

Douglas Biber

Paquetes léxicos en textos de estudio universitario: Variación entre disciplinas académicas
Revista Signos, vol. 38, núm. 57, 2005, pp. 19-29,
Pontificia Universidad Católica de Valparaíso
Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=157013764002>



Revista Signos,
ISSN (Versión impresa): 0035-0451
revista.signos@ucv.cl
Pontificia Universidad Católica de Valparaíso
Chile

¿Cómo citar?

Fascículo completo

Más información del artículo

Página de la revista

www.redalyc.org

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Paquetes léxicos en textos de estudio universitario: Variación entre disciplinas académicas*

Douglas Biber

*Universidad del Norte de Arizona
Estados Unidos de Norteamérica*

Resumen: Durante estas dos últimas décadas los investigadores han estado crecientemente interesados en el estudio de expresiones prefabricadas de varias palabras. Un acercamiento productivo a esta área de pesquisa ha sido investigar las secuencias frásticas que realmente ocurren con mayor frecuencia en un registro dado, las que se denominan paquetes léxicos (*lexical bundles*). Los paquetes léxicos tienen ciertas características estructurales sistemáticas y sirven para importantes funciones discursivas: al expresar posturas, señalar la organización discursiva o proveer marcos referenciales. Estudios anteriores han documentado el uso de paquetes léxicos en conversaciones, prosa académica, libros de textos universitarios, escritos de alumnos y en la enseñanza en el aula. El presente estudio extiende esta línea de investigación, comparando el uso de paquetes léxicos en libros de textos de cinco principales disciplinas académicas: negocios, ingeniería, ciencias naturales, ciencias sociales y humanidades. Lo encontrado aquí muestra que existen importantes diferencias entre estas disciplinas, tanto en el uso que hacen de estos paquetes, como en las funciones que estos cumplen.

Palabras Clave: Paquetes léxico, expresiones frásticas, registro, libros de textos, disciplinas académicas.

Lexical bundles in university textbooks: Variation across academic disciplines

Abstract: Over the past two decades, researchers have become increasingly interested in the study of multi-word prefabricated expressions. One productive approach to this research area has been to investigate the multi-word sequences that actually occur most commonly in a given register, referred to as 'lexical bundles'. Lexical bundles have systematic structural characteristics, and they serve important discourse functions, expressing stance, signaling the discourse organization, or providing referential frames. Previous studies have documented the use of lexical bundles in conversation, academic prose, university textbooks, student writing, and classroom teaching. The present study extends this line of research by comparing the use of lexical bundles in textbooks across five major academic disciplines: business, engineering, natural science, social science, and humanities. The findings presented here show that there are important differences across academic disciplines, both in the extent to which they rely on lexical bundles and in the functions typically served by those bundles.

Key Words: Lexical bundle, multi-word expressions, register, textbooks, academic disciplines, stance, discourse organization.

Recibido: 15 de noviembre de 2004

Aceptado: 10 de marzo de 2005

Correspondencia: Douglas Biber (douglas.biber@nau.edu). Department of English, Northern Arizona University, PO Box 6032, Flagstaff, AZ 86011-6032, Estados Unidos de Norteamérica.

* Traducido por la Dra. Marianne Peronard.

INTRODUCCIÓN

A lo largo de estas dos últimas décadas, los investigadores han mostrado un interés creciente en el estudio de expresiones frásticas prefabricadas (Weinert, 1995; Ellis, 1996; Howarth, 1996, 1998; Wray & Perkins, 2000; Gras, 2002). Las secuencias frásticas han sido estudiadas bajo diversas denominaciones, incluyendo "frases léxicas", "fórmulas", "rutinas", "expresiones fijas", "patrones prefabricados" (*prefabs*) y, en castellano, "locuciones". Estos enfoques definen su objeto de estudio en términos un tanto diferentes, de modo que ofrecen diferentes perspectivas acerca del uso de estas secuencias de múltiples palabras. Por ejemplo, algunos estudios describen secuencias de palabras que son idiomáticas (idiotismos) como *a ojos vista*, mientras que otros estudios se centran en secuencias frásticas que se ciñen a las reglas de la lengua, pero que se destacan perceptualmente (e.g. *si sabes a qué me refiero*).

Un enfoque complementario consiste en describir las secuencias frásticas que de hecho aparecen frecuentemente en un registro dado, denominadas "locuciones" (*en caso de*). Se han investigado secuencias frásticas en varios estudios, incluyendo a Salem (1987), Altenberg (1993, 1998), Butler (1997) y De Cock (1998). El término "paquete léxico" (*lexical bundle*) fue utilizado por primera vez por Biber, Johansson, Leech, Conrad y Finnegan (1999) en el *Longman Grammar of Spoken and Written English* (LGSWE), en el cual se describían las secuencias recurrentes de palabras en conversaciones y en textos académicos. Este marco teórico ha sido aplicado en diversos estudios subsecuentes (Biber & Conrad, 1999; Partington & Morley, 2002; Cortés, 2004; Biber, Conrad & Cortés, 2003).

