

EVALUACIÓN DE RECURSOS LÉXICOS *ON-LINE* PARA IFE: DATOS EMPÍRICOS E INTROSPECCIÓN LINGÜÍSTICA

ALEJANDRO CURADO FUENTES
Universidad de Extremadura

RESUMEN. *Este estudio se centra en una evaluación de recursos léxicos electrónicos para IFE (Inglés para fines específicos). Se describe principalmente la gestión de datos empíricos en un corpus de inglés específico y la información lingüística obtenida a través de tareas comunicativas. El objetivo es formular criterios generales y específicos para el desarrollo de recursos léxicos (ejercicios y textos). Las actividades on-line se evalúan según el grado de efectividad comunicativa que implican en las tareas, para lo cual la capacidad introspectiva del profesor/investigador es también fundamental. En el proceso de investigación, se reconocen dos pasos fundamentales: primero, el diseño de material adecuado, integrando datos léxico-gramaticales mediante recursos electrónicos, y, seguidamente, la evaluación del aprendizaje o asimilación en tareas comunicativas específicas.*

PALABRAS CLAVE: *corpus, tarea comunicativa, introspección, recursos, evaluación.*

ABSTRACT. *This paper focuses on an assessment of electronic lexical resources for ESP (English for Specific Purposes). We describe how both corpus-driven empirical data and linguistic information obtained from communicative task performance may contribute to formulating general and specific criteria for the development of such lexical resources (exercises and texts). The on-line activities are assessed according to the degree of communicative effectiveness that the tasks involve, and for which the teacher / researcher's introspective skills are equally noteworthy. The two steps encompassed are the provision of material integrating lexical-grammatical content, and, secondly, the evaluation of this content intake as demonstrated in communicative tasks.*

KEYWORDS: *corpus, communicative task, introspection, resources, evaluation.*

1. INTRODUCCIÓN

La importancia creciente del sector de IFA (*Inglés para fines académicos*) como parte de IFE (*Inglés para fines específicos*) conduce, entre otros aspectos, a un énfasis

especial en análisis de género y tipos de textos (cf. Johns 1991; Jordan 1997; Dudley-Evans & St. John 1998). Se trata de una relación entre nivel textual y discente, en que el docente funcionaría como proveedor clave del material textual, un “*carrier content*” que se explotaría mediante tareas dirigidas al aprendizaje de lenguas especializadas (Dudley-Evans & St. John 1998: 174). Un ejemplo lo constituirían las carreras técnicas e ingenierías, donde resulta fundamental ofrecer no sólo una actualización de conceptos sino también de competencias académicas y profesionales. Como señalan diversos autores (entre otros, Flowerdew 2001), en un enfoque comunicativo acertado de dichos elementos discursivos, el diseño de material docente a partir de corpus específicos puede resultar muy útil para trabajar los géneros académicos.

En este contexto de IFE, como en todo entorno de aprendizaje que se preste a algún tipo de medición socio-lingüística (entre tipos de hablantes con funciones comunicativas concretas), resulta positiva una evaluación empírica de factores lingüísticos y paralingüísticos (Ooi 1998: 14). Esta observación del lenguaje en uso podría iniciarse como respuesta a tener que hacer frente a las dificultades propias de todo alumno de IFE en el desempeño de tareas de comprensión/producción de discurso académico o especializado (Cheng et al. 2003: 174). Tal conducción de tareas muchas veces encuentra, entre otros escollos, un obstáculo lingüístico considerable en el llamado nivel inferior o “*bottom-up*” de datos léxico-gramaticales, como indican diferentes autores (cf. Conrad 1996; Aston 1997; Smith 2003).

En nuestra propia experiencia investigadora, el uso de material diseñado a partir del cotejo empírico de datos en corpus específicos suele proporcionar un encauzamiento positivo para el desarrollo comunicativo de tareas propias de IFE (p.ej., presentaciones orales de temas específicos en Informática –cf. Curado 2002– y Empresariales –cf. Curado 2003–). También como resultado de nuestra línea de trabajo, se comprueba que una integración adecuada de recursos multimedia en red tiende a facilitar el acceso y manejo del material derivado de corpus para las tareas en IFE. Según los resultados obtenidos en nuestros estudios de casos concretos de aprendizaje, para proporcionar una condición rentable de trabajo, la gradación del material y los recursos electrónicos basados en datos de corpus es importante. Se trata, pues, de un enfoque clasificatorio llevado a cabo a partir de la adecuación de actividades y recursos *on-line* a las necesidades comunicativas percibidas entre el alumnado.

Este trabajo describe principalmente la gestión de datos empíricos a partir de un corpus específico así como la información lingüística derivada de tareas comunicativas para el desarrollo de recursos léxicos (ejercicios y textos) electrónicos en cursos de IFE. Estas actividades *on-line* se evalúan en acuerdo con la efectividad comunicativa que pudieran tener para el alumnado. Para gestionar adecuadamente el vocabulario de interés con el fin comunicativo académico y/o profesional del estudiante, la capacidad introspectiva del profesor/investigador es también significativa. Los objetivos específicos del estudio serían proveer mejor y más material basado en datos léxicos de corpus mediante recursos electrónicos, y, seguidamente, una evaluación de su aprendizaje o asimilación léxico-gramatical en tareas comunicativas específicas.

2. DISEÑO DEL CORPUS

2.1. ASPECTOS ACADÉMICOS Y TIPOS DE TEXTOS

Como criterio general a la hora de establecer una referencia temática en nuestro corpus, se abarca una serie de contenidos (*carrier content*) comunes a las carreras de Informática y Empresariales. Se examinan los intereses académicos de los discentes mediante el contraste de programas y planes de estudios, observando que un elemento afín es la gestión empresarial mediante las nuevas tecnologías, en asignaturas como Sistemas de gestión informática, Contabilidad, Gestión empresarial, Marketing y Estadística. Los temas comunes (entre al menos dos) reseñados en programas y lecturas de estas cinco asignaturas aparecen en la tabla 1:

Gestión de bases de datos.
Soporte técnico en empresas.
Programas multimedia.
Aplicaciones de Office.
Desarrollo de software para contabilidad.
Administración y gestión de red.
Aplicaciones en red.
Comercio electrónico.
Software de finanzas.
Publicidad en Internet.
Comunicación y publicación electrónica.
Tecnología audio-visual.
Comercialización de software.
Aplicaciones para gestión de proyectos.
Análisis de datos contables en empresas.
Recursos para la empresa pública.
Librerías virtuales.

Tabla 1: Lista de temas abarcados en nuestro corpus.

Este esquema del corpus por temas sirve fundamentalmente como guía o pauta para la selección de material electrónico durante la construcción de dicho corpus especializado (actualmente en fase de publicación [Curado, en prensa]). Así pues, nos basamos en un entorno de comunicación específico que se rige, principalmente, por dos funciones: información y persuasión. Se incluyen tanto textos escritos como orales precisamente para reflejar lo más fielmente posible los diferentes encuentros

comunicativos: No sólo lecturas de textos descriptivos (*chapters* [capítulos en libros de texto] y *reports* [informes técnicos]) sobre conceptos puntuales de gestión de la información (función informativa), sino también la presentación a modo de monólogo (p.ej., *reviews* [revisiones]) o diálogo (p.ej., *discussions* [debates o entrevistas]) de la revisión de un artículo o producto en la tecnología de la información, analizando aspectos positivos y negativos (fines de persuasión). Esta combinación puntual de registros escritos y orales orienta la inclusión de diferentes textos en las categorías temáticas anteriormente aludidas. La figura 1 resume nuestra clasificación de los textos del corpus, unos tipos textuales seleccionados para los fines comunicativos del entorno especializado.

