

Vocabulario académico de artículos de investigación de ciencias agrarias:

Resultados preliminares de un estudio de corpus

Iliana Martínez – Silvia Beck

Universidad Nacional de Río Cuarto

Introducción

El estudio del vocabulario especializado utilizado en textos académicos en inglés ha sido objeto de numerosas investigaciones recientemente debido a la creciente demanda de instrucción por parte de los hablantes nativos de inglés en diferentes contextos. Debido al enorme tamaño del vocabulario de una lengua y para maximizar la efectividad de su enseñanza, los investigadores han producido listas de palabras del vocabulario recurrente en textos académicos. Se considera que estas listas proveen el vocabulario necesario para funcionar en contextos académicos (Coxhead & Nation, 2001). La más reciente de estas listas es la Lista de Palabras Académicas (LPA) de Coxhead (2000), la cual consiste de 570 familias de palabras y 3,107 formas (types). Para construir la lista, Coxhead utilizó como punto de partida la Lista de Palabras Generales (LPG) realizada por West (1953). La Lista de Palabras Generales comprende las palabras más frecuentes y ampliamente utilizadas del lenguaje, incluyendo palabras gramaticales. Coxhead construyó su lista excluyendo las palabras de la lista general de West, e incluyendo aquellas palabras que aparecían frecuentemente y con cierto grado de uniformidad en un amplio rango de disciplinas y géneros en cuatro áreas, artes, comercio, leyes, y ciencias. La lista que resultó de su estudio, la Lista de Palabras Académicas, proveía una cobertura de alrededor de 8.5% a 10% de las palabras de un texto académico y se esperaba que combinadas con las palabras de la Lista General de West, proveyeran una cobertura de alrededor de 90% de cualquier texto académico.

A pesar de esta importante cobertura, la eficiencia de la LPA como instrumento de enseñanza de vocabulario académico en cursos con fines específicos ha sido recientemente cuestionada, demostrándose que las diferencias lexicales que existen entre las disciplinas pueden ser mayores que las similitudes. En relación a esto, Hilad (2002; 2006) señala la necesidad de especificidad en los cursos de inglés con fines específicos, enfatizando las complejidades de la comunicación disciplinar y la necesidad de relacionar la enseñanza lingüística y retórica con los usos, convenciones, valores y expectativas de las comunidades disciplinares específicas en donde estos cursos son funcionales. El argumento de Hilad está sustentado por otras investigaciones que han demostrado que cada disciplina tiene sus propias formas de explicar las experiencias y sus propias formas de argumentar (e.g., Bloor, 1998; Holmes, 1997; Hyland, 1999; 2000; 2001; Samraj, 2002).

Además, estudios recientes en vocabulario académico utilizando diferentes corpora (Chen & Ge, 2007; Hyland & Tse, 2007; Paquot, 2007), cuestionan la utilidad de la LPA en cursos específicos argumentando que, dada la generalidad de la lista, puede resultar por una parte, en la exposición a más vocabulario del que los alumnos necesitan, y por otra parte en la falta de exposición al vocabulario que efectivamente necesitarán en sus disciplinas.

Los alumnos para quienes la especificidad es más relevante son probablemente los graduados e investigadores de contextos en donde el inglés es una lengua extranjera, ya que en esos contextos la necesidad de publicar investigación en inglés es tan importante como en otros contextos. En nuestra universidad latinoamericana, los investigadores necesitan publicar en revistas científicas internacionales con referato tanto para avanzar en sus carreras como para obtener apoyo financiero. En las ciencias agropecuarias esto se ha convertido en un problema ya que las posibilidades de publicar en español en revistas

prestigiosas indexadas se han tornado dificultosas debido a la tendencia de las revistas plurilingües de Europa y Japón a optar por el inglés como única lengua de publicación (Hyland, 2006). Por lo tanto, revistas científicas de agricultura que solían publicar en español o en español e inglés, ahora utilizan sólo inglés.