Biber, Conrad y Cortés (2004) extienden estos primeros estudios en dos direcciones: 1) describen el uso de paquetes léxicos en dos registros universitarios –enseñanza académica y textos de estudios– comparando esos patrones de uso con los de conversaciones y prosa académica; y 2) desarrollan una taxonomía funcional para una clasificación de las locuciones. Se proponen tres categorías principales: locuciones de postura (*stance*), locuciones organizadoras del discurso y locuciones referenciales. Además, se distinguieron varias sub-categorías que reflejan las particulares funciones de las locuciones en el discurso (e.g. locuciones de postura epistémica versus locuciones de postura actitudinal/modalidad).

La presente investigación aplica esta taxonomía funcional al uso de las locuciones en textos de estu-

dio en cinco disciplinas académicas principales: negocios, ingeniería, ciencias naturales, ciencias sociales y humanidades. Biber, Conrad y Cortés (2004) muestran que las locuciones no son muy comunes en los textos de estudio (en comparación con otros registros). Sin embargo, los resultados que presentamos aquí, muestran que hay importantes diferencias entre las disciplinas académicas, tanto en el grado en que se apoyan en las locuciones, como en las funciones que estas, característicamente, cumplen.

METODOLOGÍA

1. El corpus usado en el estudio

Varios estudios en torno al lenguaje universitario se han basado en el análisis del Corpus T2K-SWAL (TOEFEL, 2000 y Spoken and Written Academia Language Corpus; véase Biber, Conrad, Reppen, Byrd & Helt, 2002; Biber, Conrad, Reppen, Byrd, Helt, Clark, Cortés, Csomay & Urzúa, 2004). Este corpus fue construido para representar la gama de registros orales y escritos que deben abordar los estudiantes universitarios en los Estados Unidos. El corpus completo incluye, pues, tanto registros académicos (e.g. enseñanza en clases y libros de estudios) y registros no académicos (e.g. discursos en torno a aspectos administrativos de las clases, encuentros de servicios, *syllabus* de los cursos, escritos institucionales).

La presente investigación de los textos de estudios se basa en el sub-corpus mostrado en el Cuadro Nº 1. Las muestras para cada disciplina académica son relativamente pequeñas para el propósito del análisis de locuciones (ver 2) y el resultado debiera ser considerado como exploratorio, ya que requiere la confirmación basada en el análisis de un corpus mayor.

2. Identificación y análisis funcional de los paquetes léxicos

Los paquetes léxicos son simplemente secuencias de palabras que ocurren con mayor frecuencia en un corpus de textos. Biber et al. (1999) y Biber y Conrad (1999) documentan el uso de estos paquetes léxicos en conversaciones y en prosa académica; Biber, Conrad y Cortés (2004), además, describen el uso de paquetes léxicos en la enseñanza en clases y en textos de estudio.

Debido a que los paquetes léxicos se definen con base en frecuencias, no son necesariamente unida-

des estructurales completas o expresiones fijas. Por ejemplo, una secuencia de cuatro palabras como *que les parecería que* o *como si esto fuera poco*, son paquetes léxicos comunes en la conversación. Con el fin de delimitar el alcance de esta investigación aquí solo consideramos secuencias de cuatro palabras.

En general, la mayoría de los paquetes léxicos no son idiotismos. Es decir, la mayoría de ellos son demasiado poco frecuentes como para ser considerados paquetes. Para el presente análisis, los paquetes léxicos se definen como secuencias frásticas que ocurren por lo menos veinte veces por millón de palabras, aunque muchos de ellos son muy frecuentes (ocurren más de 200 veces por millón de palabra). En cambio, idiotismos como *estirar la pata* (con el significado de morir) o *lo sentí como una bofetada* (para significar sentirse ofendido), rara vez se encuentran en el habla o la escritura natural. Los idiotismos y las

fórmulas fijas se utilizan generalmente en textos literarios más que en conversaciones cara a cara. Estos ejemplos: “estirar la pata” (*kick the bucket*) y “una bofetada en la cara” (*a slap in the face*) ocurren alrededor de 5 veces por millón de palabras en textos literarios y rara vez han sido encontradas en una conversación real (Biber & Conrad, 1999).

La mayoría de los paquetes léxicos no representan una unidad estructural completa. Puesto que la mayoría une dos unidades estructurales: estos comienzan en el límite de una cláusula o frase, siendo las últimas palabras del paquete los primeros elementos de una segunda unidad estructural. En las conversaciones, la mayoría de los paquetes léxicos sirven de puente entre dos cláusulas (e.g. *quiero saber, bueno eso era lo que yo*), mientras que en la prosa académica suelen enlazar dos frases (e.g. *en caso de, la base de*).

El análisis cuantitativo de los paquetes léxicos se desarrolló con programas computacionales desarrollados en la Universidad del Norte de Arizona. Al identificar los paquetes léxicos nos atuvimos a las palabras en cuanto unidades ortográficas, aun cuando a veces estas combinan palabras separadas en forma bastante arbitraria. Por ejemplo: *autocontrol, cortauñas, milhojas*, y *marcapasos* serían consideradas como una sola palabra en nuestro análisis. Se eliminaron los paquetes léxicos que abarcaban un cambio de turno en la conversación o un signo de puntuación.

En comparación con estudios previos, basados en grandes *copora*, el análisis del cual informamos aquí se basa en muestras relativamente pequeñas y dispares de las diversas disciplinas consideradas (ver Cuadro N°1). En un extremo, los textos de ingeniería están representados solo por 72.000 palabras y 9 textos. En el otro extremo, los textos de ciencias sociales están representados por 213.000 palabras y 21 textos.