El tamaño y distribución de los textos obedecen también a criterios básicos de estilo y géneros académicos abarcados en cursos de IFE, donde hemos observado que el nivel intermedio y/o intermedio-avanzado mayoritario (en segundo y tercer curso) no encuentra demasiada dificultad con este tipo de lectura recopilada. Esta combinación de géneros académicos (p.ej., capítulos de libro) y tipos de textos en el entorno específico (p.ej., debates) aporta una sincronización importante que refleja en gran parte lo que ocurre en el espacio de la comunicación académica, coincidiendo en nuestra apreciación con Conrad (1996: 301).

La inclusión de debates y entrevistas (*discussions*) sobre temas especializados interesa especialmente para el marco de tareas comunicativas orales en tercer curso, donde los alumnos tienen que hacer uso de un lenguaje especializado y, a la vez, de frases y usos, giros o circunloquios para encauzar sus ideas u opiniones. En este sentido, este tipo de texto oral puede contener diferentes “voces de géneros” (mi traducción) (Cook 2001: 194) para no sólo informar sino también persuadir al oyente/receptor del mensaje. De ahí que un hablante/emisor pueda resultar más “académico”, apoyando su discurso en datos fehacientes y de investigación, mientras que otros usuarios puedan preferir un discurso más coloquial, haciendo mayor uso de ejemplos que fortalezcan sus ideas, e ilustrando la variedad y riqueza del discurso académico (Conrad 1996: 302). Todo ello tiende a favorecer las habilidades de pensamiento crítico entre los alumnos, que se sienten así animados a formar una opinión mediante el contraste de otras (finalidad importante en la asignación de tareas comunicativas en nuestros cursos, como veremos en el apartado de Resultados).

La complementación de estos tipos de textos orales y escritos también interesaría en el sentido que Jolliffe y Brier (1988: 43) apuntan: “*students cannot learn to write in their disciplines simply by reading textbooks*”. Además, como afirma Cook (2001: 201), la comunicación entre emisor y receptor en el marco especializado se ve favorecida cuando “se identifica un tipo de texto determinado y se adopta una posición específica de interpretación” (mi traducción). En este sentido, se estaría haciendo a los discentes partícipes de este entorno para que activen roles comunicativos en las tareas académicas. De ahí que el diseño del corpus se haga considerando estos factores de aprendizaje relacionados con los diferentes tipos de textos, lecturas, y las tareas comunicativas posibles asociadas a los registros abarcados.

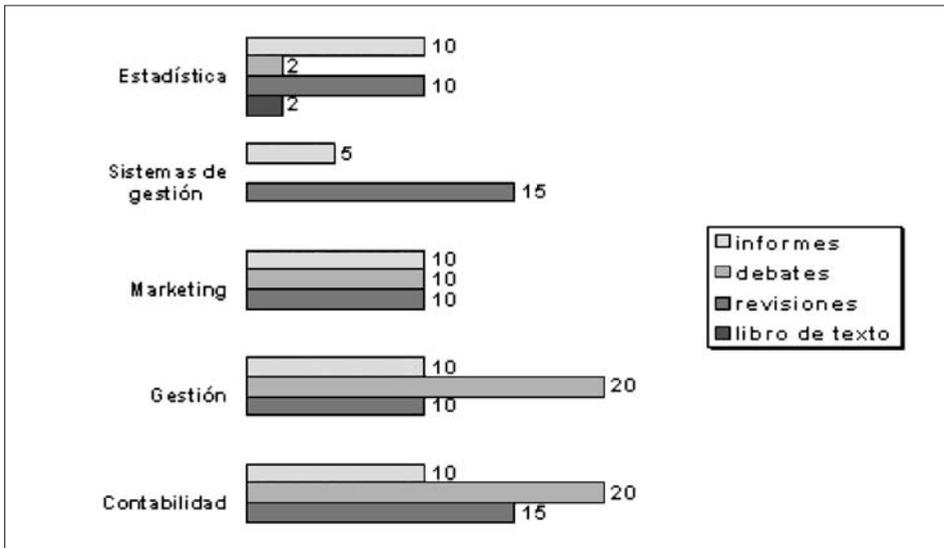


Figura 1: Número de fuentes y distribución de textos en corpus de Inglés de la gestión empresarial y tecnología informática.

2.2. ASPECTOS LÉXICO-GRAMATICALES

El análisis de datos léxico-gramaticales en el corpus constituye un paso crucial en su diseño y futura utilización como base/fuente de material lingüístico. La complementación de los factores de frecuencia y dispersión léxicas sirven para determinar qué se entiende por lenguaje típico en nuestro espacio comunicativo. Asimismo, el léxico ha de interpretarse dentro del contexto temático y de encuentro sociolingüístico en que tiene lugar (Sinclair 2004: 128). Nos apoyamos, por tanto, en una primera fase, en los aspectos empíricos que guardan las asociaciones léxico-gramaticales, claves en nuestra colección textual (colocaciones, frases léxicas, recursos léxico-retóricos). En una fase posterior, cabría hablar de la introspección del profesor o investigador para dilucidar qué tipo de construcciones léxico-gramaticales resultan más importantes de cara al desarrollo real de tareas comunicativas (según temas predominantes, tipo de comunicación favorecida entre emisor y receptor, y finalidad principal de dicha comunicación).

2.2.1. Aspectos empíricos

En el primer caso, las listas de frecuencia y dispersión léxicas deben cotejarse con listas de corpus afines al nuestro, pero de mayor envergadura. La afinidad de la colección textual se define, sobre todo, en cuanto a campo o área temática se refiere, pues se trata, en este caso, de encontrar elementos léxico-gramaticales clave o relevantes en nuestro marco temático general (“*key words signalling the ‘aboutness’ of the texts*” –Scott 1997: 234–).

Así, se ha utilizado como corpus de referencia para nuestro análisis el de James y Purchase (1996) *English in Business Studies and Economics* (con 1,002.594 palabras, 29.025 palabras individualmente diferentes, y una ratio de 0,017 entre total de palabras y palabras distintas). Nuestro corpus de inglés de la gestión empresarial y tecnología informática, por otra parte, contiene 372.761 palabras y 12.923 vocablos individuales (una ratio de 3,47 por ciento). De entrada, por tanto, aparecen menos repeticiones léxicas en nuestros textos en comparación con el corpus de referencia elegido, y en contraste, más vocabulario de corte específico.

En general y a modo de resumen de nuestro análisis, se observa que, al realizar la comparación de las dos listas de frecuencia léxica mediante la herramienta de concordancia *WordSmith Tools* (Scott 1996), aparece sobre todo vocabulario de contenido en las posiciones superiores de la lista: *design, value, example, model, analysis, data, profit, factors, available, different*, etc: Son sustantivos en un 62 por ciento de los casos, y a continuación adjetivos y verbos (*available, domestic, international, run, computed, see*, etc), también modales/verbos auxiliares como *can, is, are*. Finalmente, es relevante la ocupación entre las 100 primeras palabras clave en nuestro corpus de pronombres personales como *you* y *I* (muy frecuentes y típicos, en efecto, en el discurso académico oral [cf. Fortanet 2004]).

