

LA LINGÜÍSTICA DE CORPUS AL SERVICIO DE LA ENSEÑANZA DE UNA LENGUA EXTRANJERA: UN ESTUDIO COMPARADO DE RESEÑAS CRÍTICAS

Área Temática:

- Norma y variaciones en la enseñanza/aprendizaje de lenguas extranjeras y segundas. Lugar de la fonética, la gramática y el léxico
- Didáctica del nivel superior

Magdalena Zinkgräf
maguizinkgraf@gmail.com
Universidad Nacional del Comahue

Introducción

La lingüística de corpus ha sido empleada como herramienta al servicio de disciplinas relacionadas con estudios del lenguaje como la enseñanza de un idioma extranjero (Granger y Meunier, 2009: xix), fortaleciéndolas con investigaciones en la fraseología y el análisis del discurso. La fraseología caracteriza, a través de estudios computacionales, el lenguaje empleado en general, y/o en circunstancias particulares definidas a través de un género discursivo (Henry y Roseberry 2001, Ellis 2008, y Meunier y Granger 2008). Sus hallazgos difuminan las divisiones clásicas entre léxico y gramática, favoreciendo la identificación de unidades pluriverbales (Sinclair 2009) que sistemáticamente recurren en textos de diversa índole.

Este trabajo se centra en el papel dual que cumple la lingüística de corpus al a) detectar frases en un corpus de reseñas por hablantes nativos del inglés para un proyecto de escritura y b) facilitar la comparación entre dicho corpus y su contrapartida por estudiantes universitarios de inglés para el Profesorado y Traductorado y así evaluar la efectividad del proyecto.

1. El contexto educativo y la lingüística de corpus

El objetivo general de la asignatura Inglés V del Profesorado y Traductorado en Inglés es lograr que los alumnos alcancen un nivel semejante al de un hablante nativo en aspectos relacionados con el uso del idioma extranjero en los modos oral y escrito, y monitoreen su producción para corregir los errores típicos del output de estudiantes de una L2.

En los cursos 2006-2008, se propuso la escritura de reseñas críticas de micro-relatos enmarcada en el “enfoque del género” (*The genre approach*) (Swales 1990; Lin 2006; Rabbini 2003). Las reseñas permiten la integración de habilidades para la redacción de un texto original y auténtico, con un propósito comunicativo genuino: reseñar (para su posible publicación) micro-relatos de la revista *Smokelong Quarterly*.

2. El Corpus de Expertos Nativos como modelo

Los estudiantes fueron expuestos a un corpus de reseñas críticas escritas por tres o más periodistas reconocidos del diario británico *The Guardian* y del *Observer*. Este corpus se presentaba paulatina y sistemáticamente durante el año a través de diferentes tipos de actividades que los alumnos debían completar. Además, los textos eran analizados en clase para desarrollar una conciencia crítica sobre los recursos lingüísticos empleados por los autores nativos, focalizando estructuras y aspectos gramaticales (uso de tiempos verbales y conectores), giros idiomáticos y colocaciones.

Así se fueron sumergiendo en un “ambiente” donde el género les proporcionaba el input necesario para la adquisición implícita de la lengua extranjera propuesta en Ellis (2008), Kennedy (2008) y Coxhead (2008). Hacia mediados de año, los estudiantes comenzaron con la escritura de sus propias reseñas, que fueron recopiladas en el Corpus de Estudiantes (CEs) para su comparación con el Corpus de reseñas escritas por Expertos Nativos (CEN).

3. Comparación de los corpus

3.1. Estadísticas generales y palabras frecuentes

A través del software *Wordsmith Tools 5.0*, a) se analizó tanto el CEN, de 50 reseñas de cine y ficción, identificando aspectos que caracterizan el género, como el CEs, de 54 reseñas críticas escritas por alumnos avanzados de inglés y b) se compararon estos corpus con respecto a 1) las palabras más frecuentes, 2) las secuencias de tres palabras con mayor repetición y 3) las líneas de concordancia que caracterizan el uso de estas secuencias.