Lo encontrado por Biber, Conrad y Cortés (2004) muestra que los paquetes léxicos son generalmente escasos en textos de estudio, comparado con otros registros universitarios. Por ello, este estudio usa un límite de corte relativamente bajo de frecuencia, incluyendo todas las secuencias que ocurren más de 20 veces por millón de palabras para considerarlas paquetes léxicos, lo que corresponde a un mínimo de 4 ocurrencias.

Otra característica definitoria es que la secuencia frástica debe ser usada en diferentes textos para ser considerada un paquete; esta restricción evita los usos

Cuadro N° 1
Composición de los sub-corpus de los libros de estudio usados en el Corpus T2K-SWAL.

Categoría	N° de textos	N° de palabras
Negocios		
Primer año	4	29.700
Segundo año	4	28.400
Graduado	7	58.100
Total	15	116.200
Ingeniería		
Primer año	3	18.600
Segundo año	3	24.900
Graduado	3	28.500
Total	9	72.000
Ciencias Naturales		
Primer y segundo año	6	53.600
Tercero y cuarto año	6	42.600
Graduados	6	49.000
Total	18	145.200
Ciencias Sociales		
Primer y segundo año	7	75.300
Tercer y cuarto año	7	71.200
Graduados	7	66.500
Total	21	213.000
Humanidades		
Primer y segundo año	6	56.300
Tercer y cuarto año	6	52.900
Graduados	6	54.900
Total	18	164.100

idiosincráticos por hablantes o autores individuales. Las muestras individuales de los textos de estudios en el corpus T2K-SWAL tienden a ser largas (generalmente más de 5.000 palabras por texto), lo que da como resultado muestras que tienden a ser de un número relativamente reducido de textos en las distintas disciplinas. Para compensar esto, se siguieron los siguientes criterios:

- Cualquier secuencia que ocurra solo en 4 ocasiones debería estar distribuida al menos en 3 textos diferentes.
- Cualquier secuencia que ocurra 5 o más veces debe estar distribuida al menos en 2 textos diferentes.
- Debido a que los textos de ingeniería están representados por una muestra mucho menor que las otras disciplinas, las secuencias que ocurran solo 3 veces en este registro, tendrán una proposición normada de 40 por millón de palabras. Es decir, límite de corte de frecuencia más bajo que para el resto de los registros.
- Cualquier secuencia que solo ocurra 3 veces en los libros de ingeniería debe estar distribuida en 3 textos.
- Cualquier secuencia que ocurra 4 o más veces en textos de ingeniería debe estar distribuido a lo menos en 2 textos diferentes.

3. Clasificación funcional de los paquetes léxicos

Biber, Conrad y Cortés (2004) desarrollaron una taxonomía funcional de los paquetes léxicos, distinguiendo entre tres funciones primarias: 1) expresiones de posturas; 2) organizadores del discurso; y 3) expresiones referenciales. Los paquetes posturales expresan actitudes o evaluación de certeza que enmarcan otras proposiciones. Los organizadores discursivos reflejan relaciones entre lo anterior del discurso y el discurso posterior. Los paquetes referenciales denotan directamente entidades físicas o abstractas; o al mismo contexto textual, ya sea para identificar una entidad o para distinguir un atributo particular de la entidad como especialmente importante.

La taxonomía fue desarrollada inductivamente. Es decir, agrupamos paquetes que servían para la misma función, basados en el significado y uso típico de cada uno. Usamos listas de concordancia para examinar el uso de cada paquete en su contexto discursivo. Una vez que los paquetes fueron asignados a los grupos, tratamos de determinar las funcio-

nes discursivas asociadas a cada uno de los grupos. Obviamente, este último paso estuvo influido por los estudios teóricos previos acerca de las funciones discursivas de los rasgos lingüísticos (Hymes, 1974; Halliday, 1978; Brown & Fraser, 1979; Biber, 1988).

En algunos casos, un mismo paquete tiene múltiples funciones, incluso al interior de una misma ocurrencia. Por ejemplo, paquetes como *echa una mirada* o *¿te parece que miremos?* funcionan como directivos e introductor de un tópico. En otros casos, un mismo paquete sirve para diversas funciones dependiendo del contexto. Por ejemplo, paquetes como *el comienzo de* y *al final de* pueden funcionar como referencia temporal o espacial, o como referencia deíctica en el texto. En general, sin embargo, la mayoría de los paquetes tienen una sola función. Examinamos los que eran potencialmente multifuncionales de acuerdo a nuestras listas y los clasificamos de acuerdo a su función más frecuente.

Las tres principales categorías funcionales tienen varias sub-categorías asociadas a funciones y significados más específicos. El Cuadro N° 2 muestra un listado de estas sub-categorías junto con paquetes ejemplificadores. La siguiente sección, compara el uso de los paquetes con esas funciones en las diferentes disciplinas.