La observación de cómo se combinan estos vocablos en los textos resulta positiva para discernir el corte más o menos técnico de su comportamiento léxico. Así, por ejemplo, estudiando las colocaciones de un sustantivo común en el corpus, *data*, se observa que *data transfer* es una construcción típica en nuestro corpus, apareciendo en diversos textos y categorías temáticas. Se trataría, por tanto, de una unidad de corte semitécnico, común en el campo de la tecnología de la empresa, y que hace referencia a la actividad de transmitir información entre equipos informáticos. Este matiz del vocablo no aparecería en un corpus sobre economía como el utilizado de referencia (James & Purchase 1996). Tampoco aparecería (con *data*) una cadena léxica como *arrangement of the data*, con significado semitécnico en nuestro corpus. De otro lado, una utilización más restringida y, por tanto, técnica de *data* sí se vería en *data mining techniques*, pues aparece con asiduidad, pero sólo en tres de los textos pertenecientes a la categoría temática “Sistemas de gestión informática”. No se da ninguno de estos casos en el corpus de James y Purchase (1996)¹.

2.2.2. Introspección lingüística

Apoyándonos en las palabras de Leech (1991: 14) sobre la experimentación con datos derivados del corpus, en un consecutivo nivel de análisis lingüístico, percibimos la necesidad de introspección lingüística, que pasaría por dilucidar qué expresiones aportan un valor tangible para la formación del alumnado. Esta capacidad de análisis formado en la intuición no es posible, sin embargo, sin un estudio léxico previo que el profesor debe realizar con el corpus específico. Así, con el cotejo de la lista de frecuencia/dispersión léxica, las expresiones léxico-gramaticales se analizarían como construcciones

requeridas a partir de evaluaciones de distintas tareas académicas tanto escritas como orales. El profesor, en este sentido, construiría una base de datos con los errores léxico-gramaticales claros cometidos por los alumnos, pero también con sus aciertos y lenguaje efectivo, como se verá en el siguiente apartado de Resultados.

Asimismo, es tarea también importante en este nivel de introspección lingüística calibrar la necesidad de algunas construcciones para la comunicación de ciertas tareas. Un ejemplo sería la importancia del uso de marcadores en debates cortos sobre temas semi-técnicos en clase. Hemos observado que, en efecto, el empleo de elementos de transición o enlace como *you know, I think, for instance*, y otros de cohesión en el discurso, resulta importante en un diálogo académico para explicar las ideas más fluidamente (cf. Curado 2004(a)). En cualquier caso, también hemos notado que, para las tareas, la información léxico-gramatical se valoraría en función de factores como el tipo de carga temática abarcada (que se traten conceptos más o menos técnicos), el tipo de comunicación favorecida entre emisor y receptor (según sea un monólogo o diálogo, una composición formal o un borrador de un informe), y la finalidad principal de dicha comunicación (que se evalúen tanto el conocimiento del tema como la capacidad del alumno para desarrollar ideas, su proceso de realizar una investigación concreta y sintetizarla en un proyecto, etc).

3. RESULTADOS

Como se ha mencionado, nuestro mecanismo o método de trabajo es la integración adecuada de información empírica y datos internos sobre el uso del lenguaje en el aula. No podemos pasar por alto, por tanto, ni los datos léxico-gramaticales del corpus ni la capacidad para interpretar esos datos a la luz de la producción comunicativa del alumnado en tareas específicas. Los objetivos específicos de nuestra línea de actuación son, primeramente, la elaboración de recursos electrónicos léxicos apoyada en una gradación meditada de actividades, y a continuación, la ponderación de resultados positivos y/o negativos mediante la evaluación de los alumnos en una tarea comunicativa.

3.1. GRADACIÓN DE RECURSOS ELECTRÓNICOS

3.1.1. *Base de datos de producción léxico-gramatical en tareas*

Como se ha dicho, la tarea de recopilación del profesor/investigador de IFE necesita bastante de su introspección en el uso del lenguaje especializado. De ahí que la información lingüística derivada del corpus se complementa con su inspección de cómo se utiliza ese tipo de discurso para fines específicos. En nuestro caso, hemos llevado a cabo dicho estudio mediante el análisis contrastivo de los datos del corpus en dos tipos de tareas fundamentalmente: presentaciones orales cortas de un tema relacionado con la tecnología en la empresa, y composiciones en foros electrónicos académicos sobre un tema científico-técnico.

Nuestra metodología para contrastar la producción de elementos léxico-gramaticales en las tareas se fundamenta en tres módulos: lenguaje efectivo (cuando las expresiones utilizadas constituyen un lenguaje típico en el corpus: p.ej., *data transfer*), correcto (es el caso de construcciones léxicas válidas pero no típicas o específicas del corpus: p.ej., *transmission of data*), y erróneo (en el caso de errores léxico-gramaticales: p.ej., **sends of the data*). Los diferentes elementos lingüísticos se recopilan y catalogan convenientemente para su recuperación posterior con el fin de calibrar los distintos recursos *on-line*. La tabla 2 resume los datos léxicos recogidos en tres grupos de alumnos de IFE: Inglés para Informática (segundo [2ºI] y tercero [3ºI]) e Inglés para Empresariales (tercero [3ºE]) (en cada grupo hay diez alumnos con las dos tareas –escrita y oral– asignadas).

<i>Lenguaje</i>	2º I	3º I	3º E
Efectivo	12	18	13
Correcto	230	215	198
Errores	89	83	72

2º I = segundo curso de Informática / 3º I = tercer curso de Informática / 3º E = tercer curso de Empresariales

Tabla 2: Número de expresiones efectivas, correctas y erróneas producidas por los alumnos.

Cuantitativamente, los datos de la tabla 2 apuntan a dos o tres posibles lecturas: de un lado se cometen más errores en el curso más bajo, pero también es mayor en éste el número de expresiones correctas. De otro lado, los alumnos del curso de Empresariales cometen menos errores léxico-gramaticales, y sin embargo, producen menos expresiones (tanto correctas como efectivas). Finalmente, otra posible interpretación de los datos es que los alumnos del curso superior de Informática reflejan un conocimiento mayor de construcciones típicas del inglés de la gestión empresarial y tecnología informática. Su uso de expresiones correctas es igualmente notorio en comparación con los otros dos grupos.

En cuanto a la valoración cualitativa que sugieren las expresiones utilizadas, llegaríamos a dos posibles hipótesis en torno al conocimiento y uso de los datos léxico-gramaticales en tareas específicas: se tiende a cometer más errores en la producción de combinaciones de palabras de corte semitécnico o académico cuando se trata un tema técnico o especializado en una tarea oral, y por el contrario, en la tarea escrita de composiciones sobre temas menos especializados (de carácter científico-técnico más general), se hace un mayor uso de expresiones semitécnicas correctas, mas se falla considerablemente en cuestiones de tipo gramatical. Los tres grupos de alumnos podrían ilustrar esta diatriba, pero, por fijarnos en el que más elementos efectivos produjo en su discurso –los de tercero de Informática– cabe resaltar un gran dominio de expresiones semitécnicas en las composiciones escritas (12 efectivas y 46 correctas), pero también más errores gramaticales; mientras, en la tarea oral, el lenguaje efectivo de tercero de Informática es muy inferior en comparación con el cotejado en sus tareas escritas (tan sólo seis expresiones).

