



Estudios Filológicos

ISSN: 0071-1713

efil@uach.cl

Universidad Austral de Chile

Chile

Parodi, Giovanni

Discurso especializado y lengua escrita: foco y variación
Estudios Filológicos, núm. 41, septiembre, 2006, pp. 165-204
Universidad Austral de Chile
Valdivia, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=173414185012>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

The screenshot shows the header of the website with links for 'Inicio Web Revistas', 'Web Biblioteca', and 'Contacto'. Below the header, there's a banner for 'Revistas Electrónicas UACH'. A navigation bar below the banner includes 'Artículos' (with sub-links 'Tabla de contenido', 'Anterior', 'Próximo'), 'Búsqueda artículos' (with sub-links 'Autor', 'Materia', 'Búsqueda'), and 'Inicio' and 'Lista'.

Estudios filológicos

ISSN 0071-1713 versión impresa

Como citar este artículo
Agregar a favoritos
Enviar a e-mail
Imprimir HTML

Estud. filol. n.41 Valdivia sep. 2006

ESTUDIOS FILOGICOS 41: 165-204, 2006

Discurso especializado y lengua escrita: foco y variación *

Specialized discourse and written language: focus and variation

Giovanni Parodi

Pontificia Universidad Católica de Valparaíso, Programa de Postgrado en Lingüística. Avenida Brasil 2830, Valparaíso, Chile. gparodi@ucv.cl

Este artículo explora la variación entre registros escritos y orales así como entre especializados y no especializados. Siguiendo el método propuesto por Biber (1988) para el estudio de la variación entre oralidad y escritura, se indagan estas y otras cuestiones pero para la lengua española y con un énfasis en una perspectiva textual. Para cumplir este objetivo se recolectó un corpus de 90 textos auténticos y se determinan 65 rasgos lingüísticos con implicancias funcionales y comunicativas que operan no sólo al nivel de palabra y de oración, pero que buscan algún tipo de proyección textual. Las cifras normalizadas de la ocurrencia de estos rasgos en cada texto se introducen en un análisis factorial del cual emergen cinco dimensiones factibles de interpretación funcional. Estas dimensiones tienden a identificar variación entre los registros orales y escritos, y entre los textos técnicos y no técnicos, siendo la Dimensión 5 "Foco Informacional" la más relevante para la descripción del corpus escrito técnico-científico.

This paper investigates the variation across written and spoken registers and specialized and non-specialized texts. The same method was used as in Biber's (1988) research on linguistic variation across speech and writing, but the focus was the Spanish language with emphasis on a text-oriented perspective. A corpus of 90 authentic texts was collected and 65 salient linguistic features, with functional and communicative implications, were determined, most of them not only operating at word and sentence level, but in search for a text projection. Normalized frequencies of these features in each text were then entered in a factor analysis, from which five dimensions emerged. The emerging dimensions tend to identify variation between written and spoken registers and technical and non-technical texts, being the case that the "Informational Focus" turned out to be the most relevant dimension to account for the written technical-scientific corpus.

Key words: specialized discourse, textual dimensions, written language, corpus linguistics.

Introducción

Durante las últimas dos décadas se ha producido un interés creciente en el estudio de los textos especializados, los que han sido indagados desde enfoques heterogéneos y bajo denominaciones diversas (López 2002; Ciapuscio 2003; Flowerdew 2004; Parodi 2005a). Paralelamente, los desarrollos en las ciencias de la computación, junto al advenimiento de nuevas interdisciplinas lingüísticas, entre otras, la lingüística computacional, la ingeniería lingüística y la lingüística de corpus (Moreno 1998; Jurafsky y Martin 2000; Tognini-Bonelli 2001; Parodi 2006) han generado un cambio sustantivo en el modo de almacenar, procesar y analizar los datos lingüísticos. Los corpus electrónicos o digitales de variada índole y características son desde hace ya algún tiempo un modo cotidiano de enfrentar la investigación lingüística, pero tal vez sean el tamaño de los mismos y la tecnología disponible para operar sobre ellos los que presentan rasgos vanguardistas y de gran versatilidad. Desde este contexto, sin lugar a dudas, el escenario de la investigación lingüística ha cambiado o, de modo menos extremo, se encuentra en un proceso decidido de renovación.

Como se sabe, desde hace ya más de veinte años se vienen realizando investigaciones focalizadas en el inglés, tanto como lengua materna o como segunda lengua, a través del uso de grandes corpus de textos digitales y con el apoyo de programas computacionales automáticos para el etiquetamiento morfológico y sintáctico de los documentos. En estas investigaciones se ha explorado una amplia gama de rasgos lingüísticos a través de enormes cantidades de textos pertenecientes a variadas clases textuales (Biber 1988; Louwes, McCarthy, McNamara y Graesser 2004). Ello ha dado origen, entre otros, a varias gramáticas, construidas desde los principios de la lingüística de corpus, que reúnen y distinguen variantes regionales y usos de la lengua oral y la escrita (Quirk, Greenbaum, Leech y Svartvik 1985; Biber, Johansson, Leech, Conrad y Finegan 1999). Estos avances para la lengua inglesa tienden a superar de cierto modo la clásica tendencia en la elaboración de gramáticas hacia una concentración preferente si no exclusiva en el modo escrito de la lengua, con base en un solo registro y desde enfoques eminentemente normativos.

Desde la lengua española, la investigación reciente ha revelado la necesidad de enfatizar el uso de corpus digitales progresivamente más amplios y diversos, con el fin de avanzar en descripciones lingüísticas más profundas y robustas, y también como un medio eficaz de comprobar las hipótesis de los investigadores. Las distinciones, por ejemplo, entre un tipo de discurso especializado y uno de índole más general o entre un tipo de texto escrito y uno oral sólo últimamente han logrado ser descritas de manera más acuciosa, aunque aún de modo preliminar. Desafortunadamente, ello todavía no logra materializarse en la forma de una gramática del español que dé cuenta de estructuras y usos diversos de esta lengua particular y que muestre la heterogeneidad de registros y modos actuales, incluso incorporando

Ahora bien, en este artículo indago en la variabilidad lingüística y funcional existente al interior del corpus El Grial PUCV-2003, compuesto por 90 textos y tres registros diversificados. Con el objetivo general de describir de manera profunda un grupo de textos especializados y escritos de naturaleza técnico-científica (PUCV-CTC), se realiza un análisis multirrasgos y multidimensiones con apoyo de técnicas estadísticas multivariadas (Análisis Factorial de Componentes Principales). Los objetivos más específicos que se persiguen en este artículo son: (a) determinar, desde una perspectiva empírica cuantitativa, estructuras y usos lingüísticos relevantes en el corpus El Grial PUCV-2003; (b) comparar sistemáticamente tres corpora: uno especializado técnico-científico escrito, otro no especializado literario escrito y un tercero oral no especializado de entrevistas; y (c) identificar similitudes y diferencias _en términos de dimensiones_ entre los textos orales y los escritos y entre los especializados y los no especializados (o generales).

En la primera parte del trabajo entrego antecedentes relevantes al tema del análisis multirrasgos y multidimensiones. En el marco metodológico de la investigación se detallan los corpora, los rasgos lingüísticos explorados y las técnicas estadísticas ejecutadas. Posteriormente se presentan los resultados, entre los que destacan los cinco factores de agrupamientos de rasgos, interpretados en cinco dimensiones relevantes. La dimensión denominada "Foco Informacional" resulta ser la que _estadísticamente_ mejor da cuenta del discurso técnico-científico y de la modalidad escrita de la lengua. Por último se adjuntan dos anexos que juzgo de alta relevancia, ya que ellos presentan información funcional acerca de los sesenta y cinco rasgos lingüísticos e información descriptiva y explicativa acerca de las cinco dimensiones exploradas.

1. Antecedentes: lingüística de corpus y análisis multidimensional

Mi interés por el discurso especializado (DE) y la lingüística de corpus (LC) radica, por un lado, en aspectos teóricos y aplicados y, por otro, en implicaciones para la metodología de la investigación lingüística (Parodi 2005a, 2006). Desde la primera, me preocupa aportar a la comprensión del denominado DE y a la descripción de sus rasgos característicos e identificadores; al mismo tiempo, estoy seguro que esta indagación nos lleva a mejorar el conocimiento deficiente que actualmente poseemos acerca de los textos que se leen y se escriben, tanto en el medio educativo de formación especializada sea técnico-profesional secundario, superior o universitario, como en el ámbito laboral mismo. La arista más aplicada de mi interés está en la búsqueda de medios didácticos que _a partir del conocimiento descriptivo recién comentado_ posibilite y pavimente el desarrollo de competencias comunicativas del discurso escrito especializado en lectores y escritores de comunidades discursivas particulares.

Por último me interesan el DE y la LC, puesto que observo en su complementariedad una fortaleza singular. Ya sea que uno adscriba a un enfoque "basado en corpus" (*corpus-based*) o a otro decididamente "guiado por el corpus" (*corpus-driven*), capitalizar de la LC en su versión más metodológica que de paradigma teórico (Parodi 2005, 2006) brinda beneficios certeros para el estudio del DE. Es muy cierto que, según Tognini-Bonelli (2001), tal distinción entre las diversas investigaciones orientadas por el corpus compromete una aproximación en que las categorías de análisis están previamente determinadas y enmarcadas en una opción teórica y otra que busca que las categorías emergan del análisis y den sustento a la construcción de una teoría guiada por los datos. Desde la primera opción, el corpus y las herramientas disponibles actúan como un método de indagación y corroboración de ideas preexistentes; en la segunda, el corpus se constituye en parte integral de la investigación y las categorías deben emerger del mismo.

Al respecto, cabe destacar que Biber (1994) defiende dos ideas significativas respecto a la representatividad de los estudios lingüísticos basados en corpus: 1) todo corpus debe tener una amplitud importante, y 2) un corpus debe contener registros o categorías textuales diversificadas. Ello conlleva mayor validez de las conclusiones y permite la comparación y la generalización. La primera de estas ideas ha sido destacada en la bibliografía acerca de lingüística de corpus (Francis 1979; Leech 1991, 1992; Sinclair 1982, 1991; Johansson 1991; Gibbons 1996), la segunda, en la medida en que el corpus debe ser diversificado para que las categorías

Biber (1988, 2003, 2005) y Biber, Conrad y Reppen (1998) han dado cuenta de interesantes variaciones sistemáticas de orden gramatical y léxico en diversos registros del inglés oral y escrito. Dos hallazgos entre muchos de los reportados parecen relevantes: por un lado, los rasgos lingüísticos individuales presentan una ocurrencia diversa en variados registros; por otro, los mismos o similares rasgos lingüísticos pueden tener funciones diferentes en textos pertenecientes a diversos registros. En este sentido, una de las fortalezas de este enfoque metodológico se funda en un principio lingüístico-comunicativo que resulta extremadamente sensato: la variación entre registros no se explica únicamente por un solo parámetro o dimensión, lo que equivale a sostener que existen múltiples distinciones situacionales entre registros. Dicho de otro modo, no es posible pensar que un rasgo lingüístico o, incluso, unos pocos de ellos puedan explicar exclusivamente una determinada variación entre registros (por ejemplo: oral/escrito, formal/informal). Las investigaciones en que se ha aplicado el análisis multivariado han revelado que diferentes dimensiones se construyen a partir de conjuntos diferentes de rasgos lingüísticos co-ocurrentes, reflejando así diversas interpretaciones funcionales subyacentes (por ejemplo: objetividad, abstracción de información, modalización). Del mismo modo, las tradicionales distinciones de índole más dicotómica (interactivo/no-interactivo) se ven desafiadas por los estudios con análisis multidimensional, ya que se ha llegado a demostrar que existe un *continuum* de variación lingüística a lo largo de los registros.

Un supuesto teórico fundamental del enfoque multidimensiones lo constituye el principio de que la co-ocurrencia de rasgos lingüísticos (determinada mediante procedimientos estadísticos) refleja funciones comunicativas compartidas, es decir, que estos patrones de co-ocurrencia de rasgos se interpretan en términos de funciones situacionales, sociales y cognitivas comunes. En otras palabras, queda claro que los rasgos lingüísticos no co-ocurren aleatoriamente en determinados textos, sino que ellos tienden a aparecer conjuntamente, ya que apuntan a construir funciones compartidas específicas. Por ejemplo, las oraciones pasivas, las nominalizaciones y los participios pasados en función adjetiva (verboides) aportan a la *informatividad* y la *focalización del objeto*, así como también a la *concisión* y *despersonalización*; del mismo modo que los pronombres de primera y segunda persona singular, el tiempo presente, el modo imperativo y los adverbios de lugar y de tiempo están ligados a la *interactividad*.

En esta investigación seguimos en líneas generales la propuesta metodológica de Biber (1988) respecto del Análisis Multirrasgos (AMR) y Análisis Multidimensiones (AMD). Ello, dado que tal enfoque fue creado originalmente como un método analítico para el estudio detallado de las variaciones entre registros; como se aprecia en nuestro caso, nos resulta de utilidad para describir y comparar los textos de los tres corpora. El AMR y AMD utilizan las herramientas metodológicas de la lingüística de corpus según las últimas tendencias (Sinclair 1991; Leech 1991; Svartvik 1992; Tognini-Bonelli 2001; Parodi 2005a, 2006), esto es, requieren de un diseño de corpus representativo, programas computacionales automáticos para etiquetar morfosintácticamente los textos y programas para analizar y cuantificar la ocurrencia de los rasgos lingüísticos previamente marcados en el corpus (Biber, Reppen, Clark y Walter 2001).

2. Metodología

Los objetivos y los corpora. Como se indicó en la introducción, en este trabajo se busca determinar estadísticamente mediante análisis factorial los patrones lingüísticos sobresalientes y co-ocurrentes en el corpus El Grial PUCV-2003 y realizar un estudio comparativo a partir de los tres diferentes grupos de textos recolectados, con base en las dimensiones a determinar e interpretar funcionalmente. De manera más específica, a través de este enfoque multidimensional se comparan las tres áreas técnicas del Corpus Técnico-Científico (CTC).

A continuación entregamos doce puntos que resumen indicaciones técnicas (Biber 1988; Biber et al. 1998; Conrad y Biber 2001) y que orientan los pasos metodológicos a seguir:

- de acuerdo a los registros involucrados que serán considerados en el análisis.
3. Caracterización funcional de los rasgos lingüísticos seleccionados.
 4. Disponibilidad de programas computacionales capaces de analizar automáticamente los textos en formato plano (ASCII o txt).
 5. Marcaje estructural o etiquetado morfológico y/o sintáctico de los textos del corpus.
 6. Interrogación manual o (semi)automática de cada uno de los textos a partir de los rasgos en estudio para determinar su ocurrencia.
 7. Construcción de bases de datos normalizadas, dado el número de palabras divergente entre los textos.
 8. Aplicación, con asistencia de programas computacionales, del análisis factorial a las frecuencias de ocurrencia de los rasgos. Ello con el fin de reducir las variables involucradas y determinar patrones de co-ocurrencia entre los rasgos lingüísticos.
 9. Establecimiento de un conjunto de factores (cada factor queda conformado por un conjunto de rasgos lingüísticos) mediante el análisis factorial, con algún tipo de rotación (Varimax, Cuantrimax, Oblimin, etc.).
 10. Interpretación funcional de los factores, producto del análisis factorial, a partir de la co-ocurrencia de rasgos, constituyendo así una dimensión subyacente de variación.
 11. Confirmación o refutación de la interpretación de los factores mediante el cálculo de los puntajes factoriales.
 12. Cálculo de los puntajes de dimensión para cada texto respecto de cada una de las dimensiones. En esta fase, se comparan los puntajes de cada registro en cada dimensión y se estudian similitudes y/o diferencias lingüísticas y funcionales.

Ahora bien, el corpus El Grial PUCV-2003 se compone de noventa textos que alcanzan un total de 1.466.744 palabras, el cual se subdivide en tres registros o subcorpus (Corpus Técnico-Científico _CTC, Corpus de Literatura Latinoamericana _CLL, y Corpus de Entrevistas Orales _CEO). Cabe destacar que, en un primer momento, sólo se construyó el CTC y, posteriormente, con el objetivo de desarrollar un enfoque contrastivo entre diversos registros de naturaleza especializada y no especializada y entre una modalidad oral y escrita, se procedió a recolectar otros dos corpus, a saber, el CLL y el CEO. La siguiente tabla muestra su distribución por número de textos y palabras.