Los investigadores y graduados que demandan instrucción para la escritura en nuestra universidad y probablemente la mayoría de los hablantes no nativos de inglés que necesitan publicar en inglés obtienen la mayor parte de su información de bibliografía en inglés. A pesar de que su manejo del lenguaje suele ser pobre, los participantes en nuestros cursos pueden ser descriptos como lectores eficientes de artículos de investigación en inglés en sus disciplinas, y pueden por lo tanto aportar su vasta experiencia en lectura a la escritura en inglés. Debido a que estos investigadores no tienen el tiempo necesario para dedicar al aprendizaje de inglés por las actividades propias de sus carreras, es importante que los docentes los ayuden a obtener el máximo provecho de su tiempo en el curso de inglés. En nuestro contexto, tenemos la ventaja de que, debido a la gran demanda de instrucción en escritura, podemos tener cursos homogéneos respecto a áreas disciplinares (Martínez, 2002). Otra ventaja es que los alumnos comparten el lenguaje oficial de educación, español. Esta homogeneidad no es común en los contextos de inglés como segunda lengua en donde los cursos de inglés con fines específicos se dictan a alumnos de diferentes disciplinas y de diferentes lenguas maternas (Bloor, 1998; Dudley-Evans, 2002; Swales and Lindemann, 2002). Nuestra situación, por lo tanto justifica la investigación que provea descripciones más específicas del vocabulario académico que debe ser enseñado en contextos específicos.

En este trabajo, presentamos resultados parciales de un estudio mayor sobre vocabulario académico en un género específico, y una disciplina específica. En particular, este estudio reporta aspectos de frecuencia, distribución, y significado de palabras de la Lista de Palabras Académicas en un género en particular, el artículo de investigación, de un campo específico, las ciencias agrarias. Se provee una selección de ítems de la LPA que serán específicamente útiles para escribir artículos de investigación en ciencias de la agricultura. Se espera que estos resultados contribuyan al argumento en favor de la especificidad en cursos de inglés con fines específicos.

1. El estudio

Este estudio integra enfoques basados en corpus y en género (Flowerdew, 2005), analizando el artículo de investigación para revelar características específicas del vocabulario académico, utilizando como base la Lista de Palabras Académicas. Este estudio basado en corpus (Scott & Tribble, 2006) reporta acerca de la frecuencia, cobertura, y distribución de las palabras del vocabulario académico en artículos de investigación de agricultura, tanto en el artículo completo como en sus secciones.

2.1. *El corpus*

Para el estudio se construyó un corpus de artículos de investigación de ciencias agrarias de 826 416 palabras (AgroCorpus). El corpus se diseñó siguiendo los criterios propuestos por Sinclair (1991, 2005) y Barnbrook (1996), considerando representatividad, especificidad del corpus, uso de documentos completos, y disponibilidad en forma electrónica. El corpus es representativo de un género, el artículo de investigación experimental (Swales, 1990), y de una disciplina, agricultura. Los artículos se obtuvieron de las versiones on-line de journals indexados por Science Citation Index (SCI) Report. El corpus contiene 218 artículos producidos en universidades de habla inglesa y fueron

seleccionados de revistas científicas publicadas entre 2000 y 2003, específicamente recomendadas por especialistas de nuestra universidad. Los artículos tienen el formato de Introducción-Método- Resultados-Discusión (IMRD, Swales, 1990). Los documentos se prepararon para poder ser analizados como textos completos o como secciones individuales (sub-corpora). Se excluyeron los números, resúmenes, agradecimientos, referencias, leyendas y apéndices del conteo de palabras.

Las unidades de análisis utilizadas fueron formas (types), casos (tokens) y familias. Las formas (types) se refieren a cada palabra; casos (tokens) a la cantidad de repeticiones de cada forma; y familias al grupo de formas de palabras formal y semánticamente relacionadas (Bauer & Nation, 1993, in Chung & Nation, 2003).

2.2 El software utilizado para el análisis

Para el análisis, se utilizó el software computacional Wordsmith Tools (WST) (Scott, 2004), que comprende un grupo de programas para el análisis del comportamiento de palabras en textos. En este estudio se utilizó la herramienta Wordlist, la cual genera listas de palabras en orden alfabético y en orden de frecuencia que sirven para la comparación léxica de textos. Entre sus funciones, esta herramienta permite la comparación entre listas de palabras (Scott).

