

El discurso especializado escrito en el ámbito universitario y profesional: Constitución de un corpus de estudio

Specialized written discourse at university and professional domains: Composition of a corpus

Giovanni Parodi

Pontificia Universidad Católica de Valparaíso, Chile

RESUMEN

Puede decirse que la alfabetización lingüística del discurso especializado (tanto académico como profesional) ha comenzado solo recientemente a ser explorada en Chile. Todo avance en esta línea, a nuestro entender, debe partir de una mirada empírica profunda de los datos reales. Así, una forma de acceder a los géneros especializados escritos de la academia es partir desde el supuesto de que los materiales escritos empleados por los estudiantes en su formación universitaria revelan información relevante acerca del discurso y del conocimiento disciplinar. En el presente artículo aportamos antecedentes acerca de un proyecto en marcha en la Pontificia Universidad Católica de Valparaíso, Chile, centrado en la recolección, construcción y descripción de un corpus de discurso escrito a partir de textos recolectados en el ámbito académico y en el profesional en cuatro áreas disciplinares del conocimiento: Química Industrial, Ingeniería en Construcción, Trabajo Social y Psicología. De este modo, aportamos una revisión de los conceptos de discurso especializado, discurso académico y discurso profesional. En la parte empírica establecemos los procedimientos de recolección y constitución del Corpus Académico PUCV-2006, el cual alcanza cerca de los 100 millones de palabras. También se avanza preliminarmente una tipologización de los 652 textos recolectados.

Palabras Clave: Discurso académico, discurso profesional, discurso disciplinar, lingüística de corpus, Corpus PUCV-2006.

ABSTRACT

It could be said that the specialized discourse literacy (academic and professional) has just began to be explored recently in Chile. Advances in this line must start from a deep and empirical analysis of real data. Thus, one way to access to the specialized written genres employed by the academia is to begin from the tenant that all materials read by students in university training reveal relevant data about disciplinary discourse and knowledge. This article gives information about a research Project, in its first steps, at Pontificia Universidad Católica de Valparaíso, Chile, focused on the collection, construction and description of an academia corpus based on texts collected in the academic and professional areas of four disciplinary domains of knowledge: Industrial Chemistry, Construction Engineering, Social Work, and Psychology. So, a revision of the concepts of specialized, academic, and professional discourse is presented. Then, the procedures of collecting and organizing the Academic Corpus PUCV-2006 are described, which is formed by

almost 100 million words. Also, a preliminary text typology of the 652 text corpus is advanced.

Key Words: Academic discourse, professional discourse, disciplinary discourse, Corpus linguistics, Corpus PUCV-2006.

INTRODUCCIÓN

Un asunto de gran impacto durante la última década ha sido el interés creciente por el estudio de la variación lingüística a través del discurso de las disciplinas científicas. Este foco en la diversidad se constituye como un factor explicativo de la construcción divergente del conocimiento especializado al interior de las comunidades discursivas o de práctica. La percepción bien fundada de que no existe tal cosa como un discurso disciplinario unitario y de que es más conveniente hablar de **discursos disciplinarios** en forma plural (Hyland, 2000), se vuelve cada vez un hecho más reconocido entre los investigadores (Bhatia, 2004). Los hallazgos empíricos -desde diversos enfoques lingüísticos- han documentado la relevancia de los análisis basados en corpus como un modo de avanzar y describir en detalle las variaciones lingüísticas y discursivas a través de las disciplinas y a través de los tipos de textos prototípicos (Biber, 1988, 1994, 2005, 2006; Martin & Veel, 1998; Wignell, 1998; Williams, 1998; Swales 2000, 2004; Flowerdew, 2002; Parodi 2005a, 2006, 2007).

La teoría del género ha emergido en los últimos años con gran fuerza como una forma de contribuir a la comprensión del modo en que en el discurso y (a través de este) el conocimiento disciplinar se construye y se emplea en contextos académicos y profesionales. No obstante ello, los avances y desarrollos en el nivel de análisis y descripción empírica en que se estudian los rasgos lingüísticos, sus funciones y distribuciones en el discurso natural en contextos académicos institucionales y profesionales han iniciado un camino fructífero solo recientemente, en especial desde el enfoque de la lingüística de corpus. Resulta así interesante y estimulante que existan diversas preguntas no respondidas acerca de, por ejemplo, los patrones característicos del uso lingüístico en español especializado disciplinar y del tipo de conocimiento asociado a ellos en áreas especializadas tanto en ámbitos académicos de formación universitaria como en entornos profesionales.