Tabla 1. Constitución del corpus El Grial PUCV-2003

Tipo de corpus	Número de archivos o textos	Total de palabras
Corpus PUCV-CTC	74 (82%)	626.790 (42%)
Corpus PUCV-CLL	12 (13%)	459.860 (32%)
Corpus PUCV-CEO	04 (5%)	380.094 (26%)
Totales	90 (100%)	1.466.744 (100%)

profesionales, a saber _sector marítimo (Especialidad Operación Portuaria), sector metalmecánico (Especialidad Mecánica Industrial), y sector de administración y comercio (Especialidad Contabilidad). Los textos recopilados corresponden a aquellos que se entregan a los alumnos como parte de lecturas obligatorias o complementarias en cada área técnica, esto es, son parte importante del acceso de estos estudiantes al conocimiento especializado.

Por su parte, el Corpus de Literatura Latinoamérica (CLL) se constituyó con doce textos escritos. Ellos resultaron ser material coincidente, utilizado por los profesores de la asignatura de Lengua Castellana y Comunicación de los establecimientos técnico-profesionales en estudio (entregados como lectura a sus alumnos de 4º año de Enseñanza Media). Esto quiere decir que este corpus se conformó a partir de las obras literarias a que todos los alumnos estaban expuestos.

Por último, el tercer Corpus de Entrevistas Orales (CEO) está formado por dos entrevistas orales en profundidad realizadas a un total de setenta y cinco alumnos de 4º año de Enseñanza Media de establecimientos técnicos y no técnico-profesionales de la ciudad de Valparaíso (educación diferenciada y no-diferenciada). La primera entrevista, de tipo semidirigida, consistió en una conversación acerca de técnicas de estudio y estrategias de lectura y comprensión. La segunda se estructuró según algunas de las temáticas abordadas en la primera conversación y tuvo un carácter más abierto y menos dirigido que la primera. Las entrevistas se realizaron por alumnos y alumnas de último año de la carrera de Pedagogía en Castellano de la PUCV.

Rasgos lingüísticos. Respecto a los rasgos lingüísticos a indagar, se llevó a cabo un rastreo bibliográfico con el fin de identificar categorías gramaticales representativas de relevancia funcional en español. Con esta información disponible, se construyó una matriz con un total de sesenta y cinco rasgos lingüísticos del español. En el Anexo 1 se entrega un listado detallado de estos sesenta y cinco rasgos y su descripción funcional, agrupados en torno a quince categorías más generales, conjuntamente en su relación con ciertos registros y un detallado apoyo bibliográfico referencial. Este apartado ha sido desarrollado por el equipo de investigadores de este proyecto en la Pontificia Universidad Católica de Valparaíso (PUCV). Estamos seguros que este anexo constituye un aporte en sí mismo al estudio de las categorías lingüísticas del español que presentan importancia funcional.

Análisis factorial. Como se sabe, el análisis factorial es un procedimiento estadístico que permite identificar agrupamientos de rasgos lingüísticos que co-ocurren frecuentemente en los textos. Este análisis identifica correlaciones entre un número amplio de variables que se distribuyen de modo similar. Cada grupo de variables co-ocurrentes resulta ser un factor, el cual es posteriormente interpretado en términos de categorías funcionales como una dimensión de variación. Este procedimiento tiende _entre otros_ a la reducción de las variables involucradas en virtud de su co-ocurrencia significativa (Oakes 1998; Hair, Anderson, Tatham y Black 1999).

Una vez que se realizó el análisis factorial (Factores Principales), se determinaron siete factores principales con una rotación tipo Oblimin (Oakes 1998; Hair et al. 1999). Estos factores fueron confirmados a través de los puntajes factoriales. Se obtuvieron así siete posibles dimensiones de las cuales sólo fue factible interpretar funcionalmente cinco de ellas; por tanto, dos factores no permiten esbozar una dimensión consistente debido tanto al reducido número de rasgos constitutivos como a su naturaleza heterogénea.

3. Resultados

En lo que viene, entregamos tres tipos de resultados empíricos: (a) se determinan las dimensiones textuales y se analizan las funciones comunicativas compartidas por los conjuntos de rasgos en co-ocurrencia; (b) a través del enfoque multimensional se comparan los registros orales y escritos, técnicos y no-técnicos, y (c) se analiza la distribución de los grupos de rasgos a través de los tres registros y de los tres ámbitos técnico-científicos del CTC (marítimo, industrial y comercial). Dicho de otro modo, se interpretan, funcionalmente, los parámetros estadísticos encontrados (dimensiones) y se estudia su incidencia en cada uno de los sistemas comunicativos analizados. Siguiendo

cinco factores óptimos. En ellos no se incluyeron aquellos rasgos que obtuvieron un valor absoluto menor a 0,40, dado que normalmente en este tipo de estudios se estiman sin importancia relativa de cara a la interpretación, incluso si fueran estadísticamente significativos. Otras investigaciones utilizan como valor de referencia la cifra de 0,35 (Biber 1988); no obstante ello, en este estudio se decidió usar un puntaje de corte superior. Una explicación pormenorizada de cada una de las cinco dimensiones así como del listado completo de rasgos y pesos estadísticos se adjunta en el Anexo 2.

A continuación se presenta cada factor compuesto por un listado de rasgos lingüísticos agrupados estadísticamente y, junto a cada uno de ellos, se entrega _en cada caso_ una cifra numérica, normalmente, con decimales. Esta cifra facilita el peso factorial que indica una medida de fuerza de la relación entre el rasgo en cuestión y el factor como un todo.

FACTOR 1

Dimensión 1: Foco Contextual e Interactivo

Subordinadas adverbiales de causa efecto	0,945
Adverbios de tiempo	0,934
Adverbio de negación	0,928
Pronombres segunda persona singular	0,911
Pronombres primera persona singular	0,828
Desinencias de segunda persona singular	0,813
Pronombre de negación	0,731
Adverbios de lugar	0,723
Modo indicativo	0,693
Desinencias primera persona singular	0,668
Futuro perifrástico	0,662
Enfatizadores	0,652
Formas activas "ser"	0,637
Verbos modales de volición	0,630
Pronombres demostrativos	0,592
Pronombres de segunda persona singular	0,531
Subordinadas adverbiales condicionales	0,523
Adverbios de negación temporal	0,503
Subordinadas sustantivas	0,497
Subordinadas adverbiales de tiempo	0,487
Verbos privados	0,474
Frases infinitivas en función nominal	0,466
Presente	0,424
Frases preposicionales complemento del nombre	-0,545
Nominalizaciones	-0,479
Sustantivos	-0,443
Participios en función adjetiva	-0,437

Los rasgos que constituyen el Factor 1 son los más numerosos y ostentan los pesos estadísticos más altos. Los rasgos que se reúnen en este factor denotan una gran relación funcional entre la mayoría de ellos. Su interpretación no resulta compleja. Como se aprecia, los rasgos con peso negativo son un grupo menor e indican una clara interpretación. Ellos son clásicamente considerados portadores de la carga referencial del texto; permiten la integración de grandes cantidades de información, y una alta frecuencia de ellos apunta a una fuerte densidad y concisión de la información. Contrariamente, los rasgos con altos puntajes positivos apuntan hacia una referencia directa del contexto físico y temporal, determinan marcos de orden de la sucesión de hechos, establecen una vinculación con la acción y expresan motivos y consecuencias. También, a través de estos rasgos, se hace referencia a los participantes y existen suficientes evidencias de que la marca de lo situado está presente. Esta dimensión se concreta en la acción, en la sucesión de acontecimientos y en las relaciones interpersonales de tipo dialógico. Así, los rasgos lingüísticos involucrados, en su conjunto,

asocia con un foco en la explicitud y dependencia del contexto y en la activa participación de los interlocutores, rasgos clásicos del discurso oral y dialógico.

Esta Dimensión 1 representa un parámetro importante de variación, compuesto por un conjunto amplio de rasgos lingüísticos y definitorios de una clara distinción entre dos polos como son, por un lado, lo contextual e interactivo (oralidad: conversaciones) y, por otro, lo informacional, altamente planificado y cohesionado (escritura: exposición). Por ello, se la nombra como Dimensión 1 "Foco Contextual e interactivo".

FACTOR 2

Dimensión 2: Foco Narrativo

Pronombres segunda persona plural	0,842
Pronombres primera persona singular	0,828
Futuro perifrástico	0,823
Pretérito imperfecto	0,820
Pronombre tercera persona plural	0,708
Modo indicativo	0,686
Desinencias primera persona plural	0,667
Verbos modales de volición	0,651
Pretérito indefinido	0,614
Pronombre de negación	0,590
Verbos privados	0,577
Adverbios de lugar	0,533
Pronombres segunda persona singular	0,529
Verbos perceptivos	0,496
Adverbio de negación	0,493
Adverbios de negación temporal	0,482
Formas activas "estar"	0,460
Verbos públicos	0,445
Pronombres primera persona plural	0,431
Desinencias tercera persona singular	0,423
Conjunciones advers., disy. y aditivas	0,411
Frases infinitivas en función nominal	0,405
Conjunción "ni"	0,402
Nominalizaciones	-0,581
Frases preposicionales complemento del nombre	-0,562
Adjetivos atributivos	-0,442

Los rasgos positivos constitutivos del Factor 2 denotan un marcado acento en la identificación de las personas del discurso. También se aprecia la co-ocurrencia de los tiempos verbales del pasado: el pretérito imperfecto y el pretérito indefinido. Esta dimensión se asocia con una sucesión de acontecimientos, que implica la precisión de circunstancias de tiempo y lugar, como también la participación de las personas del discurso. Todo ello apunta hacia una determinada trama textual de tipo narrativo. Estas características co-ocurrentes llevaron a denominar este factor como Dimensión 2 "Foco Narrativo". Ellas permiten identificar textos literarios orales o escritos de orden general, a diferencia de textos altamente especializados.

FACTOR 3

Dimensión 3: Foco Compromiso

Verbos privados	0,824
Pronombres primera persona singular	0,789
Pretérito indefinido	0,705
Verbos modales de volición	0,655
Desinencias primera persona singular	0,640
Modo indicativo	0,630

Pronombres segunda persona singular	0,563
Desinencias segunda persona plural	0,562
Frases infinitivas en función nominal	0,518
Subordinadas sustantivas	0,467
Subordinadas adverbiales de concesión	0,452
Formas activas estar	0,435
Pronombres segunda persona plural	0,427
Adverbio de negación temporal	0,411
Pronombres primera persona plural	0,402
Frases preposicionales complemento del nombre	-0,457

El tercer factor ha sido interpretado a partir de la alta ocurrencia de los llamados verbos privados (decidir, adivinar, sentir, determinar, demostrar, estimar, reconocer) y de volición (querer+infinitivo), de los pronombres personales y de las desinencias verbales de primera persona que constituyen marcas relevantes de la expresión del "yo". La clara identidad de quien escribe o habla queda manifiesta en el texto de manera explícita y quien participa se compromete e involucra con lo que dice y hace. Este compromiso con el discurso y su contenido revela los afectos y los propósitos del escritor/hablante. Este conjunto de rasgos co-ocurrentes, sistemáticamente, permite denominar al tercer factor como Dimensión 3 "Foco Compromiso". Esta dimensión está asociada a textos en los que sobresale la intención y la actitud del emisor que revela su voluntad de involucrarse en el discurso de manera explícita y de asumir un rol preponderante. Dicho de otro modo, esta dimensión caracteriza a textos en los que aparecen participantes reales que expresan intenciones y actitudes proposicionales frente a lo dicho.

FACTOR 4

Dimensión 4: Foco Modalizador

Formas activas "ser"	0,671
Atenuadores	0,656
Verbos modales de posibilidad	0,641
Adverbios de modo	0,606
Adjetivos predicativos	0,565
Desinencias tercera persona plural	0,549
Subordinadas adjetivas	0,514
Desinencias tercera persona singular	0,405
Sustantivos	-0,494

Para construir la interpretación del cuarto factor, se atendió a la co-ocurrencia significativa de atenuadores (parecer que, creer, tal vez, a lo mejor, quizás, quizá), verbos modales de posibilidad (poder) y adverbios modales (probablemente, posiblemente). Ellos revelan un parámetro funcional muy preciso: la regulación y atenuación de la información entregada, es decir, la expresión de la probabilidad y la incertidumbre de los hechos o acontecimientos descritos o narrados. La conjunción sistemática de estas marcas lingüísticas tiende a darse en textos con énfasis en cómo (*modus*) se dicen las cosas, más que en lo dicho (*dictum*). Esta distribución de rasgos presenta el contenido de un discurso como incierto y abierto a la verificación; por el contrario, se aleja de la supuesta "objetividad" que otros grupos de rasgos típicamente tienden a representar; en estos últimos, las restricciones impuestas llevan a textos con escaso compromiso explícito del hablante/escritor, con alta precisión léxica y densidad informativa. De este modo, se constituye la Dimensión 4, llamada "Foco Modalizador". Así, la Dimensión 4 agrupa rasgos característicos de textos orales y escritos, clásicamente ligados a la narración y descripción; no obstante ello, también puede marcar textos expositivos y argumentativos, más representativos de la alta especialización.

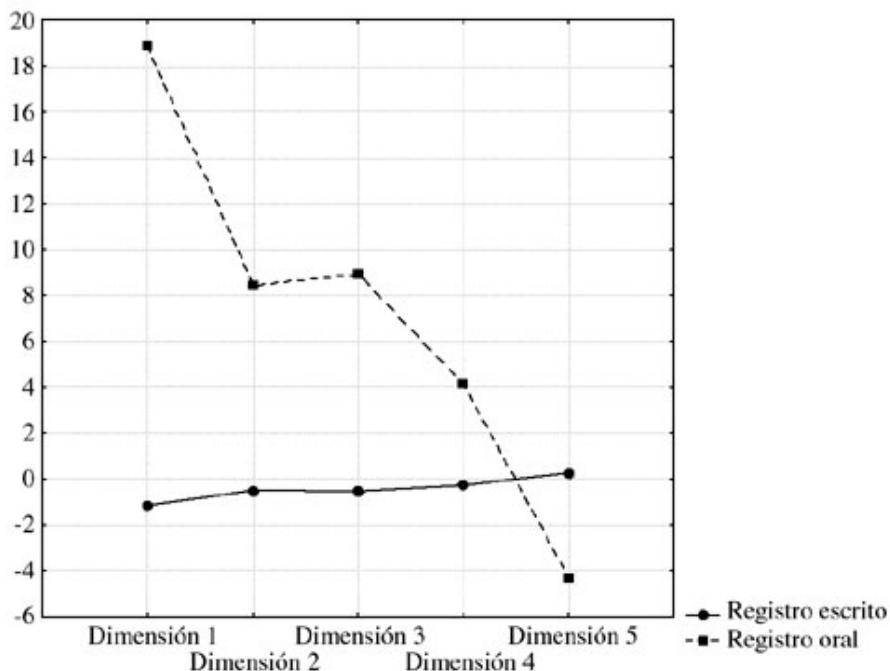
FACTOR 5

Verbos modales de obligación	0,496
Modo subjuntivo	0,494
Nominalizaciones	0,456
Participios en función adjetiva	0,413
Frases preposicionales complemento del nombre	0,413
Desinencias tercera persona singular	-0,632
Pretérito indefinido	-0,630
Forma estativa activa "estar"	-0,595
Verbos privados	-0,575
Pronombre de negación	-0,572
Verbos modales de volición	-0,503

Finalmente, el último y quinto factor se constituye a partir de once rasgos: cinco positivos y seis negativos. Los rasgos positivos en co-ocurrencia como verbos modales de obligación revelan la necesidad y certeza de los juicios expresados; el modo subjuntivo remite a organizaciones de mayor complejidad sintáctica. La presencia de nominalizaciones junto a participios en función de adjetivo, sustantivos (comunes y propios) y frases preposicionales como complemento del nombre son todos rasgos indicadores de integración y compactación de información altamente abstracta, típica del discurso especializado escrito. En resumen, los rasgos positivos agrupados en torno a esta última dimensión se encuentran básicamente orientados hacia la informatividad, entendida esta como la concentración de información en unidades y estructuras lingüísticas compactas, que presentan los datos lo más concisa y precisamente posible. Por otra parte, los seis rasgos negativos aquí presentes apuntan hacia una contextualización de eventos señalados en el discurso, tendencia atenuada en textos de alta informatividad. Esta caracterización del factor 5 derivó en su interpretación como Dimensión 5 "Foco Informacional".