	CEN	CEs
Número de textos	50	54
Número de palabras (<i>tokens</i>)	24.465	17.938
Tipos de palabras (<i>types</i>)	6.747	3.407
TTR (coeficiente <i>tokens/types</i>)	26,46	18,99
Longitud media de palabra	4,75	4.95
Desviación estándar	+/- 2.75	+/- 2.58

Tabla 1: Estadísticas generales de los micro-corpus

En la tabla 1 las estadísticas generales de ambos corpus muestran una importante diferencia entre el número de palabras distintas (*types*) de cada corpus. Mientras que el CEN presenta una gran variedad léxica según el valor del coeficiente TTR, el CEs es

significativamente menor. Así, el léxico empleado por los hablantes nativos en sus textos es, comprensiblemente, más variado que el de los alumnos, quienes se apoyan en mayor medida en la repetición de las palabras que dominan.

Otra medición interesante es la longitud media de palabra. Scott y Tribble (2006) demuestran que, a mayor largo de palabra, mayor complejidad del vocabulario (Hunston 2002). La tabla 1 indica que ambos corpus son semejantes en este respecto (nótese la longitud media y su desviación estándar).

La comparación entre las primeras veinte palabras frecuentes revela una importante correlación entre las bases de datos. Con cierta variación en el orden de los distintos lexemas en las listas, entre 75 y 80% las palabras de clase cerrada (preposiciones, conjunciones, artículos y como verbos auxiliares) se emplean en medidas similares para ambos corpus. Se registran diferencias en los vocablos HER, STORY, NOT y BE, ausentes en las 20 palabras más frecuentes del CEN. Esto responde a una característica intrínseca al corpus: mientras que el CEN se compone de reseñas de películas y novelas de ficción, el de los estudiantes contiene reseñas de micro-relatos o *short short-stories*.

BE también aparece más frecuentemente en el CEs en compañía de verbos cópula (atenuadores) y verbos modales que indican el grado de certeza con que se asevera una proposición (Gráficos 1 y 2 respectivamente).

Esta importante disparidad con el CEN puede tener sus orígenes en su condición de críticos “amateurs.” Estos alumnos universitarios optan por mitigar sus opiniones al emitir juicio y seleccionan una gran proporción de construcciones de elevación para ello.

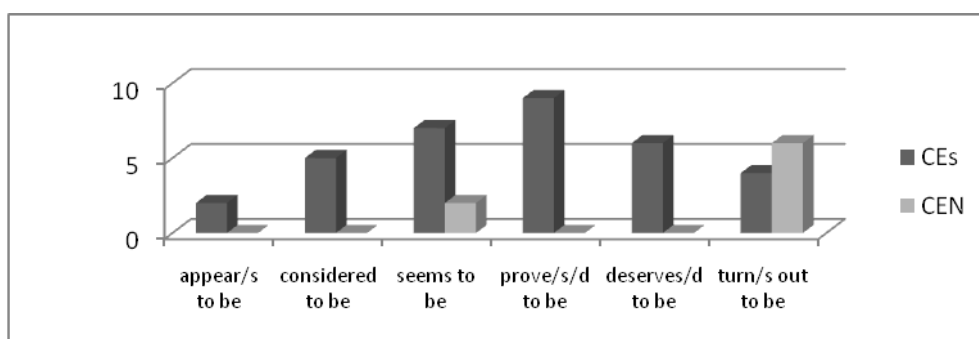


Gráfico 1. Distribución de verbos cópula + BE en los corpus

En el Gráfico 2 la distribución de los verbos modales SHOULD, WOULD y MIGHT + BE en el CEs muestra similitudes con el CEN, pero un sobreuso de CAN, WILL y MAY. Este exceso podría explicarse por la relativa frecuencia general y el amplio espectro de uso de estos tres modales en el inglés. A raíz de su “facilidad,” los estudiantes se han apoyado en estos modales básicos. Otra alternativa es que los aprendientes no hayan captado la característica estilística del género.