2.3 Procesamiento y análisis de datos

Para este estudio, primero se determinaron la frecuencia y distribución de formas de palabras (types) y casos (tokens) en el corpus. Luego utilizando la LPG y la LPA como lista de coincidencia (matchlists), se identificaron las palabras generales y las académicas en el AgroCorpus y su cobertura. Debido a que el foco de nuestra investigación eran las palabras de la Lista de Palabras Académicas, y que la mayoría de los estudios de palabras académicas ha utilizado las familias de palabras como unidad de análisis, construimos las familias con las palabras más frecuentes identificadas en el AgroCorpus. Para determinar las familias más frecuentes, utilizamos el criterio de Hyland y Tse, (2007), quienes consideraron como frecuentes aquellas palabras que aparecían por sobre la media del total de palabras académicas identificadas. Se determinaron las medias tanto para el texto completo como para las secciones individuales, y se construyeron las familias con aquellos ítems que estaban por encima de la media.

2. Resultados

En este trabajo se presentan resultados parciales. Se investigó la frecuencia, cobertura y distribución de las palabras académicas en el AgroCorpus. El resultado más importante es una lista de palabras frecuentes de la LPA en el corpus, que es aún más restringida que las listas de Hyland y Tse (2007) y Chen y Ge (2007), que puede ser utilizada para responder a las necesidades específicas de los escritores de artículos de investigación en ciencias agrarias.

3.1 Frecuencias de palabras en el AgroCorpus

La distribución general de casos (tokens) en el AgroCorpus y su distribución en las secciones aparecen en la Tabla 1. Los ítems se agruparon en palabras generales, palabras académicas, y otros casos. Estos incluían principalmente palabras técnicas, pero también

otras palabras como fórmulas y nombres propios. La LPG (67.53%) y la LPA (9.06%) proveían una cobertura acumulativa del 76.59%.

Tabla 1: cobertura de los ítems lexicales el AgroCorpus (números crudos y porcentajes)

	Introducción	Método	Resultados	Discusión	Total
Casos de la LPG	83 082 65.71% (14.89%)	160 196 65.36% (28.70%)	145 117 68.99% (26.00%)	169 694 69.39% (30.41%)	558 089 67.53% (100%)
Casos de la LPA	11 966 9.46% (15.99%)	21 294 8.69% (28.45%)	18 548 8.82% (24.79%)	23 030 9.42% (30.77%)	74 838 9.06% (100%)
Otros casos	31 389 24.83% (16.22%)	63 605 25.95% (32.87%)	46 664 22.19% (24.12%)	51 831 21.19% (26.79%)	193 489 23.41% (100%)
Casos totales	126 437 100% (15.30%)	245 095 100% (29.66%)	210 329 100% (25.45%)	244 555 100% (29.59%)	826 416 100% (100%)

En cuanto a las formas (types) utilizadas, el total en el corpus es de 23 682 con una distribución diferente en las secciones, que revela menor variabilidad en la sección Resultados y mayor en los Métodos del artículo de investigación de agricultura (Tabla 2). El subgrupo de formas (types) de palabras de la LPA reveló que del total de 3107 formas de la LPA, sólo se encontraron 1941 en el AgroCorpus, lo cual significa que 1166 ítems (37.50%) no aparecían. En el AgroCorpus, el número más bajo de formas (types) de palabras académicas se observó en los Resultados y el más alto en la Discusión. Esta diferencia puede ser atribuida al carácter argumentativo de la sección discusión de los artículos de investigación (Martínez, 2003).

Tabla 2: Distribución de las formas (types) de palabras en el AgroCorpus.

	SECCIONES				TEXTO
	Introducción	Métodos	Resultados	Discusión	COMPLETO
Formas de la LPG	2630	3008	2568	3061	4009
Formas de la LPA	1312	1283	1160	1566	1941
Otras formas	7389	9929	5416	8320	17 732
Total de formas	11 331	14 220	9144	12 947	23 682