Variación al interior del corpus El Grial PUCV-2003. Focalizamos, ahora, las similitudes y diferencias entre los registros y las tres áreas de especialización técnico-científicas con base en las cinco dimensiones previamente comentadas. En un primer momento, nos interesa indagar el impacto de las dimensiones detectadas en la distinción entre oralidad y escritura. Como bien se sabe, esta distinción ha resultado polémica llegado el momento de determinar rasgos específicadores y, en algunos casos, los autores no alcanzan acuerdo (Halliday 1985; Chafe y Danielewicz 1987; Biber 1988; Louwerse et al. 2004). El Gráfico 1 muestra los resultados obtenidos en este punto.

Comparar un corpus de registro escrito (conformado por textos técnico-científicos de divulgación didáctica y textos de literatura latinoamericana) con otro oral (compuesto por entrevistas semidirigidas a estudiantes secundarios) revela diferencias interesantes. Cabe destacar, primeramente, que las dimensiones construidas a partir de los 65 rasgos lingüísticos permiten distinguir un impacto diferenciador entre oralidad y escritura, hecho que en sí mismo constituye un primer hallazgo. El registro oral aquí estudiado muestra un comportamiento muy variado a lo largo de las 5 dimensiones, pero relativamente fácil de interpretar. La alta variación que se observa entre la Dimensión 1 y la Dimensión 5 es prueba inequívoca de que los rasgos constitutivos de una y otra se oponen en la identificación de textos prototípicamente orales de otros evidentemente escritos. En este sentido, la mayor ocurrencia del modo subjuntivo, verbos modales de obligación, nominalizaciones, participios en función adjetiva y frases preposicionales como complemento del nombre no aparecen como rasgos que tienden a co-ocurrir sistemáticamente en los textos orales aquí estudiados; ellos adquieren _de este modo_ un carácter identificador de modalidades escritas de la lengua y se distancian prototípicamente de la oralidad.



De modo muy claro se puede caracterizar la oralidad a partir de las 4 primeras dimensiones, pues todas ellas presentan índices positivos. Al mismo tiempo, se hace evidente que la interactividad y contextualización (Dimensión 1) son los rasgos identificadores más representativos de un discurso oral dialógico como el que constituye este corpus. Los rasgos de la narratividad y del compromiso aparecen también como caracterizadores del tipo de textos orales aquí descritos. No es una casualidad que la Dimensión 5 "Foco Informacional" sea la única que registra un peso negativo y una diferencia estadísticamente significativa con el resto de las otras dimensiones y, muy en particular, con la Dimensión 1.

En cuanto al registro escrito, se detecta un comportamiento relativamente homogéneo de las cifras a lo largo de las 4 primeras dimensiones. Coincidientemente con lo dicho antes respecto de la oralidad, las cuatro primeras dimensiones presentan puntajes negativos y sólo la Dimensión 5 alcanza un peso positivo, distanciándose de las primeras y constituyéndose en la principal caracterizadora de la escritura. Llama la atención, eso sí, la escasa variabilidad entre dimensiones en el registro escrito. Si bien es claramente explicable para las dimensiones 1, 2 y 3, habría sido esperable que la Dimensión 5 mostrara mayor diferencia con el resto y posicionara a la escritura de modo más distante de las otras. No obstante lo anterior, la diferencia resulta estadísticamente significativa.

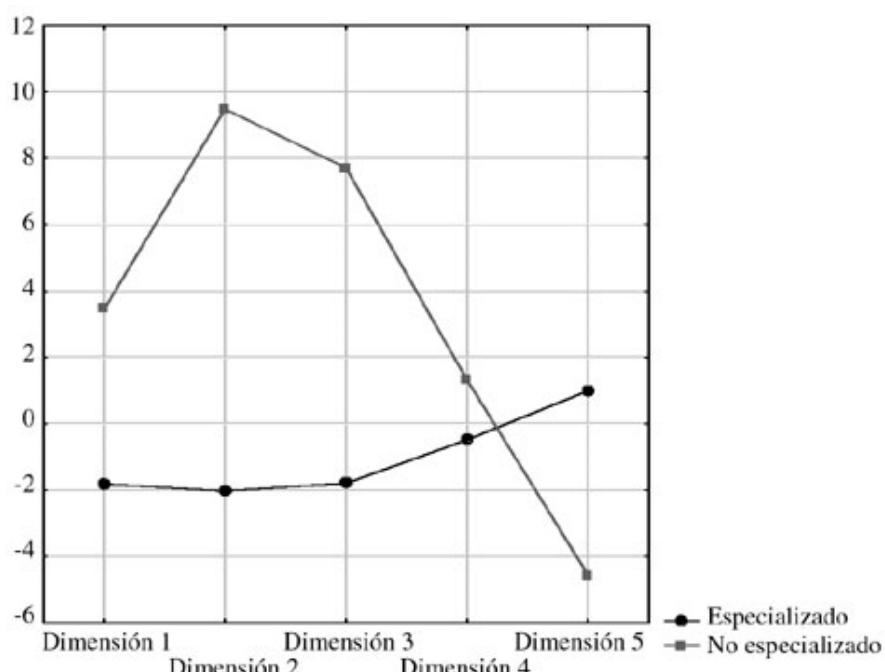
Ahora bien, si para el idioma inglés Biber (1988) identificó un conjunto de rasgos y dimensiones que logran capturar la diferencia entre diversos tipos de registros a lo largo del *continuum* oralidad/escritura, Louwes et al. (2004) sostienen que el estudio de Biber se concentra en el nivel de la palabra y que no alcanza un impacto que involucre al texto como una unidad mayor. Louwes et al. (2004) dan cuenta de una investigación que siguió los mismos pasos metodológicos y se apoyó en el mismo corpus que Biber (1988), pero utilizó un conjunto mucho más amplio de categorías e incluyó aspectos de cohesión textual en varios niveles. En esta nueva indagación para la lengua inglesa, se logra determinar una dimensión específica que distingue oralidad de escritura y que se apoya en un amplio grupo de variables que intentan capturar relaciones textuales mucho más allá del horizonte de la palabra.

Hacemos referencia a estos estudios de la lengua inglesa con el fin de precisar el alcance de nuestros datos. Es bien sabido que para el español no contamos con antecedentes en esta línea de investigación; por lo tanto, no disponemos de otros estudios como el aquí presentado

ostentan la productividad de las empleadas por Louwerse et al. 2004) buscan de modo certero _desde el plano de la palabra y la oración_ una proyección en la dimensión textual. Esto se explica a partir del principio metodológico de la construcción interpretativa de una dimensión, es decir, en la interpretación funcional y discursiva del conjunto de rasgos co-ocurrentes y en distribución complementaria. Desde esta mirada, los datos aportados en el gráfico 1 constituyen un indicador importante para el español, aunque queda claro que la cantidad de textos y los tipos constitutivos en cada modalidad de lengua no son un conjunto suficientemente variado y amplio que permitan una mayor distinción intragrupal.

El siguiente gráfico entrega datos relativos al registro especializado (escrito) y a los registros no-especializados, tomados estos últimos en su conjunto (orales y escritos). Para ello, se han calculado los puntajes promedio por dimensión (determinados por medio del puntaje factorial o *factor score*).

Gráfico 2
Comparación entre registros especializados y no-especializados



Como se aprecia en el gráfico, la comparación de los pesos factoriales entre las cinco dimensiones con base en los textos de especialidad y los considerados aquí como no especializados (orales y literarios) muestra una evidente diferenciación. Los textos no especializados destacan por pesos positivos en las cuatro primeras dimensiones en donde la narratividad aparece como la dimensión más identificadora junto al Foco Compromiso. Son las marcas de persona singular y plural (pronombres y desinencias verbales), el pretérito imperfecto y el indefinido, el modo indicativo y los verbos modales de volición, entre otros, los que identifican los textos de literatura latinoamericana escrita y al corpus de textos orales, es decir, los que representan la ausencia de especialización temática. Junto a ellos, los rasgos característicos de la interactividad, contextualización temporal y espacial, y modalización también se constituyen en prototípicos de los discursos no especializados. Por el contrario, la Dimensión 5 se destaca por un peso bastante negativo y estadísticamente significativo respecto de las otras cuatro dimensiones, hecho que indica que la compactación de información y alto grado de referencialidad no se evidencian como característica de estos textos no especializados.

alcanzan pesos totalmente opuestos a los obtenidos por el corpus no especializado. En este caso, los datos estadísticos son muy similares a los del corpus escrito, entregados en el gráfico 1. En las cuatro primeras dimensiones se registran pesos negativos y sólo emerge la Dimensión 5 "Foco Informacional" como la distintiva para este grupo de textos. Estos datos son congruentes con la teoría de corte lingüístico y funcional que hemos venido sustentando para diferenciar e identificar a los textos prototípicos del discurso especializado, de modo que los hallazgos entregados se constituyen en una corroboración empírica relevante que pone de manifiesto las diferencias entre un determinado registro y otro. Estos valores permiten avalar empíricamente la postura a favor del denominado discurso de especialidad y muestran que las 65 categorías exploradas son un conjunto de rasgos valiosos como indicadores de variabilidad.

Un hallazgo importante lo constituye el que, según las cifras entregadas en los dos primeros gráficos, la especialización se detecta como estrechamente vinculada a la modalidad escrita de la lengua. Tal como lo ha mostrado el puntaje factorial promedio respecto de la Dimensión 5 "Foco Informacional", es esta dimensión y el conjunto de rasgos lingüísticos asociados a ella, los que establecen la distinción fundamental entre, por una parte, lo que está más cerca de la escritura que de la oralidad y, por otra, lo que resulta más prototípicamente especializado o no. La explicación que se desprende de estos datos apunta claramente a que las características típicas de un discurso especializado coinciden con los rasgos clásicos de la escritura en su versión más prototípica. De este modo, queda claro que la supuesta especialización a través de la modalidad oral de la lengua no encontraría apoyo en nuestros datos y cabría suponer, entonces, que cuando se habla de oralidad especializada, a lo que se tiende mayoritariamente es a un tipo de discurso escrito que se dispone de manera oral (el ejemplo clásico es una conferencia académica que se dispone oralmente, pero que se produce de modo escrito).

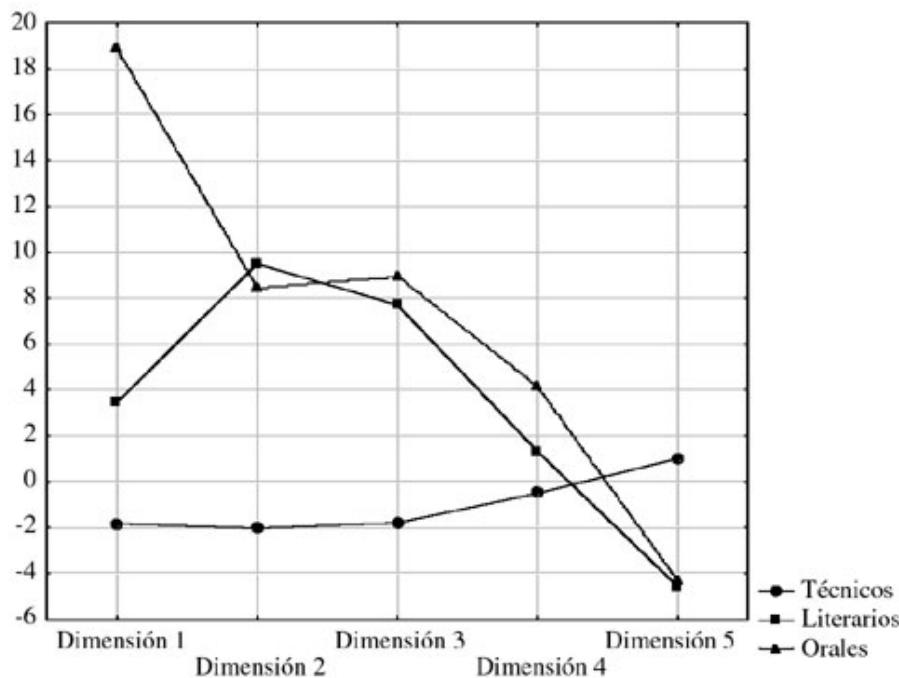
También vale la pena llamar la atención sobre el comportamiento estadístico de la Dimensión 5. Tal como se observa en los valores hasta ahora disponibles, es la función informacional de la lengua, manifiesta a través de diversos mecanismos, la que se hace presente como rasgo distintivo entre las diversas aproximaciones exploradas. Es la Dimensión 5 la que obtiene puntaje negativo para el registro oral y la que puntúa positivamente para la escritura, mostrando una diferencia estadísticamente significativa entre ambos valores. Es, del mismo modo, la Dimensión 5 la que arroja puntajes factoriales positivos para el corpus especializado y la que muestra puntajes negativos elevados para el corpus no-especializado, siendo la diferencia entre ambos estadísticamente significativa. Todo ello implica, desde la perspectiva de esta investigación interesada en el discurso especializado escrito, que la Dimensión 5 "Foco Informacional" juega un rol central en la descripción de este tipo de discurso y que sus rasgos lingüísticos constituyen un conjunto de características nucleares en la definición del mismo. Así, entre otros, la monorreferencialidad, la precisión, la compactación, la carencia de emotividad, son rasgos relevantes que se desprenden como prototípicos del discurso focalizado en esta investigación.

A continuación, en el [gráfico 3](#), se presentan los puntajes promedio por dimensión para cada uno de los tres registros del corpus El Grial PUCV-2003 (técnicos, literarios y orales), con respecto a las cinco dimensiones en cuestión.

Resulta interesante comprobar que los textos técnicos del CTC obtienen un puntaje factorial negativo muy similar a lo largo de las cuatro primeras dimensiones. Esta homogeneidad revela que, por un lado, los rasgos que identifican estas cuatro dimensiones deben presentar una ocurrencia similar en los textos de este corpus; según estos datos, la interactividad, contextualización, modalización y compromiso no serían rasgos característicos ni de alta ocurrencia sistemática en los textos del ámbito técnico y científico, comparado con los otros dos registros. Por otra parte, esclarecedor que la Dimensión 5 "Foco Informacional" se presente con un puntaje promedio positivo (0,9), lo que permite identificar y distinguir significativamente el CTC del CLL y CEO. Al mismo tiempo, como se observa, los textos de literatura latinoamericana y los textos basados en entrevistas orales poco planificadas, con un grado importante de espontaneidad, muestran puntajes negativos en la Dimensión 5 y, en cambio, obtienen los más altos puntajes promedio positivos en el resto de las otras cuatro dimensiones. Entre ellos, cabe destacar el elevado puntaje factorial que los textos orales alcanzan en la Dimensión 1 "Foco Contextual de Interactividad" (1,0), lo que evidencia, entonces,

pronombres y adjetivos demostrativos se constituyen en prototípicos.