Un ejemplo de construcción semitécnica que se presta a errores y/o variaciones válidas sería *exchange + data with*, que es recurrente en nuestro corpus. Tanto los alumnos de Informática como de Empresariales recurrieron a esta construcción en sus presentaciones orales, pero sólo dos alumnos de tercero de Informática la utilizaron tal y como se refleja en el corpus. Los de Empresariales emplearon *exchange + data to* y *exchange + information with*, mientras que en segundo de Informática, las variantes fueron *share + data between*, *take + information to*, y *replace + information*, todas consideradas correctas. No obstante, también se dieron algunos errores como **interchange the data* (en tercer curso de Informática) y **go the information with* (en segundo curso de Informática). Otros errores registrados como elementos semitécnicos fallidos serían **treat with the problem* (en lugar de *deal with the issue*), **arrive in the computer* (en lugar de *transfer to the PC*), **situated on* (por *located in*), y **run in the PC* (por *run on the PC*).

Los dos últimos ejemplos ilustran fallos de corte léxico-gramatical. En éstos, aparecería una palabra gramatical o funcional (preposiciones, conjunciones, artículos, verbos auxiliares) empleada en un contexto equivocado. Otros ejemplos comunes de malformación léxico-gramatical incluirían el orden sintáctico y la conjugación de los tiempos verbales. La figura 2 especifica el número de errores cometidos por los tres cursos en la tarea oral. Los fallos de la tarea escrita son menos en cuanto a léxico semitécnico se refiere, y por el contrario, más en cuanto a gramática. Obviamente, los errores gramaticales son tan importantes como los léxicos (de hecho, inseparables) pero para los fines de nuestro estudio de recursos léxicos *on-line*, nos fijamos sobre todo en datos como los de la figura 2, que contribuyen a dirigir nuestro enfoque².

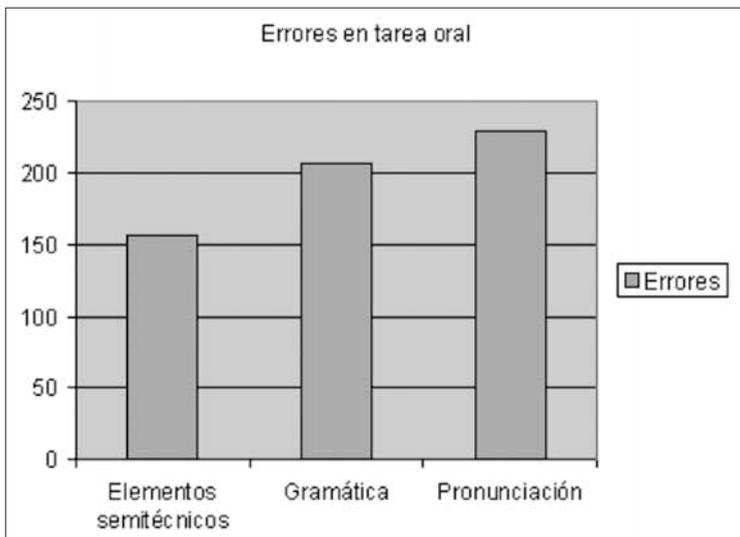


Figura 2: Número de errores lingüísticos registrados en 30 presentaciones orales sobre temas especializados relacionados con tecnología de la empresa.

3.1.2. Propuesta de actividades y recursos léxicos on-line

Nuestra propuesta para hacer frente a los escollos léxico-gramaticales destacados en la sección anterior se centra en la elaboración de material *on-line* para explotar vocabulario semiespecializado y técnico en contexto. En concreto, presentamos actividades, en un primer estadio, que midan la capacidad del alumno para discernir o reconocer la combinación de palabras semitécnicas en ejercicios de *fill-in-the-gap* y *matching*, resueltos *on-line* bajo la supervisión del profesor. A continuación, se introduce al alumno al nivel del párrafo/texto con actividades de concordancia electrónica, mediante las que tiene acceso al corpus de inglés para la gestión empresarial y tecnología informática, así como a diferentes posibilidades de explotación de lenguaje semitécnico y técnico en contexto.

3.1.2.a. Ejercicios de “Fill-in-the-gap” y “Matching”

Con estos dos tipos de actividades, correspondientes al nivel sintagmático del uso del léxico, se pone especial énfasis en la disposición y acceso a vocablos de alta frecuencia y dispersión en el corpus. Coxhead (2002) hace algo parecido con sus listas de palabras aca-

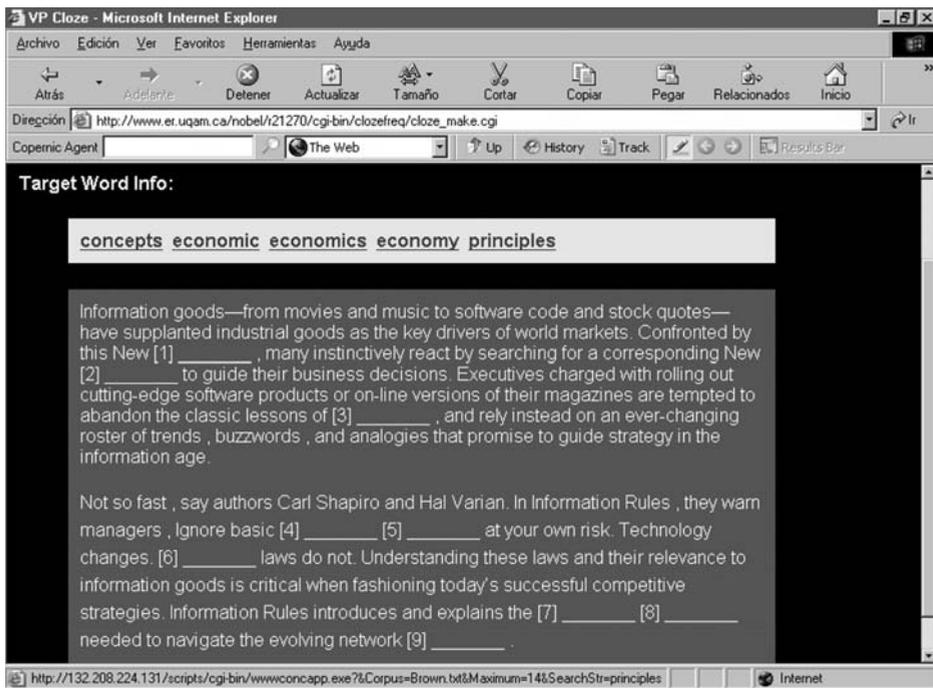


Figura 3: Sitio web donde se ofrece una serie de herramientas para elaborar *Fill-in-the-gaps* automáticos a partir de listas de palabras.

démicas en la red, donde utiliza dichas listas como referencia de qué términos va a explorar en los ejercicios *fill-in-the-gap*. En nuestro caso, elegimos las 100 primeras palabras de contenido de nuestra lista, que se incluirían en una base de datos que emplea el software libre *Cloze* (Cobb 2002). Los vocablos introducidos se eliminarían automáticamente de cualquier texto administrado que contuviera cualquier elemento de la lista. La figura 3 muestra cinco sustantivos (que aparecen en la lista de 100 palabras) suprimidos de un texto y por tanto debiendo ser seleccionados para completar la actividad. Además, el programa de Cobb (2002) ofrece la posibilidad de realizar enlaces electrónicos con la base de datos *Wordnet*, que proporcionaría el significado y múltiples ejemplos de uso del término en contexto. También existe el conocido programa *shareware Hot Potatoes 6.0*. (Half-baked software Inc., 2003), que tiene prestaciones parecidas para crear actividades de *Fill-in-the-gap* y *Matching*.