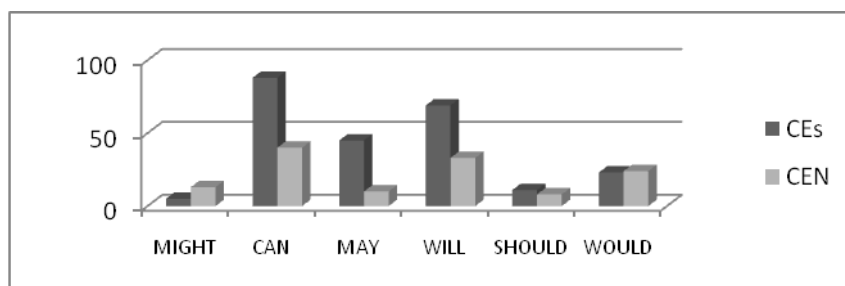


Gráfico 2. Distribución de verbos modales acompañados de BE en los corpus

En general, cabe recalcar que, en su proceso de aprendizaje, los estudiantes fueron advertidos sobre su dependencia atípica de algunas palabras, y fueron invitados durante el curso a evitar las “palabras prohibidas,” reemplazándolas por giros idiomáticos más frecuentes en el CEN. Gradualmente, se desprendieron del sobre-uso y la repetición de vocablos como STORY, READER, READ, AUTHOR y BE, si bien el CEs no alcanza a reflejar este proceso de disminución¹.

3.2. Frases pluriverbales

Al contrastar las listas de frases pluriverbales compuestas de 3 a 5 palabras², se observa que sólo algunas son comunes a ambos corpus, si bien en distintas proporciones (Gráfico 3). La mayor diferencia se registra en AS WELL AS, que es mucho más frecuente en el CEs (0,34% vs. 0,10%), y en ONE OF THE, cuya recurrencia se destaca en el CEN. La función coordinante de la primer secuencia absorbe la proporción que cumple AND en el corpus experto, 3,22%, en contraste con el CEs (2,17%), disparidad inducida también por la técnica de las “palabras prohibidas”.

¹ Los textos que componen el CEs corresponden a distintos momentos durante el curso, incluyendo los primeros intentos.

² Mínimo de frecuencia de aparición de 3 en cada corpus.

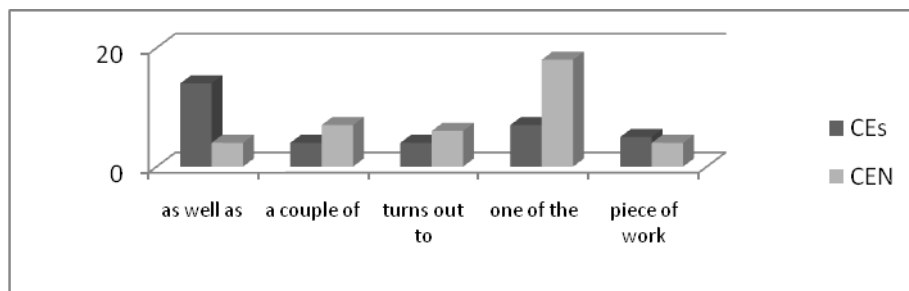


Gráfico 3. Distribución de frases pluriverbales comunes a ambos corpus

En cuanto a ONE OF THE, descubrimos en las líneas de concordancia arrojadas por el programa (*concordance lines*) que, si bien existe la misma cantidad de usos de esta frase con una forma superlativa (ONE OF THE *MOST* + *ADJ*) en ambos corpus, en el CEN se registran 10 usos más sin superlativo.

4 ng? Are Mr and Mrs Youngling going to come home to find a nursery bloodbath? **One of the things about** the previous film, Attack of the Clones, that made you t
 5 t is triumphant English history. Lisa Jardine is **one of the editors as well as one of the authors** of this 'Making History' series, which aims to take great tur

Otra divergencia se observa en frases que aparecen en ambos corpus con adjetivos de por medio en el CEs. A SENSE OF, A KIND OF, THE WORLD OF son ejemplos de discrepancias en el corpus inducidas pedagógicamente³. Sin embargo, la exclusividad de las frases THERE IS SOMETHING, IN A WAY, A SERIES OF, AT ANY RATE, TO SAY THAT en el CEN indica el nivel de idiomática auténtica de este corpus. A modo de ejemplo, compárese las siguientes líneas de concordancia para IN A WAY, empleada como introducción para una cláusula de resultado encabezada por el subordinante THAT en el CEN:

1 book club and of course pedantically relate the novels to their own love-lives, **in a way that's** supposed to be vulnerable and gutsy and life-affirming. The shee
 2 about whom pundits of all ages and sexes tend to get precious and proprietorial **in a way [that]** they don't about, say, Charlotte Brontë or George Eliot. I can only say

Por su parte, los alumnos emplearon la frase con el significado “de una manera (adj)”, como ilustran los siguientes ejemplos:

³ A partir de la consigna de las docentes, los estudiantes incluyeron un adjetivo por cada sustantivo que utilizaban en sus reseñas, y las frases típicas del CEN aparecen interrumpidas por un adjetivo en el CEs.