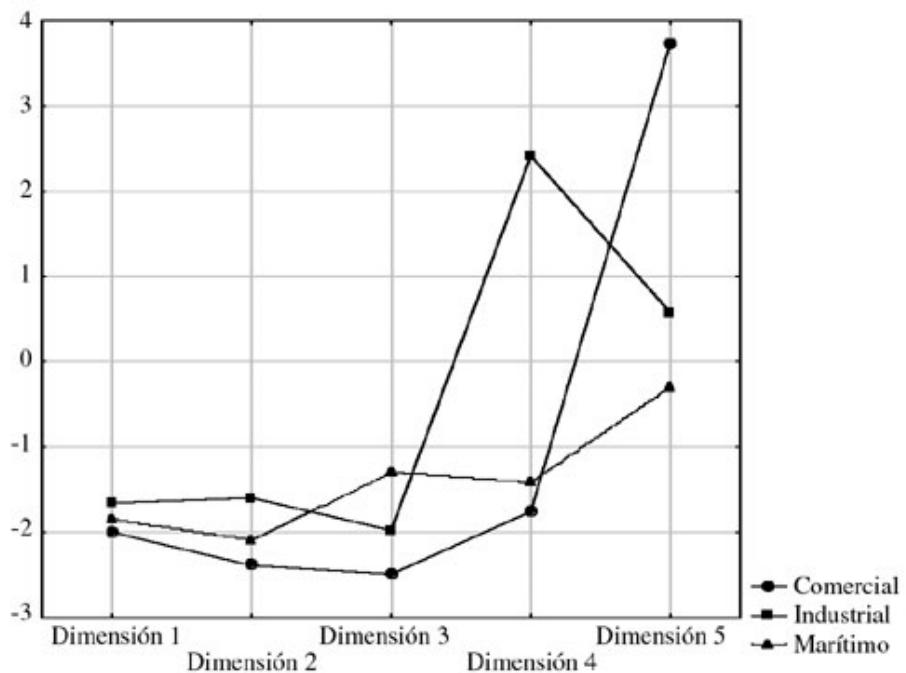
Gráfico 3
Dimensiones y registros



Este análisis comparativo arroja similitudes y diferencias muy elocuentes. Permite identificar las áreas de intersección entre la oralidad de las entrevistas y la narratividad de la literatura latinoamericana. También ayuda a separar los textos de alta especialización en los que se detectan construcciones gramaticales más complejas y de mayor empaquetamiento y reducción de información de aquellos que involucran detalladamente a los participantes y sus relaciones interpersonales. En estos últimos textos (orales y escritos), se detecta y se expresa _de modo más explícito_ el involucramiento del autor a través de marcas lingüísticas específicas (ciertos tipos de verbos, adverbios, pronombres, etc.), hecho que también puede caracterizar el discurso especializado escrito, pero que muchas veces se implica a través de otros recursos. Como se aprecia, la variación entre estos registros enfrenta un *continuum*, identificado en este caso por medio de las dimensiones y los rasgos lingüísticos que ellas capturan y permite identificarlas y caracterizarlas prototípicamente. Las implicaciones derivadas de ello son múltiples, en particular, en lo que respecta al discurso especializado de divulgación didáctica, foco de interés de esta investigación. El diseño de materiales didácticos que aborden el discurso especializado en español escrito y quienes se concentren en su enseñanza deben capitalizar estas descripciones.

Finalmente, en este último gráfico, se entregan los puntajes factoriales por dimensión para cada uno de las tres áreas de especialización del corpus PUCV-CTC-2003 (comercial, industrial y marítima).

Gráfico 4
Dimensiones y ámbitos de especialización



Es interesante observar que, también en el gráfico 4, el área comercial muestra una clara diferenciación en cuanto a la Dimensión 5 "Foco Informacional" del resto de los otros dos registros (CLL y CEO); en este nuevo análisis, más pormenorizado, es justamente esta quinta dimensión la que parece revelar la mayor distinción entre las áreas técnico-científicas. Según se aprecia, el ámbito comercial presenta el mayor puntaje promedio positivo en esta dimensión (3,8), hecho que muestra su mayor carga informacional a través de alta densidad léxica y complejidad sintáctica.

Parodi (2004) ya había detectado, a través de un estudio descriptivo simple, cierta variabilidad interna en el comportamiento de los sesenta y cinco rasgos lingüísticos entre las tres áreas técnicas. Esta investigación preliminar mostraba un patrón distintivo entre el área marítima y las otras dos. El mismo dato encuentra apoyo certero en estos análisis dimensionales con técnicas tipo *factor score*. Según se comprueba, el mayor puntaje negativo en cuanto a la carga informacional lo obtiene el ámbito de especialización marítima (-0,1), el mismo que apareciera distante de los otros dos en el trabajo que comentamos. Considerando este resultado, la Dimensión 5 parece distinguir entre textos de un ámbito y otro. Será necesario un estudio cualitativo profundo para dar cuenta con mayor detalle de qué clases textuales del área comercial se alinean en torno a este patrón dimensional.

Seguidamente en la jerarquía, la Dimensión 4 también aporta a la distinción entre especialidades. Es ahora el área industrial la que revela mayor puntaje promedio positivo en la Dimensión "Foco Modalizador" (2,5), mientras que el marítimo y el comercial obtienen puntajes negativos muy cercanos y sin significatividad estadística entre ellos. Estos datos permiten inferir que los textos del ámbito industrial contienen mayor regularidad en los patrones sistemáticos de ocurrencia en torno a los rasgos distintivos de la atenuación e incertidumbre. En el resto de las dimensiones (1, 2 y 3), las tres áreas del CTC presentan cifras negativas y relativamente similares, hechos que indican que estas dimensiones no aportarían gran explicación diferenciadora en la descripción del discurso técnico-científico de divulgación didáctico escrito. También se puede sugerir que los textos de estos ámbitos especializados no destacan por la ocurrencia de estos rasgos que denotan interacción, relaciones interpersonales e involucramiento de los participantes en el discurso. Todo ello es congruente con otros antecedentes disponibles acerca de este tipo de textos.

En términos generales, cabe destacar que el análisis multirrasgos y multidimensiones (AMR y AMD) probó ser una metodología poderosa desde el marco de la lingüística de corpus. Los beneficios del AMR y del AMD implican más que una mera descripción de los rasgos superficiales de los textos, sino que aportan a la determinación de regularidades sistemáticas que develan funciones comunicativas y aproximaciones hacia una descripción profunda de las estructuras y usos lingüísticos en contextos determinados. Al mismo tiempo, la fortaleza que radica en la muestra relativamente grande de textos especializados (comparado con análisis de textos ejemplares y con muestras de textos generales) permite alcanzar conclusiones robustas y con mayores proyecciones. Estos aspectos llevan a la lingüística _en esta línea_ no sólo a contar con una alternativa metodológica de investigación, sino también a delinejar un paradigma que brinda renovados bríos a las indagaciones venideras para el español.

De modo más específico, la determinación de las cinco dimensiones mostró ser un recurso de extraordinario potencial descriptivo; por ejemplo, entrega una distinción fundamental entre los textos estudiados de la modalidad oral y los de la escrita: los primeros aparecen anclados en la interactividad, contextualización y las relaciones interpersonales, esto es, fuertemente descritos por la Dimensión 1 "Foco Contextual e Interactivo". Del mismo modo, la oralidad aquí descrita (entrevistas) no se relaciona de modo directo con una prosa informacionalmente densa, tanto desde el punto de vista léxico como sintáctico. Si bien es cierto que el registro literario escrito coincide con el oral en las Dimensiones Foco Narrativo, Foco Compromiso y Foco Modalizador, la distinción manifiesta en la Dimensión 1 es suficiente como para establecer una diferencia importante. Evidentemente, esto no permite hablar de una fuerte polarización entre oralidad y escritura dentro del corpus general, lo que reafirma la adscripción a la idea de un desplazamiento continuo entre los registros, a modo de categorías prototípicas y también difusas.

Un aspecto particularmente relevante lo constituye la relación que se devela entre el grado de especialización de un texto y su vinculación con la modalidad escrita de la lengua. Así, los rasgos lingüísticos y funcionales asociados a la Dimensión 5 "Foco Informacional" destacan como los caracterizadores de una escritura técnica especializada, a la vez que permiten identificar algunos de los rasgos más prototípicos del lenguaje escrito. Es justamente esta dimensión la que resulta central para la descripción del corpus CTC y su distinción de los otros corpora tales como el CLL y el CEO.

Otro aspecto interesante es la continuación que aquí se realiza de la descripción del CTC, ya iniciada desde otra mirada por Parodi (2004, 2005b). Sin lugar a dudas, se hace necesario abordar un análisis más pormenorizado de las distinciones entre las tres áreas de especialización del CTC (comercial, marítimo e industrial), posiblemente desde estudios de índole más cualitativa. También queda pendiente el cálculo de los puntajes factoriales entre las diversas clases textuales que componen este corpus con el fin de indagar el aporte del AMD a la diferenciación y descripción de las mismas. Los antecedentes empíricos hasta ahora aportados dan cuenta de una interesante homogeneidad a lo largo de las Dimensiones 1, 2 y 3 entre las tres áreas técnico-científicas: estos textos no parecen tener una fuerte marca de narratividad, de involucramiento de los participantes ni de interactividad. Por el contrario, los textos del área industrial y marítima sí presentan diferencias significativas en las Dimensiones "Foco Modalizador" y "Foco Informacional". El dominio industrial muestra una tendencia hacia la modalización y el dominio marítimo se distingue de los textos de las otras dos áreas por un grado positivo mayor de densidad informativa.

Por último, los resultados aportados muestran la relevancia del estudio del uso lingüístico en entornos de formación técnico-profesional y confirman la necesidad de crear materiales especializados de instrucción sobre la base de descripciones empíricas de los registros en estudio. De acuerdo a los hallazgos, los estudiantes de las áreas especializadas requieren desarrollar la habilidad para manejar una variedad del español técnico escrito muy particular: prosa densa, académicamente, en términos léxicos, morfológicos y sintácticos, pero también textos con marcas de modalización. Esta evidencia también sugiere que existen razones significativas para atender diferencialmente al análisis, comprensión y producción del registro técnico-científico que circula y se genera tanto en el ámbito educativo como en el ambiente laboral. Todo estudiante debería practicar con una amplia variedad de registros, ya que en su

de manera casi automática la inmersión en estos tipos de textos (Parodi 2005b). Una alternativa didáctica podría ser que el material especializado contuviera una organización progresiva desde lo más divulgativo didáctico hasta aquellos textos más típicos del ámbito profesional y laboral.

Anexos

ANEXO 1: Rasgos lingüísticos corpus PUCV-2003

A. MARCADORES DE TIEMPO VERBAL

1. Pretérito indefinido (indicativo) [PRET.IND]

Expresa anterioridad del verbo al acto de la palabra. Es acción que ya no se puede volver a realizar (Contreras 1984). Funcionalmente, ha sido definido como tiempo representativo de la secuencia narrativa, porque es el tiempo en el que se explica la complicación que perturba el estado inicial (Bassols y Torrent 1997). Se le considera característico de la lengua escrita y de la prosa narrativa (De Kock y Gómez 2002).

2. Pretérito imperfecto (indicativo) [PRET.IMP]

Presenta una acción durativa, o bien, reiterada o habitual, simultánea en todo o en parte con otra también pretérita, durativa o instantánea. Por esto se dice que es un tiempo relativo (Alvar 2000; Moliner 1986). Específicamente, señala simultaneidad respecto a un momento anterior al punto central (Alvar 2000). Ha sido considerado característico de la lengua escrita, en particular de la prosa narrativa (De Kock y Gómez 2002). Bassols y Torrent (1997) señalan que es un tiempo con el que se describen los estados iniciales en la secuencia narrativa, como también los incisos descriptivos insertos en el relato. También consignan su uso en construcciones argumentativas, tal vez porque, como anota Moliner (1986), el pretérito imperfecto en los verbos modales (deber, poder, tener que) expresa opinión acerca de la conveniencia o procedencia de una cosa.

3. Pretérito perfecto (indicativo y subjuntivo) [PRET.PER]

Manifiesta una acción pasada que tiene incidencia en el momento de la enunciación, que dura hasta el momento mismo en el que se habla (Moliner 1986). Gili Gaya (1980) constata que, aunque el español moderno establece la diferencia entre el pretérito indefinido y el perfecto, extensas zonas de España e Hispanoamérica han preferido una forma por sobre la otra, debido a que en ambas predomina el aspecto perfectivo.

4. Presente (indicativo y subjuntivo) [PRES]

Presenta las acciones que coexisten con el acto de la palabra (Gili Gaya 1980). Es usado para todo lo universal. Habitual en máximas y sentencias. Cuando la acción se refiere al momento del habla, se trata del presente actual. Pero, además, con el presente se enuncian las verdades intemporales y las acciones habituales. Puede expresar acciones pasadas, en el llamado presente histórico (Moliner 1986). Asimismo, cuando la acción se siente de segura realización, puede tener valor de futuro (Moliner 1986).

5. Futuro (indicativo y subjuntivo) [FUT]

Expresa acción venidera, independientemente de cualquier otra acción. Según Gili Gaya (1980), el empleo del futuro supone una cierta capacidad de abstracción por parte del hablante, debido a esto aparece tardíamente y es poco frecuente en el habla infantil. Es preferido en ciertas modalidades, como la lengua escrita; su uso en la lengua oral aumenta conforme el grado de formalidad del habla (Arroyo 2000). El futuro del subjuntivo aparece sólo en ciertos casos: en oraciones con "si" y en las "oraciones de relativo pronominal o adverbial"; y esto solamente en lenguaje anticuado y en algunas frases estereotipadas.

6. Futuro perifrástico [FUT.PERI]

Tiempo que corresponde al mundo comentado, en oposición al imperfecto, tiempo de lo narrado (Weinrich 1974). Es más común en los registros coloquiales y descuidados (Arroyo 2000; Contreras 2000).

B. MARCADORES DE MODO VERBAL

7. Indicativo/imperativo [INDIC.IMP]

Se da una coincidencia de forma entre la segunda persona singular del imperativo y la tercera persona singular del indicativo.

8. Subjuntivo/imperativo [SUBJ.IMP]

Existe un sincretismo entre las formas del subjuntivo y del imperativo debido a que las únicas formas específicas del imperativos son las segundas personas del singular (tú) y el plural (vosotros). Las otras personas se toman del presente del subjuntivo. En las frases negativas, las segundas personas se sustituyen por las del subjuntivo.

9. Modo indicativo [MOD.IND]

Expresa la modalidad declarativa experiencial (Cepeda 2002), como también estados o acciones consideradas reales (Gómez Macker y Peronard 1988). El modo indicativo es propio del intercambio discursivo oral (Cepeda 2002). Este modo hace referencia a hechos reales localizados en un tiempo verdadero (Criado de Val 1962; Gómez Macker y Peronard 1988). Generalmente se refiere a un acontecimiento localizado a partir de un pretérito. Según Alcoba (1999), el modo indicativo se utiliza cuando hay aserción. En este modo predomina la función representativa y su modalidad distintiva es la lógica o declarativa. "Se define por la relación objetiva entre hablante y mensaje y la modalidad declarativa (...) y como la modalidad declarativa es la forma no marcada (...) el modo indicativo es siempre el más extenso" (Hernández 1996: 473).

10. Modo subjuntivo [MOD.SUBJ]

Permite especular sobre hechos inciertos, sujetos a una apreciación subjetiva (Criado de Val 1962). En este modo se expresa la subjetividad del hablante o del sujeto de la comunicación ante el enunciado, es el modo de la hipótesis y de la virtualidad. En toda expresión con subjuntivo se percibe la presencia del hablante y del enunciado. El hablante puede aparecer, además, "como agente de la enunciación, como sujeto del enunciado, hecho que no podía darse en el llamado imperativo" (Hernández 1996: 375). Sin embargo, además de estas características, se debe tener presente que el subjuntivo aparece normalmente como oración subordinada, y que esta remite a textos de mayor complejidad sintáctica (Gili Gaya 1980), cuya función es enmarcar la información del discurso de diversas maneras.

11. Modo imperativo [MOD.IMP]

Su misión propia es el mandato directo. Se limita casi exclusivamente al lenguaje hablado, locacional (Criado de Val 1962). Estas formas verbales se oponen a las demás formas personales, porque están especializadas en las expresiones de la modalidad deontica apelativa de mandato. Las formas de imperativo no transportan otro significado que el de mandato, en cambio, es posible formular mandato mediante otras formas lingüísticas como el futuro o el presente de indicativo, presente del subjuntivo, etc. (Gili Gaya 1980).