Según las respuestas e impresiones de los alumnos en clase, con este tipo de recursos se trabaja bien, proporcionando una visión clara de qué palabras se usan en un contexto determinado. Por ejemplo, un tema examinado fue la revisión de un programa para la seguridad en Internet, que contenía léxico técnico como *data encryption*, *decrypt the data*, y otros semitécnicos como *run + risk of being + verbo en participio*. En la mayoría de los



Figura 4: Muestra de ejercicio basado en las 500 primeras palabras semitécnicas del corpus.

casos, los alumnos fueron capaces de averiguar el sustantivo *data* en las expresiones restringidas, y en todos los ejercicios realizados con este texto, el verbo *run* fue identificado.

En nuestra página web (Curado 2004(b)), incluimos diversas actividades en el nivel sintagmático, tanto de *fill-in-the-gap* como *Matching*, a partir de listas de palabras. Se ha elaborado con algunas de las 500 primeras palabras de contenido de nuestro corpus listas semitécnicas útiles para la relación de combinaciones o colocaciones comunes al área de especialidad que nos ocupa. La figura 4 visualiza algunas de ellas.

El tipo de actividad *Matching* tiende a resultar más productiva para el reconocimiento de las combinaciones de palabras de contenido o colocaciones léxicas apropiadas en el contexto. Este resultado positivo se hace patente al administrar una serie de tests cortos o *quizzes* a los alumnos después de que hubiesen realizado alguna(s) de nuestras actividades de *Matching* en red. Un ejemplo es la figura 5, que recoge cinco verbos principales (en la parte superior) que podrían o no combinarse con alguno de los diez sustantivos en la parte inferior de la tabla. De los veinte alumnos que realizaron esta prueba, diez habían hecho algunos de nuestros ejercicios de *Matching* en red. De éstos, siete respondieron correctamente al test, mientras que de los otros diez (que no habían hecho ningún ejercicio de *Matching*), sólo tres consiguieron tener todas bien. A modo de indicación general, por tanto, el acceso a actividades de *Matching on-line* parece resultar positivo para incrementar o agudizar el reconocimiento/percepción de colocaciones léxicas importantes en el área como las que aparecen en la figura 5.

<i>Match each verb at the top with its best possible combination at the bottom:</i>								
Run	/	Manage	/	Store	/	Send	/	Call
<i>Files</i>	/	<i>Business</i>	/	<i>Card</i>	/	<i>Keyboard</i>	/	<i>Meeting</i>
<i>Client</i>	/	<i>Information</i>	/	<i>Messages</i>	/	<i>Issue</i>	/	<i>Advertisement</i>

Figura 5: Prueba de *Matching* para comprobar el posible uso beneficioso de los ejercicios en red.

3.1.2.b. *Concordancia, párrafo, texto*

La concordancia electrónica es otro recurso léxico *on-line* que ofrece múltiples posibilidades para la explotación de los datos léxico-gramaticales del corpus. No obstante, hemos observado a lo largo de los años que no es útil poner al alumnado frente a un programa de concordancia simplemente pidiéndole que “examine” el comportamiento de cierta palabra en contexto (una actividad que parece resultar más ventajosa en centros universitarios donde el IFA o Inglés para fines académicos se imparte en un nivel prácticamente bilingüe [cf. Thurstun & Candlin 1998]). Tanto el aspecto poco amigable de las diversas líneas y frases que aparecen en la concordancia como el tipo de pregun-

ta metalingüística que implica esa tarea hacen que el experimento fracase en nuestras aulas, como hemos observado.

Por el contrario, un enfoque más esperanzador para el aprendizaje y la motivación del alumnado se encuentra en lo que hemos denominado “concordancia cerrada”. En ésta, es muy importante elegir bien los textos para que el alumno consiga reunir unas pocas líneas representativas de cómo se utiliza un término característicamente en un corpus dado, donde indagaría mediante un programa como *WordSmith Tools* (Scott 1996) o cualquier otro disponible en red. Además, se darían algunas pistas sobre el campo semántico de dicho término en la concordancia cerrada, es decir, sobre su relación con otros posibles términos, tal y como apuntamos en la parte inferior de la figura 6³.

<p><i>sales _____ by 3 million dollars per year</i> <i>overstate that the “true” prices _____ by around 20 percent per year</i> <i>the costs _____ slowly year by year, leading to higher wages</i> <i>sales _____ by 30%, or by a factor of 1.3</i></p> <p>*Hint= this is a verb commonly used to refer to the expansion of economic activities (sales, buys, costs, prices, etc).</p>

Figura 6: Concordancia cerrada (con información muy concreta sobre uso de término).

Esta variante de ejercicio de concordancia cerrada tiende a tener éxito cuando incluye palabras gramaticales o funcionales (ver figura 7). En este caso, se hace un uso significativo no sólo de datos derivados del corpus, sino también de la competencia lingüística para resolver cuestiones léxico-gramaticales.

<p><i>CLOSED CONCORDANCING: Find two words missing from the lines (they are the same two words in all lines)</i></p> <p>(Same meaning as: regarding, referring to, in relation to)</p> <p><i>components, devices) can best be described in _____ the “bathtub” curve: Very</i> <i>the enterprises who drop out of the sample in _____ economic performance, we</i> <i>performance following their accessions in. _____ catching up, data show that</i> <i>we would expect some delay for effects in _____ improved economic performance</i></p>
--

Figura 7: Ejemplo de concordancia cerrada con elementos gramaticales o funcionales.

La gestión adecuada de párrafos en recursos léxicos electrónicos supone una aproximación importante como posible enlace cognitivo entre la unidad oracional/suboracional de la concordancia y la macroestructura del texto. Los párrafos pueden contener definicio-

nes o ejemplos completos, descripciones e incluso opiniones resumidas de ciertos conceptos e ideas en nuestro corpus especializado (cf. Trimble 1985). Desde nuestra propia experiencia, las habilidades de comprensión del discurso más técnico pueden ejercitarse de manera efectiva mediante la elaboración de actividades apropiadas de enlace entre párrafos y funciones retóricas específicas. Un ejemplo es la figura 8, que ilustra una definición técnica de un programa informático de gestión. La inclusión en el ejercicio de diferentes opciones de funciones retóricas (descripciones, ejemplos, etc) facilita que el alumno realice selecciones pertinentes para formar su idea sobre la unidad textual concerniente, discerniéndola de las demás.

Paralelamente, se podría trabajar con elementos parrafales representativos de un tipo concreto de texto académico y/o temático. En este caso, nos fijaríamos sobre todo en expresiones semitécnicas que nos guían a modo de “palabras nodo” o centrales en el párrafo sometido a escrutinio. A partir de este “vocablo-nodo” de la lista semitécnica, seleccionaríamos párrafos representativos de un tipo de texto, y el trabajo del alumno es elegir la opción adecuada, explicando el porqué de su elección. La figura 9 ofrece un ejemplo de esta actividad. En este caso, el nodo sería el verbo *achieve*, que en sus combinaciones típicas dentro del párrafo podría inducir a la relación de este fragmento de texto con un tipo o género determinado de nuestro corpus. En el ejercicio se proporcionan cuatro opciones, de las que habría que descartar dos.

