.

McCarthy, M. & O´Keeffe, A. (2010). Historical perspective: what are corpora and how have they evolved? En A. O´Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 359-370). Nueva York: Routledge.

McCullough, J. L. (2001). Los usos de los córpora de textos en la enseñanza de lenguas.En M. Trenchs Parera (Ed.), *Nuevas tecnologías para el autoaprendizaje y la didáctica de lenguas* (pp. 125-140). Barcelona: Milenio.

O´Keeffe, A., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom: Language use and Language teaching*. Cambridge: Cambridge University Press.

Parodi, G. (2008). "Lingüística de corpus: una introducción al ámbito". *RLA. Revista de lingüística teórica y aplicada*, *46* (1), 93-119. https://dx.doi.org/10.4067/S0718-48832008000100006 [Acceso 16-10-2016]

Schmidt, R. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics*, 11 (2), 127-158.

Schmidt, R. (2011). La atención, en Robinson, P. (Ed.). *La dimensión cognitiva en la enseñanza de lenguas*. Madrid: Edinumen.

Schmidt, R. & Frota, S. (1986). Developing basic conversational ability in a second language: a case study of an adult learner. En Day, R. (Ed.), *Talking to Learn: Conversation in Second Language Acquisition*. Rowley, Mass.: Newbury House.

Sharwood Smith, M. (1993). Input Enhancement in Instructed SLA. *Theoretical Bases. Studies in Second Language Acquisition* 15, 165-179.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.

Sripicharn, P. (2010). How can we prepare learners for using language corpora? En A. O´Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 371-384). Nueva York: Routledge.

Svalberg, A. M-L (2016). Language Awareness. En G. Hall (Ed.), *The Routledge Handbook of English Language Teaching* (pp. 399-412) Nueva York: Routledge.

Tsui, A.B.M. (2005). ESL Teacher´s Questions and Corpus Evidence. *International Journal of Corpus Linguistics*,10 (3), 335-356.

Willis, J. & Willis, D. (1996). Counsciousness-raising activities in the classroom. En J. Willis & D. Willis (Eds.), *Challenge and change in Language Teaching*. London: Heinemann.