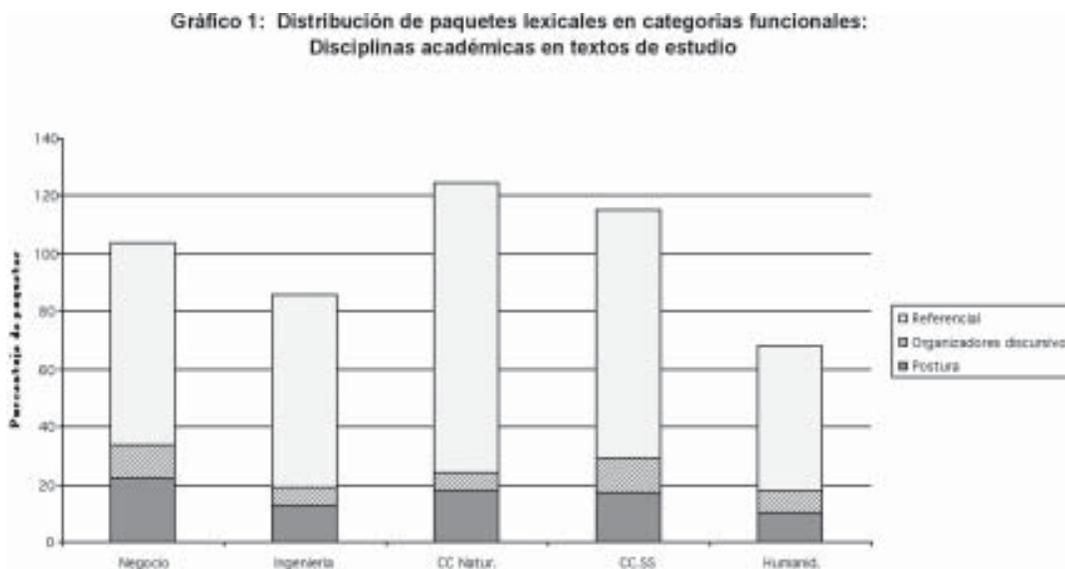
4. Paquetes léxicos en textos de estudio de diversas disciplinas académicas

El Gráfico N° 1 muestra diferencias sorprendentes entre las disciplinas académicas en el uso general de los paquetes léxicos. Las ciencias naturales y las sociales muestran mayor uso de paquetes léxicos, mientras que las humanidades se encuentran en el otro extremo con relativamente pocos paquetes léxicos diferentes.

Estos resultados requieren ser interpretados con cautela porque, tal como se ha insistido, el análisis se basa en un corpus relativamente reducido, y los ejemplos para cada disciplina no han sido normalizados en cuanto al tamaño (número de textos, número de palabras). Sin embargo, los patrones que se observan en el Gráfico N° 1, no se correlacionan con el tamaño de la muestra. En particular, los tres registros con las muestras más grandes –ciencias sociales, humanidades y ciencias naturales– ocupan los extremos opuestos del *continuum* en cuanto al uso de paquetes léxicos. Por ello, podemos tener bastante confianza en que estas son diferencias genuinas más que meros artificios producto de las diferencias en las muestras.

Cuadro Nº 2.
Categorías funcionales de los paquetes léxicos.

I. Expresiones Posturales	II. Organizadores Discursivos	III. Expresiones Referenciales
<p>A. Posturas epistémicas Personales: Yo no sé si, yo no sé qué</p> <p>Impersonales: Es más probable, el hecho de que, es muy posible que</p> <p>B. Posturas actitudinales/ modalidades</p> <p>B.1. Deseabilidad Personal: Si tú lo quieres, yo no lo quiero</p> <p>B.2. Obligación/ mandato Personal: Él o ella debieran, no estás obligado a</p> <p>Impersonal Es de absoluta necesidad, es (muy) necesario que</p> <p>B.3. Intención/predicción Personal: Yo no lo haré, nosotros vamos a hacerlo</p> <p>Impersonal: Esto va a suceder, va a depender de</p> <p>B.4. Habilidad/esfuerzo Personal: Yo te puedo dar</p> <p>Impersonal: Se puede usar para, es posible de hacer</p> <p>B.5. Importancia Es importante que lo, un papel importante en</p>	<p>A. Introducción de tópico/foco Si tú te fijas, vamos a hablar de</p> <p>B. Elaboración del tópico/aclaración/ perspectiva Desde otro punto de vista, a lo que se suma, es también pertinente para</p>	<p>A. Identificación/foco Ese es el que, uno de los más, los de ustedes que</p> <p>B. Imprecisión O algo como eso, y cosas de ese tipo</p> <p>C. Especificación de atributos</p> <p>C.1. Especificación cuantitativa El resto de los, una gran parte de, un gran número de</p> <p>C.2. Expresiones matemáticas Como una función de, en relación proporcional con, es igual a la</p> <p>C.3. Atributos enmarcadores tangibles Del tamaño de un, la composición de la, bajo la apariencia de</p> <p>C.4. Atributos enmarcadores intangibles La naturaleza de la, en el caso de, en términos de la, como resultado de la</p> <p>D. Referencia temporal, espacial, textual y otras.</p> <p>D.1. Referencias geográficas/países En los Estados Unidos, de los Estados Unidos</p> <p>D.2. Referencia o marco de locación general En la misma dirección, al lado derecho de, en la base de</p> <p>D.3. Referencia temporal La primera vez que, en el largo plazo, en un futuro cercano</p> <p>D.4. Deixis textual Como muestra el cuadro, como se discutió en el capítulo</p> <p>D.5. Referencia multifuncional Al final de la, en el comienzo de, al término de la</p> <p>D.6. Otras referencias</p> <p>Tangibles: Informes financieros desde el, miembros de una sociedad</p> <p>Intangibles: Economía como un todo, tendencia marginal a consumir, el sistema de control interno, la energía cinética de</p>