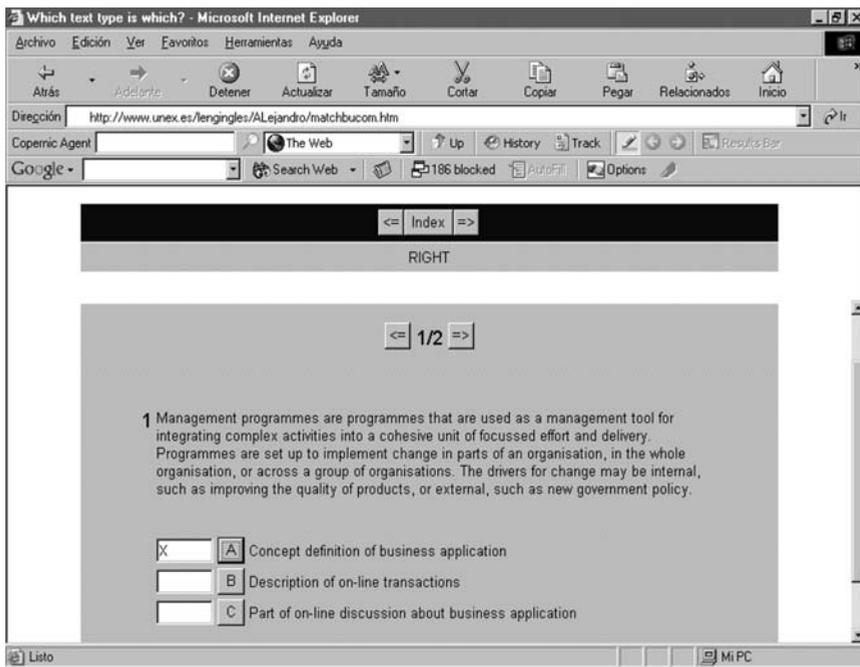


Figura 8: Ejercicio de enlace cognitivo entre párrafo y función retórico-textual en corpus.

In our organisation, we are just in the process of finalising our new 3-year rolling Strategic Plan. Crucial to achieving the objectives set in the Plan will be the implementation of a large number of new projects/initiatives that will have an impact on every part of the organisation.

Which one belongs to a business review? A piece of news? A computer technical report? An oral project discussion?

The Pet Rock, the White Power Rangers, the Beanie Babies and the Furbies were toys that achieved success in different years without coming from the rule-book or the experience database of any single company. Despite their yesterday's success, the producers of such toys are not guaranteed a place in the future that doesn't compute based on yesterday's historical data.

Figura 9: Ejemplo de actividad con párrafos y tipos de textos.

Los alumnos pueden utilizar la herramienta de concordancia electrónica para examinar posibles combinaciones de la(s) palabra(s) subrayada(s), y, para ello, les ofrecemos cuatro o cinco tipos de texto donde aparecería dicho vocablo. Mediante el cotejo de éste en las fuentes, el discente puede observar diferentes usos de las expresiones en distintos tipos o géneros textuales. En el ejercicio de la figura 9, en concreto, cuatro alumnos de tercero de Informática acertaron en su contestación (en los demás grupos, sólo dos). Para sus respuestas correctas (debate oral sobre un proyecto y revisión de un producto empresarial), los alumnos se basaron en indicadores léxicos de tipo temático (*organisation, plan, objectives, impact, success, company, producers*, etc) al igual que retórico/pragmático (*we are just in, every part of the organisation, any single company*).

Finalmente, implementamos también electrónicamente ejercicios con textos cortos pero representativos de un tema. La figura 10 muestra un ejemplo de esta actividad, donde la localización de diferentes usos de un vocablo (como, por ejemplo, el verbo *run*) es uno de los objetivos a seguir. No obstante, antes de llegar a esta observación inductiva, se explotan los diferentes conceptos técnicos, previamente resaltados en los textos (figura 10). Así, preguntas específicas sobre la relevancia de estas nociones en la tecnología empresarial pueden conducir a los alumnos a emplear un lenguaje especializado en sus contestaciones. Posteriormente, una vez aclarada la terminología o parte de ésta, se pide que se busquen repeticiones léxico-gramaticales (por ejemplo, el verbo *run* combinado con diferentes vocablos). Se induciría por tanto a una clasificación de expresiones como *run + risk of, run + test, run + matrix*, etc.

A digital signature contains a hash code derived from the data per se. Any data modification will cause a different hash code that will not match the digital signature. After the digital signature is encrypted within the message, the message is sent to the recipient, first with the sender's private key and then with the receiver's public key. Furthermore, the recipient must decrypt the message first with its private keys and then with the sender's public key. This method ensures that the message can come only from the sender. Unregistered transactions: A business transaction may run the risk of being sent but not received. This risk can be costly if the transaction is in response to a limited-time offer, such as a bid on a government contract.

A common first step of many data analyses that involve more than a very few variables is to run a correlation matrix of all variables and then examine it for expected (and unexpected) significant relations. When this is done, you need to be aware of the general nature of statistical significance (see Elementary Concepts); specifically, if you run many tests (in this case, many correlations), then significant results will be found "surprisingly often" due to pure chance. For example, by definition, a coefficient significant at the .05 level will occur by chance once in every 20 coefficients. There is no "automatic" way to weed out the "true" correlations.

Figura 10: Ejemplo de actividad con textos donde existen temas y léxico clave afín.

3.2. EVALUACIÓN

Como fase posterior a la utilización de los diferentes recursos y actividades electrónicos que llevan a cabo los alumnos, realizamos una "post-evaluación" o "post-test" que permita darnos a conocer resultados tangibles. Para ello, en la línea de Nunan (1989) o Flowerdew (1993), por citar dos ejemplos, proponemos el medio de la tarea comunicativa en entornos académicos como un elemento óptimo de comunicación especializada en el aula. Mediante un trabajo que exija rendir académicamente, los discentes deberían hacer uso de sus conocimientos léxicos, explotando, al mismo tiempo, técnicas como "*inference, linking, skimming, anticipating*" (Nunan 1989: 50).

De ahí que propongamos como tarea evaluativa final un debate corto en grupos de cinco alumnos sobre un tema técnico seleccionado del corpus (y enumerados en la tabla 1 de la Sección 2.1.). En este trabajo oral, los alumnos se enfrentan a un doble reto: tienen que hacer uso de un lenguaje especializado y/o semiespecializado, y, a la vez, de técnicas en su discurso para hacerlo comprensible, real, y convincente; para ello, deberían emplear funciones retóricas que fortalezcan sus ideas (tal y como se sugirió en los aspectos académicos de la Sección 2). Nuestro interés por el debate coincide con investigaciones recientes en Inglés Empresarial sobre *case studies*, que a menudo implican discusiones en el mundo de la empresa (cf. Jackson 1998; Almagro Esteban & Pérez Cañado 2004).