C. DESINENCIAS VERBALES DE PERSONA

12. Primera singular [DES.1S]

Refleja el carácter egocéntrico de un texto, implica una necesidad de comunicación directa.

tendencia común a evitar la referencia a la primera persona y a usar otros procedimientos de presentación de la persona del autor (Ciapuscio 1992).

13. Segunda singular [DES.2S]

Se utiliza la segunda persona con el objeto de producir un efecto determinado: "generalizar la experiencia enunciada e incluir al interlocutor de una manera personal y afectiva" (Calsamiglia y Tusón 1999: 139). Se asocia al lenguaje coloquial.

14. Tercera singular [DES.3S]

Aporta significación de tiempo, modalidad, persona y número del enunciado donde aparece la forma verbal (Alcoba 1999). Esta desinencia es característica en la lengua escrita, en particular, en el artículo científico (Kaiser 2002).

15. Primera plural [DES.1P]

La identificación de la persona que habla con la primera persona del plural incorpora al locutor a un grupo (Calsamiglia y Tusón 1999). El llamado "nosotros de modestia" reemplaza al yo porque, en ocasiones, se considera inapropiado utilizarlo en público. Otra función de la primera persona plural es incluir al interlocutor, involucrándolo y suavizando así los mandatos o las peticiones. Ciapuscio (1992) ha constatado que este uso es frecuente en el discurso de divulgación dirigido al gran público. Asimismo, la autora destaca en este tipo de textos y también en aquellos de divulgación destinado a un público más restringido, el uso de nosotros que incluye de modo aparente al interlocutor y cuya función es "ordenar la interacción y asegurarla" (Ciapuscio 1992: 198).

16. Segunda plural [DES.2P]

Al igual que en los casos de primera y segunda persona singular, esta desinencia verbal manifiesta el carácter egocéntrico y la necesidad de comunicación directa.

17. Tercera plural [DES.3P]

Esta desinencia es rasgo característico de la lengua escrita y del artículo científico (Kaiser 2002) debido a su uso para construir oraciones impersonales, en las que el sujeto se calla intencionalmente, carece de interés o es imposible de determinar (Gili Gaya 1980).

D. PRONOMBRES PERSONALES

18. Primera persona singular [PRON.1S]

Debido a que en español los verbos conjugados llevan la marca de persona en la desinencia, el pronombre sujeto es innecesario casi siempre. En primera y segunda persona su aparición es enfática y "significa insistencia particular en hacer resaltar el sujeto" (Gili Gaya 1980: 172). Refiere directamente a los participantes, marcadores de la presencia del yo (Biber 1988). En la narración, distingue al testigo-participante o al protagonista (Bassols y Torrent 1997). Abundan cuando el contexto no aclara en forma suficiente la persona verbal. Es propio de las entrevistas orales (Castellano 2000). En general, los pronombres personales de la primera y segunda persona _denominados déicticos_ remiten a los participantes en el acto de comunicación, función que les es propia. La propiedad semántica esencial de los pronombres personales es que no permiten asignar valores de verdad a los enunciados independientes del contexto (Fernández 1999). Otro aspecto importante es que los pronombres de primera y segunda persona son reversibles, en el sentido de que el yo de turno no puede menos que ceder su derecho a un tú, si quiere tener un interlocutor válido, aunque no son imprescindibles para expresar el concepto de persona gramatical (Fernández 1999).

19. Primera persona plural [PRON.1P]

Con la forma nosotros se evita utilizar el pronombre yo, sin que ello implique, en principio,

tipos de referentes, en lo que se llama "plurales ficticios" (Alcina y Blecua 1975).

20. Segunda persona singular [PRON.2S]

Las formas tú, usted son características de la lengua escrita, específicamente de la prosa narrativa y la poesía (De Kock y Gómez 2002). Requiere de un destinatario específico e indican un alto grado de interacción y acción (De Kock y Gómez 2002). En los textos narrativos, puede aparecer la narración en segunda persona singular, en caso de que los propios personajes adopten una perspectiva exterior que les dé mayor lucidez para el autoanálisis (Genette 1972).

21. Segunda persona plural [PRON.2P]

Es válido para este pronombre cuanto se ha dicho para la segunda persona singular.

22. Tercera persona singular [PRON.3S]

Los pronombres él/ella aparecen cuando existe el riesgo de malinterpretar la referencia, porque la información que aporta el contexto falla. Se asocia normalmente con entrevistas orales. (Castellano 2000). Él/ella tienen un uso deíctico que los asemeja a los demostrativos, pero también un uso referencial, esto es, tienen la capacidad de retomar los rasgos de un individuo presente en el contexto lingüístico (Fontanella 1999). En el caso de los textos narrativos, la tercera persona corresponde al narrador omnisciente (Bassols y Torrent 1997). A diferencia de los demás pronombres personales, la tercera persona singular es la "no persona", es decir, está excluida de la instancia de comunicación y remite no al enunciado, sino a una situación "objetiva". En este sentido se trata del término no marcado y de hecho no existe en todas las lenguas (Fernández 1999).

23. Tercera persona plural [PRON.3P]

Al igual que para los pronombres de tercera persona singular, el pronombre sujeto aparece cuando puede haber ambigüedad, puesto que son muchas las posibles terceras personas.

24. Pronombre demostrativo [PRON.DEM]

Al igual que otros elementos lingüísticos agrupados bajo la denominación de deícticos, los demostrativos señalan, seleccionándolos, algunos elementos del entorno contextual (Calsamiglia y Tusón 1999), adquiriendo su sentido pleno sólo en el contexto en el que se emiten. Son usuales en el lenguaje oral.

E. FORMAS NOMINALES

25. Nominalizaciones [NOMINAL]

Este término designa a los nombres derivados de bases verbales y adjetivales, así como al proceso de su formación. Los nombres derivados pueden tener como referente un evento o un proceso, un objeto, un estado, una propiedad o un producto resultante de un acontecimiento o proceso, recurso típico del lenguaje técnico (académico) para expresar significados complejos y abstractos (Picallo 1999). Permiten integrar información en pocas palabras (Chafe 1982, 1985) y reducir oraciones completas en series de frases nominales más compactas y eficientes (Janda 1985); además, tienen la función de transportar información altamente abstracta (Biber 1988). Ciapuscio (1992) incluye la nominalización dentro de los procedimientos de desagentivación, característica de la comunicación científica.

26. Sustantivos (comunes y propios) [SUST]

Según Biber (1988), los sustantivos son los principales portadores del significado referencial del texto. La ocurrencia de algunos tipos de sustantivos (los más largos y los derivacionales) se asocia con discursos cuyo foco es altamente informational, caracterizados por una cuidadosa integración de la información. El nombre significa de manera independiente y

F. FORMAS PASIVAS

27. Pasivas con "se" [PAS.SE]

Se especializan en acciones de carácter habitual, general, con objetos internos y sujetos implícitos poco delimitados, aunque pueden aparecer en todo tipo de contextos. Tiende a usarse cuando el sujeto de la frase es nombre de cosa. En general, es más frecuente que la pasiva con ser (Martín 1986). Aparece tanto en la lengua hablada como en la escrita. Se ha notado un incremento en el uso de esta construcción en el lenguaje divulgativo informativo (Mendikoetxea 1999 a y b). Ciapuscio (1992) constata su frecuencia en textos de divulgación científica con lo que se acentúa el carácter impersonal del texto. Según la autora, esta tendencia a la omisión del agente es una característica del lenguaje de la ciencia que se mantiene en la divulgación.

28. Pasivas con "ser" sin agente [PAS.SER-a]

Se especializan en acciones de carácter puntual, con objetos externos y un marcado carácter intencional que denota la existencia de un sujeto implícito delimitado. Son de mayor uso en la lengua escrita. La ausencia del agente se ha atribuido a la intención de callar u ocultar el sujeto nocional (Mendikoetxea 1999 a y b).

29. Pasivas con "ser" con agente [PAS.SER+a]

Con estas pasivas parece ser siempre gramaticalmente posible la expresión del sujeto nocional, independientemente de su papel semántico: agente, destinatario, experimentante o fuente. La expresión del agente se da fundamentalmente en textos escritos de carácter periodístico (Hernández 2000a).

30. Pasivas con "estar" [PAS.ESTAR]

Los gramáticos no se han puesto de acuerdo acerca del carácter pasivo de estas construcciones. Los que las rechazan argumentan que las pasivas con ser expresan una acción que recae en el sujeto paciente; en cambio, las construcciones con estar expresan el resultado de esa acción. Según Mendikoetxea (1999b) ambas construcciones tienen un significado pasivo.

G. ESPECIFICIDAD LEXICA

31. Relación clase/tipo (type/token) por forma [TYP.TOK.form]

Es una relación porcentual que expresa el rango de variabilidad de unidades dentro de un universo. En este caso se refiere a la especificidad léxica de las formas que aparecen en un texto. Una tasa alta se relaciona con el lenguaje escrito debido al carácter planeado de su producción. Por el contrario, una tasa baja se asocia con el lenguaje oral por el carácter espontáneo de la oralidad (Horowitz y Samuels 1987).

32. Relación clase/tipo (type/token) por lema [TYP.TOK.lem]

Se refiere a un tipo especial de relación que expresa la variabilidad o especificidad de los lemas de un texto, entendidos estos como las raíces léxicas de las formas que aparecen en un texto determinado.

H. FORMAS ESTATIVAS ACTIVAS

33. Ser [ACT.SER]

Los verbos conectores expresan básicamente relaciones de equivalencia, igualdad, similitud, dependencia o atribuyen cualidades o valores como, por ejemplo, ser, parecer, equivaler. Se considera que ser atribuye cualidades consideradas como permanentes. Este tipo de verbos,

es decir, no tienen valor especializado, pero forman parte de lo que se ha llamado unidades de conocimiento especializado (Lorente 2002).

Destacan por su frecuencia en la descripción (Bassols y Torrent 1997). Con "ser", se formulan juicios que no dependen de la experiencia inmediata (Gili Gaya 1980). Su empleo en las oraciones copulativas se reduce a servir de nexo entre el sujeto y el predicado y permiten la expresión temporal (Gili Gaya 1980).

34. Estar [ACT.ESTAR]

En tanto verbo conector cumple las mismas funciones antes mencionadas para el verbo ser, pero a diferencia de este, con el verbo estar las cualidades se consideran transitorias o accidentales y, además, se perciben como resultado de un cambio o transformación (Gili Gaya 1980). En tanto expresa estados, está léxicamente incapacitado para expresar un cambio o progreso durante el periodo de tiempo en que transcurre (De Miguel 1999).

I. TIPOS VERBALES

35. Públicos [V.PUBLIC] (Decir, explicar, admitir, estar de acuerdo, declarar, quejarse, remarcar, replicar, sugerir).

Reportan acciones propias de la actividad científica (Harvey 2002). Son verbos que denotan acciones que pueden ser observadas públicamente; son actos de hablas primarios tales como decir y explicar, y se usan comúnmente para introducir aserciones indirectas (Biber 1988).

36. Privados [V.PRIVAD] (Descubrir, creer, concluir, decidir, adivinar, encontrar, sentir, temer, determinar, demostrar, estimar, reconocer).

Se refieren a actividades perceptibles sólo por el propio hablante y se utilizan también como atenuadores (Palmer 1974; Weber y Bentivoglio 1991). Los verbos privados se destacan por expresar estados intelectuales o actos intelectuales no observables (Biber 1988). Corresponden a los procesos mentales de Halliday (1994) y a los verbos epístémicos léxicos de Hyland (1998).

37. Persuasivos [V.PERSUA] (Acordar, arreglar, preguntar, rogar, insistir, instruir, estipular, sugerir).

Implican intenciones que suponen un cambio en el futuro (acordar, arreglar, preguntar, rogar, comandar, decidir, demandar, garantizar, insistir, instruir, ordenar, comprometer, pronunciar, proponer, recomendar, requerir, estipular, sugerir, urgir) (Biber 1988). Según Bosani (2000), aluden a acciones actitudinales que podemos realizar para/sobre el contenido proposicional de lo dicho.

38. Perceptivos [V.PERCEP] (Sentir, ver, oír, oler).

Expresan estados intelectuales cuando el objeto sobre el que recae la acción es un sustantivo abstracto. Tienen un significado más concreto cuando los objetos en los que recae la acción son sustantivos concretos. Corresponden a un subtipo de los procesos mentales de Halliday (1994). Revelan focalización interna en la narración (Arianzen 2001).

J. VERBOS MODALES

39. Posibilidad [V.MOD.POS]

Expresan la opinión y la actitud del hablante/escritor respecto al contenido declarado, atenuando generalmente la fuerza de los enunciados que se emiten. Son frecuentes en los artículos de investigación científica (Hyland 1998) como una forma de adelantarse a las posibles objeciones y permite que el autor no aparezca tan taxativo.

40. Necesidad [V.MOD.NEC]

(Hyland 1998), aumentando la fuerza de lo aseverado.

41. Obligación [V.MOD.OBL]

Expresan el punto de vista del escritor que juzga la verdad de lo dicho en términos de certeza (Hyland 1998). Además, la modalidad deontica se ocupa de la necesidad o la posibilidad de los actos ejecutados por agentes moralmente responsables (Osorno 2000).

42. Volición [V.MOD.VOL]

Para Langacker (1990) los modales de capacidad y de volición, o sea, los modales dinámicos, se anclan en la realidad, mientras que los otros modales disponen de un mecanismo de proyección y parece que se anclan en la predicción. Según Olbertz (1998), los modales dinámicos (de capacidad y volición) corresponden a la modalidad orientada al participante.

K. MARCADORES DE MODALIDAD

43. Atenuadores [ATENUAD]

Son marcadores informales y menos específicos de probabilidad e incertidumbre. Marcan el contenido de una proposición como incierto. Co-ocurren con marcadores

de interactividad como los pronombres de segunda persona y los signos de interrogación. Aparecen en mayor medida en textos científicos de ciencias sociales (Markkanen y Schröder 2000). Son propios del discurso académico, pero varían notablemente su uso según el tipo de texto. Son escasos en los manuales o en los materiales de apoyo para un curso. En estos tipos de texto se transmite un saber decantado, sobre el que hay un consenso amplio: se enseña lo que la comunidad académica en cuestión tiene por seguro, y, al enseñar, se simplifica ignorando las alternativas que pueden hacer avanzar el conocimiento (Vásquez 2001). Son recursos que se ponen al servicio de la eficacia o eficiencia pragmática, es decir, se emplean para lograr efectos comunicativos que van más allá de la mera transmisión de información (Narbona 2000).

44. Enfáticos [ENFATIZ]

Acentúan el valor de los verbos. Se usan para indicar en términos positivos la confiabilidad de las proposiciones. Pueden ser usados en funciones no proposicionales para señalar solidaridad con el interlocutor. Característico de la entrevista oral (Cepeda 2002).

L. ADVERBIOS

45. De lugar [ADV.LUG]

Sitúan la significación del verbo en coordenadas espaciales y añaden información que completa la estructura argumental del predicado (Bosque 1990).

46. De tiempo [ADV.TIEMP]

Debido a su función defictiva establecen marcos de orden de la sucesión de hechos o pistas contextuales para la interpretación de lo dicho (Kovacci 1999). Son circunstanciales en posición posverbal. Actúan como circunstanciales si están en el ámbito de la interrogación y de la negación.

47. De modo [ADV.MOD]

Según Bassols y Torrent (1997), son propios de la descripción. En principio denotan la manera en que se presentan los acontecimientos o se realizan las acciones (Bosque 1990).

48. De cantidad [ADV.CANT]

a las descripciones o ampliar una explicación.

M. MARCADORES DE SUBORDINACION

49. Subordinadas sustantivas con "que" [SUB.SUST]

Las oraciones subordinadas sustantivas introducidas por "que" desempeñan las funciones propias de los sustantivos, aunque no se corresponde en todos los casos con sintagmas nominales. Por su parte, el empleo del determinante está semánticamente establecido: permite designar eventos o procesos que no se contemplan en su realización, sino en su resultado (Delbecque y Lamiroy 1999; Demonte 1997).