Posiblemente, dos factores son pertinentes aquí: el contenido técnico y las preferencias estilísticas. Los libros de ciencias naturales transmiten contenido técnico denso, y por eso utilizan términos y expresiones específicas para referirse a ese contenido, dirigidos a lograr una transmisión explícita del significado. Es posible que este apoyo en un conjunto específico de términos técnicos contribuya al uso frecuente de paquetes léxicos. Por contraste, los textos de humanidades se dedican más a la discusión crítica de ideas e interpretaciones. Además, los autores en humanidades aprecian la variación estilística, formulando, a menudo, la misma idea de múltiples formas para lograr un determinado estilo. Esta variación habría contribuido al menor uso de paquetes léxicos.

Una interpretación más completa de estas diferencias entre los registros, se puede lograr considerando la segmentación de las categorías funcionales. Tal como lo muestra el Gráfico N° 1, los organizadores discursivos corresponden a la categoría menos importante en todas las disciplinas. Este resultado es un tanto sorprendente, dado que los estudiantes necesitan signos explícitos de la organización del discurso para facilitar la comprensión; al respecto, Biber, Conrad y Cortés (2004) muestran que los paquetes léxicos, en tanto organizadores discursivos, son mucho más frecuentes en la enseñanza en el aula.

Cuando se da el caso del uso de paquetes léxicos, en los textos de estudio, generalmente funcionan como pistas explícitas de que lo que vendrá a

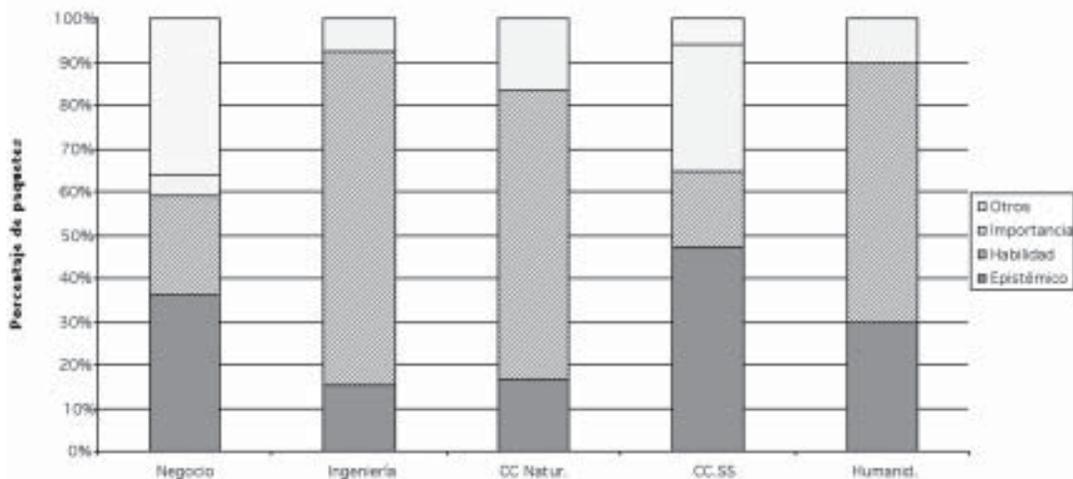
continuación será una mayor elaboración, clarificación o punto de contraste. Por ejemplo:

- *Los mundos vitales de las sociedades arcaicas son, en principio, accesibles vía el conocimiento intuitivo de sus miembros; **pero al mismo tiempo**, ellos porfiadamente escapan a nuestra comprensión.*
- *La historia de la teoría sociológica, **así como también el actual estado de la teoría sociológica**, es el precipitado de docenas de tales episodios intelectuales.*
- *Algunas normas son proscriptivas, ordenando lo que no deberíamos hacer [...] Las normas prescriptivas, **por otra parte**, establecen lo que debemos hacer [...]*

Los paquetes posturales tampoco son especialmente comunes en los textos de estudio, aun cuando son más importantes que los organizadores discursivos. Los paquetes posturales son más comunes en negocios (donde constituyen cerca del 20% de todos los paquetes).

Un análisis más detallado muestra que los paquetes posturales se usan con diferentes propósitos en las diferentes disciplinas. El Gráfico N° 2 muestra la segmentación proporcional de los paquetes posturales según su función. Varias funciones posturales casi nunca se expresan mediante paquetes en los textos de estudios, incluyendo obligación, deseo, intención/predicción, y, en general, expresiones de posturas

Gráfico 2: Distribución de paquetes posturales en diversas categorías funcionales: Disciplinas académicas en libros de texto



personales. Sin embargo, estas funciones de posturas personales son muy comunes en la sala de clases y en registros orales y escritos universitarios no académicos. En cambio, en los libros encontramos paquetes posturales que cumplen tres categorías funcionales principales: epistémico, habilidad e importancia. Es interesante observar que dos de estas categorías –habilidad e importancia– están restringidas, en primer lugar, a los textos de estudio. Es decir, mientras que los paquetes epistémicos se encuentran en la mayoría de los registros universitarios, los paquetes que cumplen funciones de habilidad e importancia se encuentran principalmente en los textos de estudio.