3.2 Comparación de frecuencias con otros estudios

La comparación de la cobertura de la LPA, y de la LPA combinada con la cobertura de la LPG con estudios previos reveló diferencias que resaltan la especificidad del AgroCorpus. La cobertura de la LPA en nuestro corpus, 9.06%, fue menor que la encontrada por Coxhead (2000) y por Hyland y Tse (2007) en sus corpora multidisciplinarios (Tabla 3) y también por Chen y Ge (2007) en su corpus de medicina. Es importante resaltar que cuando Hyland y Tse focalizaron en disciplinas específicas en su corpus, encontraron marcadas diferencias en cuanto a la cobertura de palabras académicas, con una baja cobertura, 6.2%, en biología, y una alta cobertura, 16%, en ciencias de la computación. La cobertura de la LPA y de la LPG combinadas, 76.59%, fue menor que la cobertura encontrada por Coxhead, y por Hyland y Tse en sus corpora multidisciplinarios, pero similar a la cobertura en el subcorpus de ciencias de Hyland y Tse (Tabla 3). Esta variación en cobertura contribuye al argumento de Hyland y Tse en favor de que corpora altamente específicos se benefician de manera diferente con la LPA.

Tabla 3: Comparación de la cobertura de la GSL y de la AWL en el AgroCorpus y en otros corpora

	AgroCorpus	Corpus de Hyland y Tse (2007)	Subcorpus de Ciencias de Hyland y Tse (2007)	Corpus de Coxhead (2000)
LPG	67.53%	74%	69%	76.1%
LPA	9.06%	10.6%	9.3%	10%
LPG+LPA	76.59%	84.7%	78.3%	86.1%

3.3 Frecuencia de familias de palabras

Se identificaron las familias de palabras en el corpus siguiendo el criterio de frecuencia de Hyland y Tse (2007) descrito en el Método. Por lo tanto, primero se identificó la media para las palabras académicas, tanto en el texto completo, 121, como en las secciones individuales, 129 para la Introducción, 85 para el Método, 69 para los Resultados, y 122 para la Discusión. Luego se construyeron las familias con aquellos ítems que estaban por encima de la media. El número de familias por encima de la media para todo el corpus fue de 92, un número considerablemente menor que el obtenido por Hyland y Tse (2007), quienes identificaron 192 familias frecuentes en su corpus. Las secciones individuales del corpus difirieron en el número de familias por encima de la media. Se identificaron 102 familias en la Introducción, 69 en el Método, 58 en los Resultados, y 94 en la Discusión

3.4. Comparación de familias de palabras con otros estudios

La comparación de los ítems más frecuentes en nuestro corpus con los resultados encontrados por Coxhead (2000), Hyland y Tse (2007), y Chen y Ge (2007) revelan una baja coincidencia, lo que refuerza la idea de que las diferencias se basan en la especificidad del corpus.

Comparamos las 92 familias frecuentes del AgroCorpus con las de la sub-lista 1 de Coxhead, que incluye las 60 familias más frecuentes de la LPA. La Tabla 4 muestra las 92 familias en nuestra lista con las primeras 60 sombreadas. Las palabras que coinciden con los ítems de la sublista de Coxhead están en negrita. Dentro de las primeras 60 familias, sólo 26 coinciden con las de Coxhead, con nueve ítems menos que los observados por Hyland y Tse (2007) en su corpus. Más aun, cuando se consideran todas las familias de palabras frecuentes en nuestro corpus, se observa que solo 33 pertenecen a la sublista 1 de Coxhead, un número que es aun menor que las 35 familias que coincidían en la comparación de Hyland y Tse (2007) entre los 60 ítems más frecuentes en su corpus con los de la sublista 1 de Coxhead. Estos resultados demuestran que mientras más específico es el corpus, mayor es la especificidad de los ítems, y consecuentemente, la variabilidad es menor.

Tabla 4: Familias de palabras académicas frecuentes en el AgroCorpus y en la Sublista 1 de Coxhead

1	significant	32	affect	63	Volume
2	analysis	33	individual	64	Layer
3	Data	34	isolated	65	Source
4	Site	35	components	66	Constant
5	Area	36	maximum	67	Contrast
6	variation	37	approximately	68	Overall
7	response	38	research	69	Transformed
8	similar	39	percentage	70	Role
9	sequence	40	emergence	71	Evidence
10	environments	41	conducted	72	Involved