50. Subordinadas adjetivas de pronombre relativo [SUB.ADJ]

El empleo de las subordinadas de relativo (adjetivas) "permite atribuir al sustantivo cualidades muy complejas para las cuales no tiene el idioma adjetivos o participios léxicos" (Gili Gaya 1980: 301).

51. Subordinadas adverbiales de razón o causa /efecto [SUB.ADV.c-e]

Señalan la razón del hecho mencionado en la oración principal. Permiten expresar la causa o la consecuencia (Gili Gaya 1980). La causa o la consecuencia van normalmente introducidas por un conector causal, que indica que los enunciados que lo siguen explican o dan razón de los enunciados antecedentes; o consecutivo, que indican que los enunciados que lo siguen son efecto de los razonamientos antecedentes (Calsamiglia y Tusón 1999).

52. Subordinadas adverbiales de concesión [SUB.ADV.conc]

Expresan una reserva que carece de eficacia para que se efectúe lo expuesto en la oración principal, cuya realización se cumple, a pesar del obstáculo (Bassols y Torrent 1997). Contribuyen a la estructuración del discurso argumentativo.

53. Subordinadas adverbiales condicionales [SUB.ADV.cond]

Indican la condición que debe cumplirse para que se realice lo expresado en la oración principal (Gili Gaya 1980). Contribuyen a hacer explícitos los razonamientos, por tanto, son propios de los discursos expositivos y argumentativos.

54. Subordinadas adverbiales de tiempo [SUB.ADV.tiem]

Expresan el tiempo de la acción contenida en la oración principal, determinando una relación de anterioridad, simultaneidad o posterioridad (Pérez-Rioja 1971). Su función es ubicar temporalmente la acción principal en relación con la subordinada. Con ayuda de estos adverbios es posible expresar los matices temporales para los que la sola presencia del verbo no es suficiente (Gili Gaya 1980).

55. Frases infinitivas en función nominal [FRA.INF.nom]

Indican un intento por condensar información, omitiendo los participantes (Biber 1988; Chafe 1982, 1985). Tienen un carácter eventivo relacionado con la descripción de procesos o de actividades atléticas (Demonte y Varela 1997). Desde un punto de vista lógico, son oraciones abiertas, puesto que el sujeto del infinitivo es una variable que se extrae de los sintagmas nominales que aparecen en el mismo contexto lingüístico (Gómez 1999). Según Biber (1988), estas estructuras no cumplen funciones prototípicas ni fijas.

N. FRASES PREPOSICIONALES Y ADJETIVOS

56. Frases preposicionales (complemento del nombre) [FRA.PREP.cn]

Sirven para integrar grandes cantidades de información en un texto (Biber 1988; Chafe 1982,

(Picallo 1999).

57. Adjetivos atributivos (calificativos) [ADJ.ATRIB]

El adjetivo puede cumplir la función de predicado de una oración copulativa, caracterizadora o como complemento predicativo, bien obligatoriamente escogido por el verbo, bien opcional. La construcción modificadora o atributiva tiene una estrecha relación con la construcción del predicado nominal, ya que casi todos los adjetivos que funcionan como predicados en oraciones copulativas caracterizadoras pueden ser también modificadores. Los adjetivos son palabras que se aplican a otras palabras que nombran objetos físicos o mentales; por medio de los adjetivos se adscribe a esos objetos una propiedad o un conjunto de propiedades. Más estrictamente, un adjetivo modificador adscribe propiedades, cuya especificación sirve para definir o delinear con mayor precisión a la entidad mentada para caracterizarla e identificarla entre varias similares, para clasificarla o establecer taxonomías culturales y científicas para indicar relaciones genéticas o metonímicas (relaciones parte-todo). La característica fundamental de los adjetivos (atributivos y predicativos), lo que lo diferencia de los nombres es que son términos generales y por ello pueden aplicarse a múltiples objetos y son de carácter gradual (Picallo 1999).

58. Adjetivos predicativos [ADJ.PRED]

Son adjetivos que expresan cualidades del sujeto por medio de las oraciones atributivas (Gili Gaya 1980); dicho de otro modo, su dependencia con el sustantivo se hace de forma indirecta a través de un índice verbal (Marcos Marín 1975). Hay adjetivos que se combinan con el verbo ser: pertenecen a este grupo los adjetivos de relación y procedencia. Otros adjetivos sólo se utilizan con el verbo estar. En general, son todos aquellos que indican un resultado. Por eso, los participios _salvo en construcciones pasivas_ se combinan siempre con el verbo estar. Debe tenerse en cuenta que hay ocasiones en que los participios se combinan con el verbo ser, porque de esta combinación surge un significado diferente. Otros adjetivos que sólo se combinan con el verbo estar son: contento, satisfecho. Lo mismo sucede con adjetivos que poseen significado de participios, indicando el resultado de algo. Otros adjetivos se combinan con el verbo ser y estar, el significado varía en cada caso. Con el verbo ser, el adjetivo designa: a) una cualidad inherente a lo designado por el sujeto; b) una característica más o menos permanente del sujeto, que pertenece a su descripción o que lo clasifica (lo introduce en una clase de entes). Con el verbo estar, designan cualidad adquirida, producto de un cambio real o pensado.

59. Adjetivos demostrativos [ADJ.DEM]

Los adjetivos demostrativos actualizan al sustantivo. Pertenecen a la categoría pragmática de los deícticos, por lo tanto, su interpretación está estrechamente ligada tanto al contexto textual como extratextual. Los determinantes (los artículos, posesivos y demostrativos) y los cuantificadores (los numerales e indefinidos) constituyen clases cerradas con un número fijo de miembros desprovistos de significado léxico. A estos términos, a los que no es posible darle una definición de diccionario (no corresponden en verdad a ningún campo nocional), no les falta, sin embargo, significado gramatical o funcional: los artículos marcan género y número, sirven para introducir un nombre en el discurso o para establecer una relación anafórica, expresan unicidad o presuposición de existencia (Picallo 1999).

60. Participio función adjetiva [PARTICI.adj]

Se sitúa preferentemente en el discurso escrito más que en el oral y su interpretación usual es que se usa para la integración y las elaboraciones estructurales (Biber 1988). Janda (1985) establece que se utilizan en la toma de apuntes, porque son más compactos e integrados y por ello sirven para la producción de un discurso altamente informacional cuando el tiempo es limitado. Ciapuscio (1992) destaca el uso de los verboides, entre ellos el participio, como recurso de desagentivación, procedimiento de ocultación del agente que permite condensar la información. La autora ha constatado su frecuencia en textos de divulgación científica y en textos científicos.

61. Conjunciones adversativas, aditivas y disyuntivas [CONJ.dis.adv.ad]

Son características de la lengua escrita y de la prosa narrativa (De Kock y Gómez 2002). La coordinación es el procedimiento gramatical que se usa para asociar constituyentes sintácticos sin establecer una jerarquía grammatical entre ellos (Camacho 1999). El uso de estas conjunciones es frecuente como índice de simplicidad o sencillez de la lengua hablada. Según Avila (2000), el habla utiliza con mayor frecuencia la coordinación antes que la subordinación, especialmente en las cconversaciones cara a cara y a distancia.

O. MARCADORES DE NEGACION

62. Adverbio de negación [ADV.NEG]

Importan información más coloquial y fragmentada (Biber 1988). Son indicadores de la modalidad negativa en el sentido que hacen referencia a la actitud del emisor con respecto al receptor y al propio mensaje.

63. Adverbios de negación temporal [ADV.NEG.tiem]

Los adverbios de negación temporales (nunca, jamás, tampoco) son capaces de expresar negación cuando preceden al verbo (Kovacci 1999).

64. Conjunción de negación [CONJ.ni]

La conjunción no presenta una caracterización funcional particular, adquiriendo una orientación comunicativa determinada por el rol de los elementos adyacentes que pone en relación (Hernández 1996). Tal es el caso de la conjunción negativa "ni".

65. Pronombre de negación [PRON.NEG]

Es de uso más coloquial (Tottie 1983). Son pronombres invariables indefinidos cuantificadores existenciales en el ámbito de la negación (Sánchez 1999). Puesto que funcionan como sustantivos, admiten la presencia de un adyacente adjetivo, que adopta invariablemente el significante propio de los morfemas de masculino y singular (Alarcos Llorach 1999).

ANEXO 2: Descripción de las cinco dimensiones

Dimensión 1: Foco Contextual e Interactivo

Para interpretar el factor 1 y lograr determinar la dimensión subyacente, se deben evaluar las funciones compartidas por la mayoría de los rasgos que co-ocurren. De los 23 rasgos iniciales con peso estadístico sobre 0,40, se destacan quince rasgos con pesos superiores a 0,60. Los rasgos que se reúnen en este factor denotan una gran relación funcional entre la mayoría de ellos. Su interpretación no resulta compleja y la bibliografía pertinente entrega apoyo suficiente. De este modo, en principio, se puede sostener que altas frecuencias de co-ocurrencia de adverbios de tiempo y de lugar, subordinadas adverbiales de causa/efecto y de tiempo y de pronombres demostrativos revelan una referencia directa al contexto físico y temporal, determinan marcos de orden de la sucesión de hechos, establecen una vinculación con la acción y expresan motivos y consecuencias (lo que explica la sucesión de los acontecimientos). También ellos implican un compromiso del autor del texto y contendrían la marca de lo situado (RAE 1973; Pérez-Rioja 1971; Di Tullio 1997; Kovacci 1999). Junto a estos rasgos de situacionalidad y contexto, el modo indicativo y el tiempo presente apuntan a una modalidad declarativa experiencial (Hernández 1996; Bassols y Torrent 1997; Avila 2000; Criado de Val 1962; Gómez Macker y Peronard 1988), característica de la entrevista oral (Cepeda 2002). Los pronombres de primera y segunda persona singular y las desinencias verbales de primera y segunda persona singular refieren directamente a los participantes. También son clásicamente considerados como marcadores de la presencia del sujeto en el

aluden a una comunicación directa, un estilo directo (De Kock y Gómez 2002); por ende, son representativos de la interacción dialógica.

Por su parte, los rasgos con pesos estadísticos negativos tales como las nominalizaciones y los sustantivos son clásicamente considerados portadores de la carga referencial del texto y una alta frecuencia de ellos señala una fuerte densidad informacional (Picallo 1999; Di Tullio 1997; Ciapuscio 1992; Chafe 1982, 1985; Chafe y Danielewicz 1987; Halliday 1993; Biber y Finnegan 1986; Biber 1988). Del mismo modo las estructuras adjetivas (frases de participio en función adjetiva y frases preposicionales) actúan como expansiones en el proceso de textualización, es decir, ellas permiten la integración y precisión de grandes cantidades de información en un texto (Ciapuscio 1992; Harvey 2002). Como se comprende, los textos caracterizados por estos tipos de rasgos suelen ser altamente abstractos, condensar cantidades de información importante y expresar significados complejos; por lo tanto, constituyen un recurso típico del lenguaje técnico y científico (Hernanz 1999; Biber 1988; Burdach 2000; Zarzalejos 2001; Moyano 2000).

En resumen, la dimensión que se infiere a partir de este factor se concreta en la acción, en la sucesión de los acontecimientos y en las relaciones interpersonales de tipo dialógico; lo que también se destaca a través de la deixis temporal, espacial y demostrativa. Todos los rasgos lingüísticos involucrados, en su conjunto, permiten suponer que los textos caracterizados por esta dimensión no contienen información altamente abstracta como tampoco evidencian una integración concisa y precisa de la información, sino, por el contrario, la alta frecuencia de ocurrencia de estos rasgos tipificadores y estadísticamente positivos se asocian con un foco en la explicitud y dependencia del contexto y en la activa participación de los interlocutores, rasgos clásicos del discurso oral y dialógico. Esta dimensión emergente caracteriza los discursos en que se explicitan las marcas espacio-temporal y donde existe colaboración mutua entre los participantes, lo que los torna más auténticos y menos planificados; la espontaneidad e interactividad son típicas de estos discursos. En ellos pueden co-existir múltiples organizaciones discursivas, tales como: narración, exposición, argumentación y descripción. Por todo ello, como un modo de reflejar las funciones primordiales detectadas en conjunto, hemos nombrado esta dimensión con la etiqueta "Foco contextual e interactivo".

Puntajes factoriales y relaciones entre los corpora (literario, oral y técnico)

Dado que nos interesa mostrar las relaciones entre los tres registros (técnico-científico, CTC; literatura latinoamericana, CLL; y entrevistas orales semiestructuradas, CEO) y _más específicamente_ aislar variables identificadoras del discurso técnico-científico en contextos didácticos, se comparan estos tres registros a través de cada una de las dimensiones interpretadas. Para ello se realiza un estudio estadístico a partir de los puntajes factoriales que resumen la frecuencia de cada uno de los rasgos en cada factor para cada uno de los textos. Los puntajes factoriales para cada texto se promedian con los puntajes de todos los textos de un corpus específico (CTC, CLL y CEO) y, de este modo, se obtiene un promedio del puntaje factorial para cada dimensión. Estos puntajes promedio por corpus o registro se comparan con el fin de determinar los tipos de relaciones (similitud o diferencia) existentes (Hair et al. 1999).

De la aplicación de estos procedimientos estadísticos se deduce que, en la dimensión 1 Foco Contextual e Interactivo, las entrevistas orales (CEO) evidencian el mayor peso (26,16). Los textos del corpus literario (CLL), por su parte, ostentan un peso mucho menor (7,12) y, en tercer lugar, se ubican los textos del corpus técnico-científico (CTC) con un peso negativo (-2,14). Estos datos estadísticos permiten confirmar que la Dimensión 1 aparece fuertemente ligada, en primer lugar, a los textos orales en su modalidad dialógica tipo entrevista, seguida por los textos literarios escritos; por tanto, se hace evidente que la dimensión Foco Contextual e Interactivo no es prototípica de los textos técnicos.

Dimensión 2: Foco Narrativo

El factor 2 también presenta 23 rasgos co-ocurrentes que se distribuyen entre los puntajes 0,84 y 0,40. La presencia significativa de los pronombres personales, especialmente los de tercera persona singular y primera persona plural, sujetos humanos, protagonistas de

enunciación y las ausentes en relación a aquellas (Calsamiglia y Tusón 1999). Congruente con lo anterior, surgen en co-ocurrencia los rasgos asociados a los tiempos verbales del pasado, el pretérito imperfecto, que describe situaciones y circunstancias y, su contraparte, el pretérito indefinido, que señala los sucesos y, por tanto, el dinamismo de las acciones (Kovacci 1993), denotando así una referencia directa al mundo narrado (Weinrich 1974; Arroyo 2000; De Kock y Gómez 2002). No menos importante es la co-aparición del futuro perifrástico, demostrándose que si bien los hechos narrados se sitúan generalmente en el pasado, otros transcurren también en la actualidad de los personajes (Contreras 2000). A estos tiempos verbales se asocia el uso del modo indicativo, que permite a los protagonistas expresar estados o acciones como reales (Gómez Macker y Peronard 1988). Todo esto se complementa con el uso de los verbos modales de volición, verbos privados, verbos perceptivos y verbos públicos que se orientan a dar cuenta de las actitudes subjetivas del sujeto hablante (Arianzen 2001). De hecho, lo notable de los verbos de actividad mental _empleados, sobre todo, en primera persona singular_ no es sólo su uso, según lo que Palmer (1974) denomina verbos privados y que se refieren a actividades perceptibles sólo por el propio hablante, sino que conjuntamente ellos pueden cumplir también una función atenuadora (Weber y Bentivoglio 1991).

Ahora bien, la trama personal y temporal hasta aquí conformada encuentra su anclaje en los adverbios de lugar y de tiempo que sitúan la significación del verbo en unas coordenadas espaciales o temporales (Bosque 1990). Asimismo, el adverbio de negación, la conjunción "ni" y las conjunciones aditivas, adversativas o disyuntivas sirven de apoyo a la sucesión de hechos en el pasado (Biber 1988; De Kock y Gómez 2002; RAE 1973; Pérez-Rioja 1971). Por último, el verbo estar estativo activo destaca por su frecuencia en secuencias descriptivas (Bassols y Torrent 1997; Lorente 2002) y las frases infinitivas en función nominal se explican por su incidencia en construcciones en que aparecen los verbos modales de volición y los privados, los que remiten a los participantes en un evento comunicativo real o ficcional (Hernanz 1999).