Como paso previo a la exposición del debate, se les pide a los discentes que se preparen bien durante una semana, pudiendo traer a clase, el día de su debate, algunas notas y apuntes, pero en ningún caso apoyo visual y mucho menos texto. Se les pide que hable cada uno un mínimo de dos minutos y entre ellos se hagan preguntas sobre posibles asuntos de interés dentro del tema escogido. En el experimento que hemos llevado a cabo con los tres cursos de IFE (en Informática y Empresariales), los temas elegidos han sido los siguientes: programas multimedia (elegido por un grupo de segundo de Informática y otro de tercero de Empresariales), publicidad en Internet (un grupo de tercero de Informática y el otro de Empresariales), comercio electrónico (un grupo de segundo de Informática), y tecnología audio-visual (un grupo de tercero de Informática). Al igual que en la fase de evaluación previa para la confección de la base de datos léxico-gramatical (en el apartado 3.1.1. anterior), hemos tomado como referencia tres niveles de uso léxico: expresiones efectivas (del corpus), correctas (no halladas en el corpus), y errores o fallos. La tabla 3 resume los resultados registrados de la tarea de debates cortos específicos realizados en los diferentes módulos.

<i>Lenguaje</i>	2º I	3º I	3º E
Efectivo	8	13	5
Correcto	119	204	115
Errores	47	53	52

Tabla 3: Número de expresiones utilizadas en los debates según los diferentes módulos y apartados.

El análisis y la interpretación de estos datos se realizan en consonancia con la finalidad principal de examinar la significación cuantitativa y cualitativa del uso de léxico para tareas comunicativas específicas. Dicho de otro modo, la pregunta es si existe una producción de vocabulario específico significativa tras la exposición del alumnado a las actividades y recursos *on-line*. Para contestar, nos remitimos, sobre todo, a los ejemplos de expresiones efectivas (derivadas del corpus y localizadas en los recursos electrónicos) empleadas por los alumnos. A continuación, comparando la tabla 3 con la tabla 2 (del apartado 3.1.1.), que contiene información previa al uso de los recursos electrónicos, se contrasta el número de construcciones correctas y erróneas en ambas tablas.

Como puede observarse, la categoría que tiene un mayor número de elementos léxicos efectivos corresponde al grupo de tercero de Informática, con un total de 13 expresiones, nueve de las cuales fueron producidas por el subgrupo que habló sobre tecnología audio-visual. Algunos ejemplos de este último subgrupo son *according to the output*, *the graphical output*, *the starting point*, *computer networks*, y *electronic publishing*. Dos de las construcciones restantes, *the ratio of profit* y *the rate of profit*, fueron admitidas como efectivas, pues, aunque no son completamente iguales a las del corpus (*profit ratio* y *profit*

rate), suponen un logro importante por lo insólito de su uso en nuestras clases. Algo similar ocurrió en el subgrupo que debatió sobre publicidad en Internet: la expresión *innovations in sales* fue aceptada como efectiva aunque no fuera exactamente igual a la de los ejercicios, *innovations in the sales*. Finalmente, otras dos expresiones utilizadas, *become fully equipped* y *system performance*, no fueron vistas en las actividades léxicas electrónicas ni son frecuentes en el corpus, pero sí aparecen en dos y tres textos de éste.

En segundo de Informática se emplearon ocho expresiones efectivas vistas en los recursos electrónicos: *computer application*, *computer network* (que se usó en los dos temas de debate), *large enterprise*, *make a lot of money*, *available on the web*, *e-business*, y otra construcción que se acercó, al menos, a una vista previamente (*guarantee of money back* en lugar de *money back guarantee*). En tercero de Empresariales se registró el menor número de elementos de este tipo (cinco): *domestic market*, *network marketing*, *large size*, y, otras dos –*expansion of the business* y *growth of the business*– que se consideran análogas a los ejemplos de las actividades on-line (*business expansion* y *business growth*).

En cuanto a los rasgos lingüísticos correctos y fallidos cotejados en la tabla 3, observamos que, en comparación con la tabla 2, los alumnos de tercero de Informática pasan a ocupar el primer lugar, y los de segundo de Informática, el segundo (al contrario de lo que se refleja en la tabla 2). Esto ocurre porque los debates en tercero de Informática duran dos minutos más aproximadamente que en los demás grupos. Además, por su fluidez comunicativa, los alumnos de tercero de Informática tienden a tener un diálogo más dinámico que los restantes discentes. Esta mayor capacidad en los debates podría comprobarse en su propia producción de elementos léxicos efectivos y gran cantidad de formaciones léxico-gramaticales correctas. El superior número de errores también implicaría una elaboración del discurso más ágil, con una correspondiente búsqueda rápida de palabras de contenido, lo cual llevaría a cometer fallos, según Garman [1990: 137], algo normal en el procesamiento mental léxico.

En contraste, la tabla 2 sugiere sobre todo la gran capacidad comunicativa de los alumnos de segundo de Informática (con el mayor número de expresiones correctas y errores). En su caso, las tareas orales de presentaciones técnicas no fueron significativamente más largas que las de los alumnos de tercero de Informática (aunque sí que los de Empresariales), pero escribieron composiciones extensas. En cualquier caso, su notable producción de vocabulario efectivo (tabla 3) podría indicar que su competencia léxica se ha beneficiado del acceso y explotación de los recursos electrónicos. No ocurre lo mismo con el grupo de Empresariales, que, al contrario de lo que indican sus datos iniciales (tabla 2), su dominio de unidades léxicas efectivas empeora con respecto a los otros grupos (tabla 3)⁴.

4. CONCLUSIONES

La conclusión general de nuestra investigación puede resumirse en dos enunciados principales: según los resultados obtenidos en nuestros estudios de casos concretos de aprendizaje en IFE, la elaboración de recursos léxicos a los que los discentes accederían

de forma autónoma (*on-line*) debería apoyarse, de un lado, en datos empíricos sobre el lenguaje típico del área de especialización (representado en un corpus), y de otro, en una introspección lingüística del profesor/investigador derivada tanto del análisis de esos datos léxico-gramaticales como de la producción lingüística del alumnado en tareas comunicativas.

Desde nuestro punto de vista, la gradación del material y los recursos electrónicos fundamentada en datos de corpus y tareas de clase resulta fundamental para poder adecuar dichos ejercicios a las necesidades académicas y/o profesionales del alumnado. En este sentido, resulta esencial diseñar un corpus de lenguaje especializado con unos criterios que se determinen a partir de esas necesidades en el contexto comunicativo de aprendizaje e institucional (características académicas y elementos socio-culturales o socio-lingüísticos). Así, las dos fases esenciales en el examen de esas demandas en el aula han sido, a lo largo de este trabajo, proporcionar material basándonos en datos léxicos de corpus, y, a continuación, cotejar los elementos lingüísticos en tareas comunicativas específicas.

Se ha trabajado con actividades *on-line* en dos niveles: el sintagmático/léxico-gramatical (*gap-filling, matching*), y el paradigmático/textual (datos de concordancias, párrafos y textos *on-line*). Para valorar la posible utilidad de estas herramientas y ejercicios en red para la formación léxico-gramatical de los alumnos, nos hemos apoyado en la evaluación de tres estratos lingüísticos: formaciones léxicas efectivas, expresiones correctas, y lenguaje fallido. El examen de estos resultados se realizó observando la producción lingüística en una tarea de debate oral sobre algún tema del corpus, pues pensamos que el desarrollo de esta clase de tarea engloba la competencia léxica, de contenido, y habilidades retórico-discursivas además de pragmáticas del alumnado. En consecuencia, se valoran tres resultados principales del experimento.