11	concentration	42	prior	73	Exposure
12	indicated	43	identified	74	Chemical
13	estimated	44	function	75	Bulk
14	region	45	available	76	Location
15	method	46	detected	77	Maintained
16	period	47	selected	78	Demonstrated
17	interaction	48	ratio	79	Negative
18	induced	49	major	80	Approach
19	medium	50	promoter	81	Proportion
20	previous	51	whereas	82	Minimum
21	specific	52	removed	83	Corresponding
22	range	53	design	84	Phase
23	factors	54	distribution	85	Statistical
24	culture	55	required	86	Survival
25	stress	56	duration	87	Terminal
26	accumulation	57	consistent	88	Transferred
27	occurred	58	parameters	89	Evaluated
28	maturity	59	derived	90	Assessment
29	potential	60	incidence	91	Procedure
30	obtained	61	index	92	Adult
31	initial	62	final		

a Las palabras en negrita coinciden en ambas listas

Cuando comparamos las diez familias más frecuentes en nuestro corpus con los resultados de Hyland y Tse (2007) en su sub-corpus de ciencias, encontramos que solo tres ítems coinciden: *data*, *analyze*, y *similar*.

También se compararon las 20 familias más frecuentes en las cuatro secciones del AgroCorpus con las encontradas por Chen y Ge (2007) en un corpus de medicina (Tabla 5). Nuevamente hubo baja coincidencia, sobretodo en la Introducción (25%) y en los Métodos (30%). La más alta coincidencia se encontró en la Discusión (50%), mientras que en los Resultados hubo una coincidencia del 40%. Esto permite sugerir que, al menos en estas dos ciencias, agricultura y medicina, hay más similitud de vocabulario de la LPA en los Resultados y en la Discusión que en la Introducción y en los Métodos. En ambos casos, las palabras que coincidían eran claramente académicas, pero es necesario mencionar que muchas de aquellas palabras de la LPA frecuentes en nuestro corpus que no coincidían con los estudios mencionados tienen usos técnicos en el campo de la agricultura, tales como *site*, *area*, *environments*, *stress*, *region*, *culture*.

Tabla 5: Familias de palabras académicas más frecuentes en las secciones del AgroCorpus y del corpus de medicina^a

	INTRODUCTION	METHODS	RESULTS	DISCUSSION
1	response	analysis	significantly	similar
2	area	data	data	significant
3	environmental	site	similar	response
4	stress	area	sites	sites
5	variation	method	sequence	data
6	interactions	medium	response	area
7	region	estimated	analysis	environments
8	factors	extracted	concentration	indicate
9	potential	conducted	indicated	stress
10	available	culture	range	accumulation
11	induced	sequence	period	induced
12	specific	variance	variation	sequence
13	methods	design	detected	factors
14	significant	concentration	percentage	occur
15	sequence	obtained	accumulation	analysis

16	accumulation	selected	area	variation
17	range	isolated	occurred	previous
18	data	removed	consistent	available
19	occur	prior	estimated	region
20	component	random	maximum	concentration

a Las palabras en negrita ocurren en ambas listas

4. Conclusión

Los resultados de este estudio aportan información específica acerca de aspectos del vocabulario académico de un género y disciplina particular, representado en una lista de palabras de la LPA altamente restrictiva. Nuestra lista reducida de familias frecuentes de la LPA claramente demuestra que es necesario construir listas de frecuencia a partir de los textos de los posibles usuarios, tal como lo sugieren Hyland y Tse (2007). Las listas específicas serán más eficientes con relación al propósito de las listas de palabras, el cual es ofrecer a los alumnos una lista de las palabras que encontrarán con frecuencia en los textos que ellos leen y así se podrá reducir su esfuerzo de aprendizaje (Nation and Waring, 1997).

Nuestros resultados apoyan el argumento de que el vocabulario debe ser enseñado considerando el contexto específico de los alumnos, (Hyland & Tse, 2007, p. 251). En nuestro contexto latinoamericano, con clases homogéneas respecto a lengua materna, necesidades altamente específicas y conocimientos científicos comunes, la especificidad brinda mayor aprovechamiento en el poco tiempo que los investigadores pueden disponer y puede ayudar a compensar el conocimiento usualmente limitado del inglés que tienen los alumnos. La concentración en el vocabulario académico que incluya una lista de palabras reducida pero de gran cobertura, posibilita que los alumnos contribuyan el vasto conocimiento de sus campos específicos. De este modo, tanto la motivación como la autoestima de los alumnos pueden ser incrementadas, ya que los alumnos estarán expuestos a ítems lexicales con los cuales están familiarizados y que pueden reconocer como parte de los textos que ellos manipulan.