6 please the leading role feels because of this are dealt with **in an interesting way**. Quite a satisfying short story may grip the reader and lead him to reflect

16 other set stories dealing with unconventional topics **in a brilliant and clever way**. In just one page, Smith makes rich use of language, including only the nece

Esta diferencia es indicativa de la competencia fraseológica “en desarrollo” evidenciada en el Corpus de Estudiantes.

4. Conclusión

En este trabajo se han puesto las herramientas de la lingüística de corpus al servicio de un proyecto de escritura de reseñas críticas a) como instrumento informante sobre frases y estructuras características del género propuesto y b) como método de evaluación del proyecto, buscando similitudes y diferencias entre aprendientes y expertos, para así señalar áreas de cambio y subsanar dificultades.

La comparación revela cómo la producción escrita dentro del género “reseña crítica” de este grupo de estudiantes universitarios en un contexto de inglés como lengua extranjera se asemeja a la observada en el corpus “experto” en frecuencia de palabras y secuencias fraseológicas. Existen diferencias generadas por artificios pedagógicos y otras que caracterizan el CEs como en vías en desarrollo.

Bibliografía

- Coxhead, Avril (2008): “Phraseology and English for academic purposes: Challenges and opportunities.” En Meunier, F. y S. Granger (eds.) *Phraseology in Foreign Language Learning and Teaching*, Amsterdam: John Benjamins Publishing Company, pp. 149-152.
- Ellis, Nick (2008): “Phraseology: The periphery and the heart of language”. En Meunier, F. y S. Granger (eds.) *Phraseology in Foreign Language Learning and Teaching*, Amsterdam: John Benjamins Publishing Company, pp. 1-14.
- Granger, Sylvianne y Fanny Meunier (eds.) (2009): “Introduction: The many faces of phraseology” En *Phraseology: An interdisciplinary perspective*. Amsterdam: John Benjamins Publishing Company, pp. xix-xxviii.
- Henry, Alex y Robert L. Roseberry (2001): “Using a small corpus to obtain data for teaching a genre”. En Ghadessy, Mohsen, Alex Henry y Robert L. Roseberry (eds): *Small Corpus Studies and ELT: Theory and Practice*. Amsterdam: John Benjamins Publishing Company, pp. 93-134.
- Hunston, Susan (2002): *Corpora in Applied Linguistics*, Cambridge: Cambridge University Press.

- Kennedy, Graeme (2008): "Phraseology and language pedagogy: Semantic preference associated with English verbs in the British National Corpus." En Meunier y Granger (eds.) *Phraseology in Foreign Language Learning and Teaching*, Amsterdam: John Benjamins Publishing Company, pp. 21-42.
- Sinclair, John (2009): "The phrase, the whole phrase, and nothing but the phrase". En Granger, Sylvianne y Fanny Meunier (eds.) *Phraseology: An interdisciplinary perspective*. Amsterdam: John Benjamins Publishing Company Swales (1990), pp. 407-410.
- Lin, Benedict (2006): "Genre-based Teaching and Vygotskian Principles in EFL: The Case of a University Writing Course". *Asian EFL Journal*, 8 (3) en http://www.asian-efl-journal.com/Sept_06_bl.php.
- Meunier, Fanny y Sylvianne Granger (eds.) (2008): *Phraseology in Foreign Language Learning and Teaching*, Amsterdam: John Benjamins Publishing Company.
- Rabbini, R. (2003): "The genre approach to writing assessed", *Academic Exchange Quarterly*, March obtenido en <http://www.thefreelibrary.com/The+genre+approach+to+writing+assessed.-a0101943708>
- Scott, Mike and Christopher Tribble (2006): *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.
- Swales, John (1990): *Genre analysis*. Cambridge: Cambridge University Press.
- *Smokelong Quarterly* en <http://smokelong.com/>