Primero, que los alumnos de tercero de Informática demuestran un mayor conocimiento léxico del corpus y, paralelamente, reflejan una mejor competencia comunicativa en su estructuración de los debates (más largos y fluidos). Podría decirse que llevaron a cabo un procesamiento más efectivo del léxico en el contexto específico de uso. Seguidamente, que los alumnos en segundo de Informática esgrimen un vocabulario importante, manifestando un manejo óptimo de los recursos léxicos, con datos generales (en diferentes tareas, no sólo debates) que prueban que su competencia lingüística es favorable para el desarrollo comunicativo. Finalmente, que los alumnos de tercero de Empresariales producen un inferior número de expresiones efectivas y se mantienen en unos datos lingüísticos más o menos constantes, sin demostrar un efecto positivo (aunque tampoco negativo) en su uso de los recursos electrónicos.

5. NOTAS

1. En Curado (en prensa), se lleva a cabo este tipo de análisis léxico-gramatical con 500 términos de la lista.
2. En un estudio en curso, y siguiendo en gran parte las ideas de Sinclair (2004) sobre el componente léxico-gramatical, nos centramos en la evaluación de errores gramaticales como elementos inseparables del léxico en contexto. Nuestra hipótesis se contrasta con la producción lingüística de los alumnos en foros electróni-

- cos sobre temas generales de la ciencia y tecnología (p.ej., *Chemical weapons* o *The origin of the universe*), utilizando información léxico-gramatical del corpus (p.ej., elementos que usan la voz pasiva, modales, pronombres, etc).
3. En los primeros cursos de Informática y Empresariales, utilizamos concordancias “abiertas” de corpus especializados en Internet (con categorías como *Business* y *Computing*), pidiendo a los alumnos que realicen búsquedas de términos que van con otros, algo mucho más mecánico y que tendría el objetivo principal de familiarizar al alumno de primero con estos recursos.
 4. En este sentido, nuestros datos podrían corroborar resultados de investigaciones recientes (cf. Fuertes Olivera & Gómez Martínez 2004), que aseguran que características muy particulares entre los alumnos de Empresariales tienden a desmotivarlos en los idiomas (algunas causas son su poca asistencia a clase, preferencia por charla magistral, tópicos en torno a “saber inglés”, etc).

6. BIBLIOGRAFÍA

- Almagro Esteban, A. y M.L. Pérez Cañado. 2004. “Making the Case Method Work in Teaching Business English: A Case Study”. *English for Specific Purposes* 23 (2): 137-161.
- Aston, G. 1997. “Enriching the Learning Environment: Corpora in ELT”. Eds. A. Wichmann et al. *Teaching and Language Corpora*. Londres: Longman. 51-67.
- Cheng, W. et al. 2003. “The Language Learner as Language Researcher: Putting Corpus Linguistics on the Timetable”. *System* 31 (2): 173-186.
- Cobb, T. 2002. *Cloze Builder 1.0*. Montreal: Université du Québec. http://www.er.uqam.ca/nobel/r21270/cgi-bin/clozefreq/cloze_enter.html
- Conrad, S. 1996. “Investigating Academic Texts with Corpus-based Techniques: An Example from Biology”. *Linguistics and Education* 8: 299-326.
- Cook, G. 2001. *The Discourse of Advertising*. Londres: Routledge.
- Coxhead, A. 2002. *The Academic Wordlist Gap Maker*. Nottingham: University of Nottingham. <http://www.nottingham.ac.uk/~alzsh3/acvocab/awlgapmaker.htm>
- Curado, A. 2002. “Exploitation and Assessment of a Business English Corpus through Language Learning Tasks”. *ICAME Journal: Computers in English Linguistics* 26: 5-32.
- Curado, A. 2003. “The Use of Corpora and IT in a Comparative Evaluation Approach for Business English Oral Reports”. *ReCALL Journal* 15 (2): 189-201.
- Curado, A. 2004a. “The Use of Corpora and IT in Evaluating Oral Task Competence for Tourism English”. *CALICO Journal* 22 (1): 5-24.
- Curado, A. 2004b. “ESP Activities”. Universidad de Extremadura. <http://www.unex.es/lengingles/Alejandro> y <http://158.49.123.142/admin.php>.
- Curado, A. En prensa. *El inglés de la empresa y la tecnología*. Cáceres: Servicio de Publicaciones de la Universidad de Extremadura.
- Dudley-Evans, T. y M.J. St. John. 1998. *Developments in ESP. A Multidisciplinary Approach*. Cambridge: Cambridge University Press.
- Flowerdew, J. 1993. “Content-Based Language Instruction in a Tertiary Setting”. *English for Specific Purposes* 12: 121-138.

- Flowerdew, L. 2001. "The Exploitation of Small Learner Corpora in EAP Materials Design". Eds. M. Ghadessy *et al.* *Small Corpus Studies and ELT. Studies in Corpus Linguistics*. Amsterdam: John Benjamins. 363-380.
- Fortanet, I. 2004. "The Use of 'We' in University Lectures: Reference and Function". *English for Specific Purposes* 23 (1): 45-66.
- Fuertes Olivera, P.A. y S. Gómez Martínez. 2004. "Empirical Assessment of Some Learning Factors Affecting Spanish Students of Business English". *English for Specific Purposes* 23 (2): 163-180.
- Garman, M. 1990. *Psycholinguistics*. Cambridge: Cambridge University Press.
- Half-baked software inc. 2003. *Hot Potatoes 6.0*. Victoria: University of Victoria. The Humanities Computing and Media Centre. <http://web.uvic.ca/hrd/halfbaked/>
- Jackson, J. 1998. "Reality-based Decision. Cases in ESP Teacher Education: Windows on Practice". *English for Specific Purposes* 17 (2): 151-167.
- James, G. y J. Purchase. 1996. *English in Business Studies and Economics*. Hong Kong: Hong Kong University of Science and Technology.
- Johns, T. 1991. "Should you be Persuaded: Two Examples of data-driven Learning". Classroom Concordancing. *English Language Research Journal* 4: 1-16.
- Jolliffe, D. y E. Brier. 1988. "Studying Writer's Knowledge in Academic Disciplines". Ed. D.A. Jolliffe. *Advances in Writing Research v.2: Writing in Academic Disciplines*. Norwood, NJ: ALEX. 29-55.
- Jordan, R.R. 1997. *English for Academic Purposes*. Cambridge: Cambridge University Press
- Nunan, D. 1989. *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.
- Leech, G. 1991. "The State of the Art in Corpus Linguistics". Eds. K. Aijmer y B. Altenberg. *English Corpus Linguistics*. Londres: Longman. 8-29.
- Ooi, V.B.Y. 1998. *Computer Corpus Lexicography*. Edimburgo: Edinburgh University Press.
- Scott, M. 1996. *WordSmith Tools 2.0*. Oxford: Oxford University Press.
- Scott, M. 1997. "PC Analysis of Key Words and Key Key Words". *System* 25 (1): 1-13.
- Sinclair, J. 2004. *Trust the Text. Language, Corpus and Discourse*. Londres: Routledge.
- Smith, B. 2003. "The Use of Communication Strategies in CMC". *System* 31 (1): 29-53.
- Thurston, J. y C.N. Candlin. 1998. "Concordancing and the Teaching of the Vocabulary of Academic English". *English for Specific Purposes* 17: 20-34.
- Trimble, L. 1985. *English for Science and Technology: A Discourse Approach*. Cambridge: Cambridge University Press.