Referencias

- Barnbrook, G. (1996). *Language and computers: a practical introduction to the computer analysis of language*. Edinburgh: Edinburgh University Press.
- Bloor, M. (1998). Variations in the methods sections of research articles across disciplines: The case of fast and slow text. In P. Thompson (Ed.), *Issues in EAP writing, research and instruction* (pp. 84-106). Reading, UK: CALS, The University of Reading.
- Chen, Q. & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL Word families in medical research articles (RAs). *English for Specific Purposes*, 26:502-514.
- Chung, T. & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, 15(2): 103-116.
- Coxhead, A. & P. Nation. (2001). The specialised vocabulary of English for academic purposes. In Flowerdew, J. & M. Peacock (Eds.). *Research Perspectives on English for Academic Purposes* (pp. 252-267). Cambridge University Press.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34 (2): 213-238.
- Dudley-Evans, T. (2002). The Teaching of Academic Essay: Is a Genre Approach Possible?. In Johns, A. (Ed.), 2002, *Genres in the Classroom. Multiple Perspectives*. (pp.225-235). Mahwah, N J. Lawrence Erlbaum Associates, Publishers.

- Flowerdew, L. (2005) An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticism against corpus-based methodologies. *English for Specific Purposes*, 24:321-332
- Holmes, R. (1997). Genre analysis, and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes*, 16 (4): 321-337.
- Hyland, K. & Tse, P. (2007). Is there an academic vocabulary? *TESOL Quarterly*, 41(2): 235-253.
- Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20 (3), 341-367.
- Hyland, K. (2000). Speaking as an insider: promotion and credibility in abstracts. In *Disciplinary Discourses. Social Interaction in Academic Writing* (pp. 63-84). London: Pearson Education.
- Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20, 207-226.
- Hyland, K. (2002). Specificity revisited: How far should we go now? *English for Specific Purposes*, 21: 385-395
- Hyland, K. (2006). *English for Academic Purposes. An advanced resource book*. London: Routledge.
- Martinez, I. (2002). Developing genre awareness in nonnative-English-speaking writers of experimental research articles: A collaborative approach. In Cradall, JoAnn and Kaufman, Dorit (Eds.) *Content-Based Instruction in Higher Education Settings. Case Studies in TESOL Practice Series*. Jill Burton, Series Editor. Virginia: TESOL Pp. 79/92.
- Martinez, I. (2003). Aspects of theme in the method and discussion sections of biology journal articles in English. *Journal of English for Academic Purposes*. 2: 103–123.
- Nation, P. & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: CUP
- Paquot, M. (2007). Towards a productively-oriented academic word list. In J. Walinski, K. Kredens, & S. Gozdz-Roszkowski (Eds.), *Corpora and ICT in language studies. PALC 2005. Lodz studies in LANGUAGE 13* (pp. 127–140). Frankfurt am main: Peter Lang.
- Samraj, B. (2002). Disciplinary variation in abstracts: The case of wildlife behaviour and conservation biology. In Flowerdew, J. (Ed.). *Academic discourse* (pp. 40-56). London: Longman.
- Scott, M. & Tribble, C. (2006). *Textual Patterns*. Amsterdam: Benjamins.
- Scott, M. (2004). *WordSmith Tools*. Oxford: OUP.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: OUP.
- Sinclair, J. (2005). Corpus and Text - Basic Principles. In M. Wynne (Ed.), *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow Books: 1-16. Available online from <http://ahds.ac.uk/linguistic-corpora/> [Accessed 2006-07-26].
- Swales, J. & Lindemann, S. (2002). Teaching the Literature Review to International Graduate Students. In Ann M. Johns (Ed.), *Genres in the Classroom. Multiple Perspectives*. (pp. 105/119) Mahwah, New Jersey. Lawrence Erlbaum Associates, Publishers.
- West, M. (1953). *A general service list of English words*. London: Longman.