Dado que estos textos tienen como propósito primario transmitir conocimiento, es sorprendente que los significados epistémicos no sean la categoría principal de los paquetes posturales en este registro. Tal vez sea aún más sorprendente que los paquetes epistémicos sean los menos comunes en ciencias naturales e ingeniería. En contraste, las ciencias sociales se apoyan más en paquetes posturales, y son también relativamente comunes en negocios y humanidades. La base de las ciencias sociales es usualmente un conjunto de patrones típicos más que leyes de la naturaleza que no admiten excepciones. Además, gran parte de las discusiones en ciencias sociales y humanidades comparan diferentes puntos de vista y la probabilidad de que uno u otro sea preferible. Los paquetes epistémicos son especialmente

importantes en estos discursos interpretativos, en los que los estudiantes necesitan distinguir el estatus epistémico de las diferentes proposiciones (e.g. fáctico versus probable versus dudoso). Por ejemplo!:

- Estos investigadores concluyen que, dadas las condiciones económicas de la mayoría de las sociedades en desarrollo y la incapacidad de los gobiernos de proveer programas sustanciales de servicio social, **es muy probable que** haya un número creciente de padres de edad que no poseen propiedades, pensiones ni ahorros en su ancianidad.
- Se reinforma con mayor frecuencia la cesantía entre los nacidos en la ciudad que entre los inmigrantes, esto se debe a que las familias de los primeros **tienen mayores posibilidades de** estar ya bien establecidos en la economía urbana.
- Siempre existe la preocupación de que lo que se descubre en estudios interculturales que usan estas pruebas se confundan **por el hecho de que** las pruebas puedan ser significativas solo para sujetos estadounidenses.

Debido a que los paquetes léxicos epistémicos en los textos de ciencias sociales cumplen funciones interpretativas, a veces, encontramos paquetes que expresan tanto certeza como probabilidad en el mismo discurso, ilustrando cómo una tendencia puede ser considerada como un “hecho” en dicha disciplina:

- Esta diferencia racial puede estar asociada **al hecho de que** los estadounidenses africanos **tienen mayor posibilidad de ser** pobres y sufrir separación marital.

Los paquetes posturales de habilidad son, generalmente, más importantes que los paquetes epistémicos en textos de estudio universitarios, pero son especialmente importantes en ingeniería y en ciencias naturales. La mayoría de estas son construcciones pasivas, referidas a las habilidades del investigador o las del estudiante lector para llevar a cabo los procesos o cálculos técnicos:

- La longitud de la línea de acción **puede determinarse mediante** la Ecuación 7.5
- Las apropiadas correlaciones de la libre convección n-Pentano-cobre del capítulo **pueden ser utilizadas para** estimar los coeficientes de transmisión de calor.
- El teorema establece que **existe la posibilidad de** escribir un programa computacional usando solo tres estructuras de control básicas.

Los paquetes de habilidad también son importantes en los libros de texto de humanidades, donde identifican los procesos interpretativos que los alumnos debieran aprender, como en:

- Para identificar una falacia de relevancia, pues, uno **debe ser capaz de** distinguir la evidencia legítima de diversas formas de apelación emocional.
- Siempre que se sospeche que se está cometiendo esa falacia, él o ella **debiera ser capaz de** identificar la correcta solución.

Por último, los paquetes posturales que señalan importancia son especialmente relevantes en ciencias sociales, por ejemplo:

- Los temas del federalismo siguen siendo **una parte importante de** nuestra agenda política.
- Debido a que la cultura cumple un rol tan básico y fundamental, **es de suma importancia** examinar exactamente cómo la cultura juega ese papel en nuestras vidas.

Es interesante comprobar que los libros de ciencias sociales usan la mayoría de los paquetes para expresar tanto certeza como importancia. Estos paquetes proveen señales expresas al (estudiante) lector de los puntos a los que debiera prestar especial atención. En el otro extremo, los libros de texto de ingeniería prácticamente no utilizan paquetes lexicales que indiquen certeza o importancia.

El Gráfico N° 1, muestra que los paquetes referenciales corresponden a la categoría funcional más frecuente en todas las disciplinas. Los paquetes referenciales son especialmente importantes en ciencias naturales y, en muy menor medida, en ciencias sociales. Sin embargo, el Gráfico N° 3 muestra que existen notables diferencias en las funciones específicas de estos paquetes en las diferentes disciplinas.

Un patrón importante que se observa en el Gráfico N° 3 es que los paquetes de cantidad/matemáticas y los paquetes intangibles, se encuentran en distribución complementaria. En un extremo, el 45% de todos los paquetes referenciales en los textos de ingeniería expresan significados cuantitativos/matemáticos, mientras que solo un 20% de los paquetes referenciales expresan significados intangibles. En el otro extremo, 55% de los paquetes referenciales en los textos de humanidades expresan significados intangibles (versus un 8% con significado cuantitativo). Los textos de estudio de ciencias naturales son relativamente similares a los de ingeniería en cuanto a la preferencia por paquetes cuantitativos/matemáticos, mientras que las ciencias sociales son más parecidas a las humanidades (con preferencia por los paquetes intangibles).