Los rasgos con peso negativo, tales como las nominalizaciones, las frases preposicionales y los adjetivos atributivos se complementan en la integración y densidad de la información (Chafe 1982, 1985, 1994; Janda 1985; Ciapuscio 1992; Halliday y Martín 1993; Hernanz 1999; Burdach 2000; Moyano 2000; Zarzalejos 2001). La dimensión que se perfila a partir del factor 2 se identifica con una sucesión cronológica de eventos principalmente en el pasado y descripción de todo lo que rodea a dichos eventos. Esto último se complementa con los indicadores de tiempo y lugar. Además, la fuerte incidencia de la deixis personal permite expresar la presencia de los protagonistas, ya sea por medio de la presencia de puntos de vista internos mediatizados por la conciencia del hablante o a través de puntos de vista externos, situados fuera de esta conciencia. Al primer punto de vista corresponden estados internos, perceptivos, privados y, al segundo, aquellos más públicos. En síntesis, este factor se asocia preferentemente con una sucesión de acontecimientos, que implica circunstancias de tiempo y lugar, como también la participación de las personas del discurso. Por tanto, el factor 2 permite conformar una dimensión funcional que hemos denominado Foco Narrativo.

La aplicación de los correspondientes procedimientos estadísticos entre registros revela que la Dimensión 2 Foco Narrativo presenta el mayor puntaje factorial (29,32) en el registro literario (CLL), a este le siguen los textos del corpus de entrevistas orales (CEO) con un peso positivo de 21,22 y, por último, se ubican los textos del corpus técnico-científico (CTC) con _comparativamente_ el menor peso positivo: 1,12. El resultado obtenido por el registro literario (CLL) es congruente con los rasgos co-ocurrentes y que tipifican a dicho discurso, esto es, la estructura narrativa y el mundo ficticio que instaura. La primera consiste en el ordenamiento de los eventos relatados en el pasado. Lo segundo, en cambio, en la serie de objetos, personajes, lugares, tiempos, eventos y leyes que rigen las relaciones entre ellos.

Dimensión 3: Foco Compromiso

Para interpretar el factor 3 se han considerado diecisiete rasgos que co-ocurren con valores superiores a 0,40. Los verbos privados (Biber 1988; Weber y Bentivoglio 1991), los verbos de volición (Gómez 1999) y los pronombres de primera y segunda persona singular remiten a los participantes de un acto de comunicación (Fernández 1999), concefcionamento a personajes que

como característica esencial de este factor la manifestación explícita del emisor del discurso (Calsamiglia y Tusón 1999; Crismore 1989). La presencia de marcas de pretérito indefinido sugiere que los verbos descritos anteriormente se refieren a acciones pasadas con un final temporal determinado, es decir, a construcciones que marcan el resultado de la acción que el verbo expresa. El modo indicativo, por su parte, hace referencia a hechos reales localizados en un tiempo verdadero (Criado de Val 1962; Contreras 1984) y expresa la modalidad declarativa experiencial característica del intercambio discursivo oral (Cepeda 2002). En relación a lo anterior, se postula también que el modo indicativo es un rasgo a través del cual se expresan estados o acciones como reales (Gómez Macker y Peronard 1988), es decir, este rasgo caracteriza intercambios lingüísticos cuyos referentes son hechos concretos en un aquí y un ahora determinados. En las oraciones sustantivas subordinadas, el empleo de la subjunción *que* está semánticamente condicionada: permite designar eventos o procesos que no se contemplan en su realización, sino en su resultado, o sea, como hechos ya establecidos, que se conciben como algo previo a la enunciación (Delbecque y Lamiroy 1999). La presencia de este rasgo sugiere que los interlocutores en el discurso, destacados antes, se manifiestan directamente sobre hechos específicos con un final temporal determinado. La aparición de frases infinitivas puede ser explicada a partir de su recurrencia en construcciones con los tipos de verbos que hemos presentado anteriormente, a saber, verbos privados y de volición. Aunque la función característica de estas construcciones es condensar información (Halliday 1994; Burdach 2000), en este caso esta función está subordinada a otros rasgos que aparecen con más peso tales como verbos privados y de volición, además de las desinencias y los pronombres personales en general. Por último, las frases preposicionales (Herranz 1999; Zarzalejos 2001; Harvey 2002), único rasgo que aparece con un valor negativo (-0,457) se asocian con una alta carga informativa, ya que expanden construcciones de mayor abstracción. La interpretación de los rasgos de alta ocurrencia, tales como verbos privados, verbos de volición y pronombres y desinencias verbales de primera persona en este factor constituyen la Dimensión 3 Foco Compromiso. Esta dimensión está asociada a textos en los que la intención y la actitud del emisor tienen mayor relevancia que el mensaje mismo; dicho de otra forma: esta dimensión caracteriza a textos en los que aparecen participantes reales que expresan intenciones y actitudes proposicionales frente a lo dicho. La clara presencia del "yo" a través de los pronombres y desinencias de primera persona es una evidencia clara de la voluntad de involucrarse en el discurso de manera explícita y asumir un rol preponderante.

Por último, podemos observar que el comportamiento de esta dimensión presenta su valor más alto en el CLL (9,52). En segundo lugar, se ubica el registro CEO con un valor de 7,7. A diferencia de estos puntajes, en cambio, aparece con un valor negativo (-1,99) el registro CTC. Todo ello permite señalar que la Dimensión 3 Foco Compromiso caracteriza más a los registros en los que aparecen participantes reales que expresan intenciones y actitudes frente al mensaje, a saber, el CEO y el CLL. El valor negativo del CTC en esta dimensión se explica por la naturaleza del discurso técnico-científico, el cual se focaliza en el contenido informativo del mensaje, lo que no exige la identificación de los interlocutores en el discurso, razón por la cual los rasgos que expresan compromiso comparativamente con los otros registros son menos frecuentes.

Dimensión 4: Foco Modalizador

Para el factor 4 se han considerado 8 rasgos co-ocurrentes con peso positivo y 1 rasgo con peso negativo. La co-ocurrencia de adjetivos predicativos, cuya función calificativo-atributiva se realiza de modo indirecto o no adjunto (González 2000), unido a formas activas con ser sugieren una relación con una función descriptiva (Gili Gaya 1980; Bosque 1990, 1999; Bassols y Torrent 1997). La co-ocurrencia de los rasgos anteriores se vuelve relevante al aparecer junto a subordinadas adjetivas que permiten atribuir al sustantivo cualidades muy complejas para las cuales no tiene el idioma adjetivos o participios léxicos (Gili Gaya 1980) y a adverbios de modo que denotan la manera en que se presentan los acontecimientos o se realizan las acciones (Bosque 1990). Se puede decir que ambos rasgos (subordinadas adjetivas y adverbios de modo) remiten a procesos de "*incidencia* de un segmento lingüístico sobre otro para configurar una unidad superior" (Hernández 2000b: 391); del mismo modo, su presencia permite la articulación de una sintaxis más expandida y de menor concisión. Por su parte, los atenuadores _marcadores informales y menos específicos de probabilidad e intensidad_ que aparecen en la subdimensión de la modalidad (Muller 1999)

hablante/escritor respecto al contenido declarado; en otras palabras, se destaca la visión particular del locutor respecto al contenido de los enunciados que emite (Calsamiglia y Tusón 1999; Hyland 1998). En síntesis: la conjunción sistemática de estos rasgos apunta a discursos con énfasis en *cómo* (*modus*) se dicen las cosas en relación al contenido mismo, a *lo dicho* (*dictum*).

El sustantivo ha sido clásicamente asociado a la función referencial. El peso negativo de este rasgo ($-0,494$) y los pesos positivos de los rasgos ya mencionados sugieren fuertemente la relación del factor 4 con una función más descriptiva que informativa, más expresiva que referencial. Por todo lo dicho anteriormente, como un modo de reflejar las funciones principales que aparecen en conjunto, hemos llamado a esta dimensión como Foco Modalizador, asociada principalmente a discursos con marcas actitudinales explícitas.

De la aplicación de los procedimientos estadísticos sobre la base de los puntajes factoriales, se desprende que los textos del corpus con más alto puntaje en esta dimensión corresponden al registro entrevistas orales (CEO), con 4,17. Los textos del corpus literario (CLL) presentan un puntaje bajo, de 1,31 en la dimensión Foco Modalizador (algunos de ellos tienen puntaje negativo). En cuanto al corpus técnico-científico (CTC), la mayoría de los textos muestra un puntaje negativo promedio de $-0,49$. Los puntajes obtenidos por CEO en la Dimensión 4, Foco Modalizador, son congruentes con algunas de las características comúnmente aceptadas de la oralidad. Entre estas, Narbona (2000) destaca el propósito socializador de la interacción cara a cara, que privilegia el uso de estrategias como la atenuación o la intensificación; así, los hablantes tratan de hacerse comprender y, al mismo tiempo, buscan influir sobre el oyente.

Dimensión 5: Foco Informacional

El factor 5 está conformado por la correlación significativa entre once rasgos. La presencia de los rasgos positivos co-ocurrentes como verbos modales de obligación y modo subjuntivo da cuenta de la necesidad y certeza de los juicios expresados, correspondiendo fundamentalmente a una modalidad deontica (Hyland 1998; Osorno 2000). Además, en el caso del subjuntivo, este remite a organizaciones de mayor complejidad sintáctica (Gili Gaya 1980) y a subordinación, cuya función es enmarcar la información del discurso (Delbeque y Lamiroy 1999; Galán 1999). Por otra parte, este mismo rasgo es usado para la expresión de la especulación subjetiva y el mandato (Criado de Val 1962; Gómez Macker y Peronard 1988). La co-ocurrencia de nominalizaciones, junto a participios en función de adjetivo y frases preposicionales, se presenta como indicador de integración y compactación de información altamente abstracta, propia de un discurso académico nominal (Burdach 2000; Picallo 1999; Chafe 1982, 1985; Janda 1985).

Los rasgos negativos que presentan un mayor peso en este factor son: las desinencias de tercera persona singular, las cuales se utilizan cuando existe el riesgo de malinterpretar la referencia, porque la información que aporta el contexto falla (Castellano 2000). El pretérito indefinido, el cual sitúa la acción de un suceso o evento en un espacio temporal acabado, es decir, es acción que ya no se puede volver a realizar (Contreras 1984). El verbo estar estativo activo, aunque no tiene valor especializado (Lorente 2002), forma parte de las llamadas unidades de conocimiento. Los verbos privados, que se destacan por expresar estados intelectuales o actos intelectuales no observables (Weber y Bentivoglio 1991). Los pronombres de negación, cuyo uso es preferentemente coloquial (Sánchez 1999; Tottie 1983). Y, por último, los verbos modales de volición, que dan cuenta de una modalidad dinámica anclada en la realidad (Langacker 1990) o de una modalidad orientada al participante, es decir, una modalidad que perfila el estatus del sujeto (Olbertz 1998). Cabe destacar que todos los rasgos positivos constituyentes de este factor se presentan preferentemente en discursos escritos y, en particular, en el artículo de investigación científica (Hyland 1998; Cornillie 2003; Criado de Val 1962; Harvey 2002; Burdach 2000).

En síntesis, los rasgos positivos agrupados en este factor están fundamentalmente orientados hacia la informatividad, entendida esta como la concentración de información en unidades y estructuras lingüísticas más pequeñas, que presentan los datos lo más conciso y precisamente posible (Biber 1988; Halliday y Martin 1993; Burdach 2000). Observamos así un conjunto de rasgos que dan cuenta de una alta carga informativa, asociada a una entrega regulada y

cantidades relevantes de datos y expresar significados complejos. Por otra parte, los rasgos negativos de este factor se orientan hacia una contextualización de los estados o actos intelectuales no observados de un sujeto, en un momento determinado (Ciapuscio 1992). La presencia de los rasgos antes mencionados permite distinguir claramente entre discursos con alta cantidad de información y, por ende, un mayor grado de abstracción de los que presentan menor cantidad de información, por ello es posible denominar a este factor 5 como dimensión Foco Informativo.

En relación al estudio estadístico comparativo, observamos que esta dimensión nos permite diferenciar significativamente el registro CTC (0,90) del CEO (_3,58) y del CLL (_4,28). De esta manera observamos que la configuración de los datos apoya la idea de que el registro CTC está conformado por textos cuyos rasgos lingüísticos dan cuenta de una alta cantidad de información, en tanto que los registros CLL y CEO presentan rasgos lingüísticos escasamente relacionados con la compactación y abstracción de la información.

Notas

* Investigación financiada parcialmente por FONDECYT a través del Proyecto 1020786.

Obras citadas

- Aijmer, K. 2002. "Modal adverbs of certainty and uncertainty in an English-Swedish perspective". Ed(s)., H. Hasselgard, S. Johansson, B. Behrens y C. Fabricius-Hansen. *Information structure in a cross-linguistics perspective*. Amsterdam: Netherlands. 97-113
- Alarcos Llorach, E. 1999. *Gramática de la Lengua Española*. Madrid: Espasa Calpe.
- Alcina, J. y J. Blecua. 1975. *Gramática Española*. Barcelona: Ariel.
- Alcoba, S. 1999. "La flexión verbal". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 4915-4992.
- Alvar, M. 2000. *Introducción a la Lingüística Española*. Barcelona: Ariel.
- Arianzen, C. 2001. "Las estrategias discursivas en el relato de Julio Ramón Ribeyro", [En línea]. Disponible en: <http://folk.uio.no/jmaria/lund/2000/textos/1.pdf>
- Arroyo, J. 2000. "La presencia de lo oral en la literatura: sobre la variable futuro verbal en una muestra del teatro español contemporáneo". Ed(s)., M. Muñoz, G. Fernández, A. Rodríguez y V. Benítez. *IV Congreso de Lingüística General*. Cádiz: Universidad de Cádiz. 267-282.
- Avila, A. 2000. "Hacia una caracterización gramatical del corpus de lengua hablada". Ed(s)., M. Muñoz, G. Fernández, A. Rodríguez y V. Benítez. *IV Congreso de Lingüística General*. Cádiz, Universidad de Cádiz. 151-158.
- Bassols, M. y A. Torrent. 1997. *Modelos textuales. Teoría y práctica*. Barcelona: Octaedro.
- Biber, D. 1986. "Spoken and written textual dimensions in English: Resolving the contradictory findings". *Language* 62: 384-414.
- Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. 1994. "Using register-diversified corpora for general language studies". Ed., S. Armstrong. *Using large corpora*. Cambridge: The MIT Press. 180-201.
- Biber, D. 2003. "Variation among university spoken and written registers: A new multi-dimensional analysis". Ed(s)., P. Leistyna y Ch. Meyer. *Corpus analysis. Language structure and language use*. Amsterdam: Rodopi. 47-70.