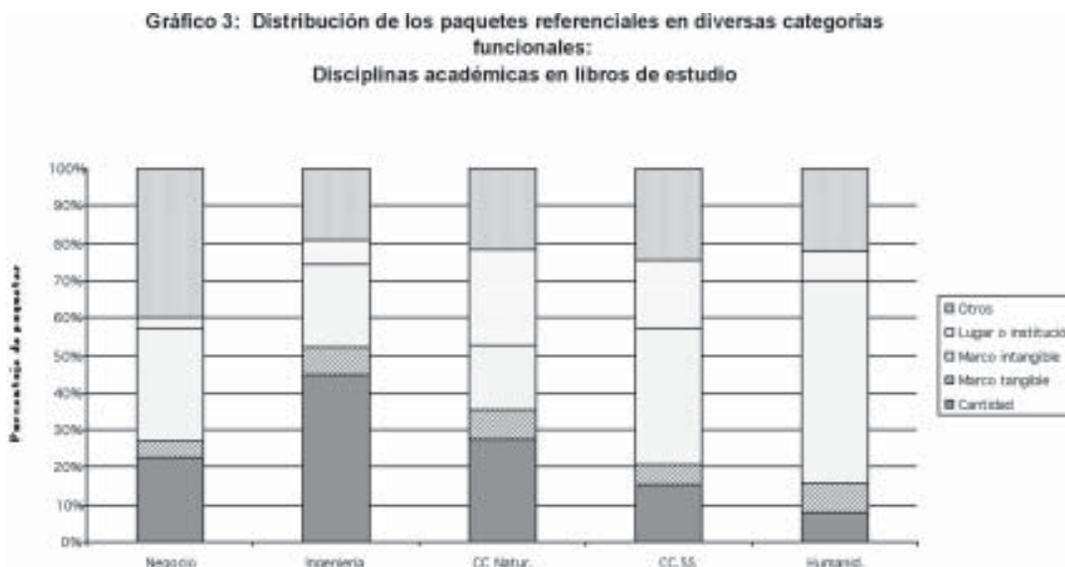
Los paquetes cuantitativos/matemáticos en los libros de textos de ingeniería a menudo identifican el estatus matemático de la frase nominal siguiente, por ejemplo:

- y el número de, a una aceleración de, por un factor de, grados de libertad de, expresado en términos de, en la magnitud de, la proporción de la, la raíz cuadrada de, el valor de la,

Estos paquetes se usan para identificar las relaciones entre los componentes y las variables utilizadas en ecuaciones matemáticas:

- La frecuencia de resonancia depende de **la proporción de las** masas con la tasa de elasticidad total de los neumáticos.
- Las propiedades de respuesta se pueden presentar examinando la ganancia de respuesta **como una función de** la frecuencia,
- Un 60% de solución azucarada acuosa a 20°C fluye por la cama **a una razón de** 244 libras minutos⁻¹ negativo.

En contraste, los paquetes referenciales intangibles predominan en los textos de humanidades. Por ejemplo:



- La naturaleza de la, en el caso de, una parte de la, en el transcurso de, en el proceso de, en el contexto de, a la luz de, la cuestión de si, de la relación entre, en respuesta a la, la idea de una, un análisis de la.

Estos paquetes proveen un marco referencial interpretativo para la siguiente frase nominal, destacando un atributo abstracto como importante:

- En segundo lugar, identifica **la naturaleza de los talentos especiales** que Dios ha dado a los artistas...
- **En el proceso de** analizarse uno mismo, uno debería contestar tres preguntas...
- Los poemas de Manfredi deben ser vistos **en el contexto de** la floreciente literatura dedicada a definir a la mujer y la feminidad...
- El Cagliostro de Houclon presenta con exquisita agudeza **el carácter de la** fraudulenta santidad.

Los textos de ciencias naturales difieren de los de las otras disciplinas en cuanto a que exhiben un uso frecuente de paquetes léxicos de locación, como en:

- En el fondo de, en el lugar de arriba, en la dirección de, en la cercanía de, a ambos lados de, en la superficie de, en el centro de.

Estos también tienen un tipo de función enmarcadora, pero tienen un significado mucho más concreto, al identificar la locación física de un referente, o la relación física de un referente con otro referente. Por ejemplo:

- Los minerales menos solubles se concentran en la orilla del antiguo lago y se encuentran **en el fondo de** la secuencia evaporadora.
- Se echa agua sobre el tefra **en la cercanía de** los pozos de vapor.
- Las áreas con conteo bacterial alto coinciden con **la posición de la** pluma...

CONCLUSIONES

Aun cuando los paquetes lexicales no son normalmente idiotismos semánticos y, generalmente, no son unidades estructurales completas, constituyen algunos de los bloques textuales más importantes usados en el discurso oral o escrito. Es decir, estas combinaciones frásticas son extraordinariamente comunes en los textos orales y escritos, en tanto sirven para una importante gama de funciones discursivas.

La presente investigación se ha enfocado en textos de estudio, contrastando el uso de paquetes lexicales en diversas disciplinas. Los paquetes referenciales son los más importantes en todas ellas, aunque los paquetes posturales son también relativamente importantes. Sin embargo, se observa que estas disciplinas usan estos paquetes con frecuencias diferentes y en diferentes funciones básicas.