- Biber, D. y E. Finnegan. 1986. "An initial typology of English text types". Ed(s)., J. Aarts y W. Meijss. *Computer linguistics. Recent developments in the use computer corpora in English language research*. Amsterdam: Rodopi. 19-46.
- Biber, D.; S. Conrad y R. Reppen. 1998. *Corpus linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D.; S. Johansson, S. Conrad y E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow, UK: Longman.
- Biber, D.; R. Reppen, V. Clark y J. Walter. 2001. "Representing spoken language in university settings: The design and construction of the spoken component of the T2K-SWAL Corpus". Ed(s)., R. Simpson y J. Swales. *Corpus Linguistics in North America. Selections from the 1999 Symposium*. Ann Arbor: The University of Michigan Press. 48-57.
- Bosani, A. 2000. "Verbos de comunicación y discurso". Ed(s)., J. Bustos, P. Charaudeau, J. Girón, S. Iglesias y C. López. *Lengua, discurso, texto: I simposio internacional de análisis del discurso V.I*. Madrid: Visor. 253-262.
- Bosque, I. 1990. *Las categorías gramaticales*. Madrid: Síntesis.
- Bosque, I. 1999. "El nombre común". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 3-76.
- Burdach, A. 2000. "El léxico científico y técnico: Un recurso publicitario persuasivo". *Onomazein* 5: 189-208.
- Cabré, M. 1999. "Hacia una teoría comunicativa de la terminología: Aspectos metodológicos". *Revista Argentina de Lingüística* 15: 24-38.
- Cabré, M. 2000. *La terminología: Representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Universidad Pompeu Fabra.
- Cabré, M. 2002. "Textos especializados y unidades de conocimiento: Metodología y tipologización". Ed(s)., J. García y M. Fuentes. *Texto, terminología y traducción*. Barcelona: Almar. 122-187.
- Cademártori, Y. 2003. "La inscripción de las personas en textos de divulgación científica". *Revista Latinoamericana de Estudios del Discurso* 3(1): 9-28.
- Calsamiglia, H. 2000. "Decir la ciencia: las prácticas divulgativas en el punto de mira". *Revista Iberoamericana del Discurso y Sociedad* 2(2): 3-8.
- Calsamiglia, H. y A. Tusón. 1999. *Las Cosas del Decir. Manual de Análisis del Discurso*. Barcelona, Ariel.
- Camacho, J. 1999. "La Coordinación". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 2635-2694.
- Cassany, D.; C. López y J. Martí. 2000. "La transformación divulgativa de redes conceptuales científicas. Hipótesis, modelo y estrategias". *Revista Iberoamericana de Discurso y Sociedad* 2(2): 73-103.
- Castellano, A. 2000. "Ambigüedad y variación del pronombre personal sujeto". Ed(s)., M. Muñoz, G. Fernández, A. Rodríguez y V. Benítez. *IV Congreso de Lingüística General*. Cádiz: Universidad de Cádiz. 521-131.
- Cepeda, G. 2002. "Entonación, actitud y modalidad". *Estudios Filológicos* 37: 7-28.

D. Tannen. *Spoken and written language: Exploring orality and literacy*. Norwood, N.J.: Ablex., 35-53.

Chafe, W. 1985. "Linguistic differences produced by differences between speaking and writing". Ed(s)., D. Olson, N. Torrence y A. Hidayard. *Literature, language and learning: The nature and consequences of reading and writing*, Cambridge: Cambridge University Press. 105-123.

Chafe, W. 1994. *Discourse, consciousness and time*. Chicago: The University of Chicago Press.

Chafe, W. y J. Danielewics. 1987. "Properties of spoken and written language". Ed(s)., R. Horowitz y J. Samuels. *Comprehending oral and written language*. New York: Academic Press. 83-115.

Ciapuscio, G. 1992. "Impersonalidad y desagentivación en la divulgación científica". *Lingüística Española Actual* 2: 183-205.

Ciapuscio, G. 2000. "Hacia una tipología del discurso especializado". *Revista Iberoamericana de Discurso y Sociedad* 2(2): 39-71.

Ciapuscio, G. 2003. *Textos especializados y terminología*. Barcelona: Universidad Pompeu Fabra.

Ciapuscio, G. y I. Kuguel. 2002. "Hacia una tipología del discurso especializado". Ed(s)., J. García y M. Fuentes. *Texto, terminología y traducción*. Salamanca: Almar. 37-73.

Conrad, S. y D. Biber. 2001. "Multi-dimensional methodology and the dimensions of register variation in English". Ed(s)., S. Conrad y D. Biber. *Variation in English. Multidimensional Studies*. 13-42.

Contreras, C. 2000. "Unidad temática y variedad textual: un tópico social en tres relatos orales". *Estudios Filológicos* 35: 23-39.

Contreras, L. 1984. *Nuevo texto gramatical práctico, reestructurado, ampliado y actualizado conforme a las recientes normas actualmente en vigencia de la Real Academia Española*. Santiago: Instituto Geográfico Militar.

Cornillie, B. 2003. "Subjetivización, predicciones de anclaje y modales del español". *Foro Hispánico* 23: 41-57.

Criado de Val, M. 1962. *Fisonomía del idioma español*. Madrid: Aguilar.

Crismore, A. 1989. *Talking with readers. Metadiscourse as Rethorical Act*. New York: Peter Lang Publishers.

De Kock, J. y C. Gómez. 2002. *Gramática Española. Enseñanza e Investigación. Apuntes metodológicos*. Salamanca: Ediciones Universidad Salamanca.

De Miguel, E. 1999. "El aspecto léxico". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 2977-3060.

Delbecque, N. y B. Lamiroy. 1999. "La subordinación sustantiva: las subordinadas enunciativas en los complementos verbales". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 1965-2082.

Demonte, V. 1997. *La subordinación sustantiva*. Madrid: Cátedra.

Demonte, V. y S. Varela. 1997. "Los infinitivos nominales eventivos del español". *Revista Signo y Seña* 7: 123-156.

- Fernández, O. 1999. "El pronombre personal. Formas y distribuciones. Pronombres átonos y tónicos". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 1209-1274.
- Flowerdew, L. 2004. "The argument for using English specialized corpora to understand academic and professional language". Ed(s)., U. Connor y A. Thomas. *Discourse in the Professions. Perspectives from corpus linguistics*. Philadelphia: John Benjamins. 11-33.
- Fontanella, M. 1999. "Sistemas pronominales de tratamiento usados en el mundo hispánico". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 1399-1426.
- Francis, N. 1979. "A tagged corpus: problems and prospects". Ed(s)., S. Greenbaum, G. Leech y J. Svartvik. *Studies in English linguistics for Randolph Quirk*. London: Longman. 192-209.
- Galán, C. 1999. "La subordinación causal y final". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 3597-3642.
- Genette, G. 1972. *Figures III*. París: Le Seuil (Points).
- Gili Gaya, S. 1980. *Curso Superior de Sintaxis Española*. Barcelona: Vox.
- Gómez Macker, L. y M. Peronard. 1988. *El lenguaje humano. Léxico fundamental para la iniciación lingüística*. Valparaíso: Ediciones Universitarias de Valparaíso.
- Gómez, L. 1999. "Los verbos auxiliares. Las perifrasis verbales de infinitivo". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 3323-3390.
- González, J. 2000. "Morfología Nominal". Dir., M. Alvar. *Introducción a la Lingüística Española*. Barcelona: Ariel. 177-190.
- Hair, J., R. Anderson, R. Tatham y W. Black. 1999. *Análisis multivariante*. Madrid: Prentice Hall.
- Halliday, M. 1985. *An introduction to functional grammar*. London: Edward Arnold.
- Halliday, M. 1993. "On language and physical science". Ed(s)., M. Halliday y J. Martin. *Writing science. Literacy and discursive power*. Pittsburgh: University of Pittsburgh Press.
- Halliday, M. 1994. *An Introduction to Functional Grammar*. London: Arnold.
- Halliday, M y J. Martin. 1993. *Writing science. Literacy and discursive power*. Pittsburgh: University of Pittsburgh Press.
- Harvey, A. 2002. "Representación e imagen del quehacer científico en los Medios de Comunicación". Ed., G. Parodi. *Lingüística e Interdisciplinariedad. Ensayos en honor a Marianne Peronard*. Valparaíso: Ediciones Universitarias de Valparaíso. 335-353.
- Hernández, C. 1996. *Gramática Funcional del Español*. Madrid: Gredos.
- Hernández, C. 2000a. "Morfología del Verbo. La auxiliaridad". Dir., M. Alvar. *Introducción a la Lingüística Española*. Barcelona: Ariel. 195-213.
- Hernández, C. 2000b. "Sintaxis: La subordinación". Dir. M., Alvar. *Introducción a la Lingüística Española*. Barcelona: Ariel. 391-407.
- Hernanz, M. 1999. "El Infinitivo". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de*

- Horowitz, R. y J. Samuels. 1987. "Comprehending oral and written language: Critical contrasts for literacy and schooling". Ed(s)., R. Horowitz y J. Samuels. *Comprehending oral and written language*. San Diego: Academic Press. 1-52.
- Hyland, K. 1998. *Hedging in scientific research articles*. Philadelphia: John Benjamins.
- Janda, R. 1985. "Note-taking as simplified register". *Discourse Processes* 84: 437-454.
- Johansson, S. 1991. "Times change, and so do corpora". Ed(s)., K. Aijmer y B. Altenberg. *English Corpus Linguistics. Studies in honor of Jan Svartvik*. London: Longman. 305-314.
- Jurafsky, D. y J. Martin. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. New Jersey: Prentice Hall.
- Kaiser, D. 2002. "La presencia del autor en los textos académicos: un estudio contrastivo de trabajos de estudiantes de Venezuela y Alemania". *Boletín de Lingüística* 17: 53-68.
- Kittredge, R. 1982. "Variation and homogeneity of sublanguages". Ed(s)., R. Kittredge y J. Lehrberger. *Sublanguages. Studies of language in restricted semantic domains*. Berlin: Walter de Gruyter. 145-189.
- Kovacci, O. 1993. "La didáctica de la lengua materna. Experiencias en Argentina". *Actas del I Congreso Internacional sobre la Enseñanza del Español*. Madrid: CEMIP.
- Kovacci, O. 1999. "El Adverbio". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 705-786.
- Lakoff, G. 1972. "A study in meaning criteria and the logic of fuzzy concepts". *Linguistics Society* 8: 183-288.
- Langacker, R. 1990. Subjectification. *Cognitive Linguistics* 1(1): 5-38.
- Leech, G. 1991. "The state of the art in corpus linguistics". Ed(s)., K. Aijmer y B. Altenberg. *English Corpus Linguistics. Studies in honor of Jan Svartvik*. London: Longman. 8-29.
- Leech, G. 1992. "Corpora and theories of linguistic performance". Ed., J. Svartvik. *Directions in corpus linguistics. Proceeding of Nobel symposium*. Berlin: Mouton de Gruyter. 105-122.
- Leech, G. 2000. "Grammars of Spoken English: New Outcomes of Corpus-Oriented Research". *Language Learning* 50(4): 275-724.
- Lledó, E. 1995. "Usos lingüísticos y género". *Textos de didáctica de la lengua y la literatura* 6: 29-34.
- Longrace, R. 1983. *The grammar of discourse*. New York: Plenum.
- López, C. 2002. "Aproximaciones al análisis de los discursos profesionales". *Signos* 35 (51-52): 195-215.
- Lorente, M. 2002. "Verbos y discurso especializado". [En línea]. Disponible en: <http://elies.rediris.es/elies16/Lorente.html>
- Louwerse, M., P. McCarthy, D. McNamara y A. Graesser. 2004. "Variation in language and cohesion across written and spoken registers". Ed(s)., K.D. Forbus, D. Gentner y T. Regier. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Mahwah, NJ.: Lawrence Erlbaum.
- Marcos Marín, F. 1975. *Aproximación a la gramática española*. Madrid: Cincel.

/markkane.html

Martín Zorraquino, M. y J. Portolés. 1999. "Los marcadores del discurso". Coord(s)., I. Bosque y V. Demonte. *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe. 4051-4213.

Martín, G. 1986. *Curso de redacción*. Madrid: Paraninfo.

Mendikoetxea, A. 1999a. "Construcciones inacusativas y pasivas". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 1575-1630.

Mendikoetxea, A. 1999b. "Construcciones con se: Medias, pasivas e impersonales". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 1631-1722.

Menéndez, S. 1999. "El discurso del libro de texto: una propuesta estratégico-pragmática". *Revista Iberoamericana de discurso y sociedad* 1(2): 85-104.

Moliner, M. 1986. *Diccionario de uso del español*. Madrid: Gredos.

Montolío, E. 2001. *Conectores de la lengua escrita*. Barcelona: Ariel Practicum.

Moreno, A. 1998. *Lingüística computacional: Introducción a los modelos simbólicos, estadísticos y biológicos*. Madrid: Síntesis.

Moyano, E. 2000. *Comunicar ciencia. El artículo científico y las comunicaciones a congreso*. Buenos Aires: Universidad Nacional de Lomas de Zamora.

Narbona, A. 2000. "Sintaxis Coloquial". Dir., M. Alvar. *Introducción a la Lingüística Española*. Barcelona: Ariel. 463-476.

Oakes, M. 1998. *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.

Olbertz, H. 1998. *Verbal Periphrases in a Functional Grammar of Spanish*. Berlín-New York: Mouton de Gruyter.

Osorno, M. 2000. "Categorías gramaticales de modalidad y aspecto en la lengua de señas colombiana". [En línea]. Disponible en: <http://www.geocities.com/fenascolv/ponencia151.html>

Otal, J., I. Fontanet y V. Codina. eds. 1997. *Estudios de Lingüística Aplicada*. Castellón: UJI. 611-618.

Palmer, F. 1974. *The English verb*. London: Longman.

Parodi, G. 2004. "Textos de especialidad y comunidades discursivas técnico-profesionales: Una aproximación basada en corpus computarizado". *Revista Estudios Filológicos* 39: 7- 36.

Parodi, G. ed. 2005a. *Discurso especializado e instituciones formadoras*. Valparaíso: Editorial Universitaria de Valparaíso.

Parodi, G. 2005b. "La comprensión del discurso especializado escrito en ámbitos técnico-profesionales: ¿aprendiendo a partir del texto?" *Revista Signos* 38(58): 221-267.

Parodi, G. 2006. *Lingüística de corpus*. Buenos Aires: Eudeba (en prensa).

Parodi, G. y A. Gramajo. 2003. "Los tipos textuales del corpus PUCV-2003: Una aproximación multiniveles". *Revista Signos* 36(54): 207-223.

Parodi, G. y R. Venegas. 2004. "Bucólico: Aplicación computacional para el análisis de textos. Hacia un análisis de rasgos de la informatividad". *Revista Lingüística y Literatura* 15: 223-251.

- Picallo, M. 1999. "La estructura del sintagma nominal: La nominalización y otros sustantivos con complementos argumentales". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 363-394.
- Portolés, J. 1998. *Marcadores del discurso*. Barcelona: Ariel Practicum.
- Quirk, R., S. Greenbaum, G. Leech y J. Svartvik. 1985. *A grammar of contemporary English*. London: Longman.
- R.A.E. 1973. *Esbozo de una Nueva Gramática de la Lengua Española*. Madrid: Espasa Calpe.
- Reppen, R., S. Fitzmaurice y D. Biber. 2002. *Using corpora to explore linguistic variation*. Amsterdam: John Benjamins.
- Sánchez, C. 1999. "La Negación". Coord(s)., I. Bosque y V. Demonte. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe. 2561-2634.
- Sinclair, J. 1982. "reflections on computer corpora in English language research". Ed., S. Johansson. *Computer corpora in English language research*. Bergen: Norwegian Computing Centre for the Humanities. 1-6.
- Sinclair, J. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stubbs. M. 1996. *Text and corpus analysis. Computer-assisted studies of language and culture*. Massachusetts: Blackwell.
- Svartvik, J. 1992. "Corpus linguistics comes of age". Ed., J. Svartvik. *Directions in corpus linguistics. Proceeding of Nobel symposium*. Berlin: Mouton Gruyter. 7-13.
- Tognini- Bonelli, E. 2001. *Corpus linguistics at work*. Amsterdam: John Benjamins.
- Tottie, G. 1983. *Much about "not" and "nothing": A study of the variation between analytic and synthetic negation in contemporary American English*. Lund: CWK Gleerup.
- Vásquez, G. 2001. *Guía didáctica del discurso académico escrito. ¿Cómo se escribe una monografía?* Madrid: Edinumen.
- Weber, E. y P. Bentivoglio. 1991. "Verbs of cognition in spoken Spanish: a discourse profile". Ed(s)., S. Fleischman y L. Waugh. *Discourse Pragmatics and the Verb: The Evidence from Romance*. London: Routledge.
- Weinrich, H. 1974. *Estructura y función de los tiempos en el lenguaje*. Madrid: Gredos.
- Zarzalejos, J. 2001. "En el titular está el periódico". [En línea]. Disponible en: <http://cvc.cervantes.es/obref/congresos/valladolid/ponencias/elespanolenlasociedad/1laprensaenespanol/zarzalejosj.htm>