En términos generales, los paquetes lexicales son más frecuentes en los libros de estudio de las ciencias naturales y menos frecuentes en los de humani-

dades. Pero las diferencias entre las disciplinas, en cuanto a las preferencias por funciones específicas, son aún más notables. Por ejemplo, las ciencias sociales tienen una fuerte preferencia por los paquetes posturales epistémicos y de importancia, en contraste con las preferencias por los paquetes posturales de habilidad en ingeniería y ciencias naturales. En cuanto a los paquetes referenciales, los textos de ingeniería tienen una fuerte preferencia por los paquetes cuantitativos/matemáticos, en contraste con la preferencia por los paquetes enmarcadores intangibles en los libros de textos de humanidades.

El estudio de los paquetes léxicos se basa en el supuesto general de que la frecuencia de los patrones o estructuras no es accidental. Sin embargo, también es un hecho de que la frecuencia no es en sí misma explicativa. Más bien, la evidencia aportada por la frecuencia basada en un *corpus* proporciona hechos descriptivos que requieren explicación. En este estudio, los hechos que requieren explicación son la existencia de secuencias frásticas frecuentes que no representan estructuras lingüísticas bien definidas. El examen de estas secuencias frásticas en contextos textuales muestra que son importantes bloques del discurso, asociados con funciones discursivas básicas. Los libros de estudio de diferentes disciplinas académicas usan estos paquetes léxicos en diferentes grados y para diferentes funciones, según el objetivo comunicativo básico de cada disciplina.

REFERENCIAS BIBLIOGRÁFICAS

- Altenberg, B. (1993). Recurrent word combinations in spoken English. En J. D'Arcy (Ed.), *Proceedings of the Fifth Nordic Conference for English Studies* (pp. 17-27). Reykjavik: Island University.
- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. En A. Cowie (Ed.), *Phraseology* (pp. 101-122). Oxford: Oxford University Press.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. & Conrad, S. (1999). Lexical bundles in conversation and academic prose. En H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honor of Stig Johansson* (pp. 181-189). Amsterdam: Rodopi.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *The Longman grammar of spoken and written English*. London: Longman.
- Biber, D., Conrad, S., Reppen, R., Byrd, P. & Helt, M. (2002). Speaking and writing in the university: A multi-dimensional comparison. *TESOL Quarterly*, 36(1), 9-48.
- Biber, D., Conrad, S. & Cortés, V. (2003). Lexical bundles in speech and writing: An initial taxonomy. En A. Wilson, P. Rayson, & T. McEnery (Eds.), *Corpus linguistics by the lune: A festschrift for Geoffrey Leech* (pp. 71-92). Frankfurt: Peter Lang.
- Biber, D., Conrad, S. & Cortés, V. (2004). *If you look at...: Lexical bundles in university teaching and textbooks*. *Applied Linguistics*, 25(3), 371-405.
- Biber, D., Conrad, S., Reppen, R., Byrd, P., Helt, M., Clark, V., Cortés, V., Csomay, E. & Urzúa, A. (2004). *Representing language use in the university: Analysis of the TOEFL 2000 Spoken and Written Academic Language Corpus*. *TOEFL Monograph Series*. Princeton, NJ: Educational Testing Service.
- Brown, P. & Fraser, C. (1979). Speech as a marker of situation. En K. R. Scherer & H. Giles (Eds.), *Social markers in speech* (pp. 33-62). Cambridge: Cambridge University Press.
- Butler, C.S. (1997). Repeated word combinations in spoken and written text: Some implications for Functional Grammar. En C.S. Butler, J.H. Connolly, R.A. Gatward, & R.M. Vismans (Eds.), *A fund of ideas: Recent developments in Functional Grammar* (pp. 60-77). Amsterdam: IFOTT, Amsterdam University.
- Cortés, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes Journal*, 23, 397-423.
- De Cock, S. (1998). A recurrent word combination approach to the study of formulae in the speech of native and non-native speakers of English. *International Journal of Corpus Linguistics*, 3(1), 59-80.
- Ellis, N. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18(1), 91-126.
- Halliday, M.A.K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

- Howarth, P. (1996). *Phraseology in English academic writing*. Tübingen: Max Niemeyer.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19(1), 24-44.
- Hymes, D. (1974). *Foundations in sociolinguistics*. Philadelphia: Pennsylvania University Press.
- Partington, A. & Morley, J. (2002). *From frequency to ideology: Investigating word and cluster/bundle frequency in political debate*. Ponencia presentada en Teaching & Language Corpora (TALC, 2002) en The University Residential Centre, Bertinoro, Italy.
- Salem, A. (1987). *Pratique des Segments Répétés*. Paris: Institut National de la Langue Française.
- Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, 16(2), 180-205.
- Wray, A. & M. Perkins. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, 20, 1-28.

NOTAS

¹ En algunos casos, se superponen dos paquetes de 4 palabras, dando origen a paquetes de 5 palabras. Por ejemplo, las secuencias *es más probable que* y *más probablemente van a ser*, ocurren más frecuentemente como la secuencia de 5 palabras *es más probable que sean*.