Un modo de aproximación a los géneros especializados escritos en la academia es partir desde el supuesto de que los materiales escritos empleados por los estudiantes en su formación universitaria cotidiana revelan tanto los tipos de textos como los rasgos representativos de su disciplina. En este sentido, partir de las lecturas obligatorias en algunas áreas disciplinares constituye un modo de construir un panorama del universo lingüístico y discursivo a través del cual estos sujetos en formación acceden al conocimiento disciplinar y, en parte, a través de ello se integran paulatinamente a sus respectivas comunidades. Del mismo modo, la entrevista a profesionales en ejercicio, egresados de estos mismos ámbitos, y que se desempeñan laboralmente dentro del perfil disciplinar, permite conocer los medios y formatos de comunicación escrita que emergen de las prácticas discursivas institucionalizadas.

En el presente artículo entrego antecedentes acerca de un proyecto en marcha (FONDECYT 1060440) en la Pontificia Universidad Católica de Valparaíso, Chile, centrado en la recolección, construcción y descripción de un corpus de discurso

escrito a partir de textos recolectados en el ámbito académico y en el profesional en cuatro áreas disciplinares del conocimiento: Química Industrial, Ingeniería en Construcción, Trabajo Social y Psicología. En la primera parte del trabajo se entregan algunos referentes teóricos que enmarcan la investigación. En la segunda parte, se establecen los parámetros de la constitución del corpus y una proyección de su tamaño. En la tercera se entrega una primera descripción cuantitativa del corpus académico El Grial PUCV-2006 y se realiza una descripción general de los tipos de textos que se han encontrado. Cierran el artículo algunas proyecciones a partir del proyecto de investigación en curso.

1. Antecedentes

1.1. Marco general del proyecto

Durante los últimos años, nuestras investigaciones han estado focalizadas en un terreno poco o nada indagado en nuestro medio, con carencias muy evidentes y de proyecciones importantes en el escenario actual chileno e internacional, a saber:

a) el estudio y análisis lingüístico del discurso especializado técnico-profesional secundario escrito en liceos de Valparaíso, Chile (Parodi, 2005a, 2006, 2007), y

b) la indagación acerca de la capacidad de comprensión lingüística que los sujetos egresados del sistema de formación técnico-profesional secundaria han logrado alcanzar en tres ámbitos del conocimiento (Marítimo, Industrial y Comercial) (Parodi, 2005b).

Dado que -como bien se sabe- es escaso aún el conocimiento de materiales y plataforma didáctica para el mejoramiento de las competencias textuales y discursivas de nuestros escolares del sistema tradicional chileno, se hace evidente que el grado de conocimiento y análisis del material escrito empleado en la formación profesional secundaria ha sido poco explorado en nuestro país. Nuestro aporte ha pretendido abrir una brecha en este terreno y brindar luces acerca de los tipos de materiales empleados (Parodi & Gramajo, 2003), así como de una caracterización lingüístico/textual inicial de los mismos (Parodi, 2005a, b, 2006).

Ahora bien, ya que el estudio que estamos finalizando acerca del discurso técnico especializado y su comprensión lingüística nos ha permitido conformar un panorama inicial de la situación de la formación lingüística de un grupo de alumnos de liceos técnicos secundarios, hemos decidido avanzar en nuestras pesquisas hacia dos ámbitos complementarios entre sí, vinculados de cierto modo a la educación secundaria diferenciada. Ellos son el ámbito académico universitario y el ámbito profesional en el terreno laboral. Nos interesa recolectar y estudiar los textos escritos que leen los alumnos universitarios (y a través de los que se vehicula el conocimiento disciplinar) en cuatro áreas académicas y los textos escritos que circulan, se constituyen en eje de la comunicación laboral y se emplean en las mismas cuatro áreas profesionales en la práctica laboral cotidiana.

Puede decirse -sin riesgo a equivocación- que la alfabetización lingüística del discurso especializado (tanto académico como profesional) aún no encuentra caminos fructíferos en Chile. Es un tanto diferente la situación en Norteamérica y en Europa, ya que en algunos países de esos ejes ya se cuenta con programas de

lectura y escritura especializada en diferentes áreas del conocimiento y con intensos programas de apoyo lingüístico a determinados ámbitos laborales. Un rápido análisis de las mallas curriculares de la mayoría de las carreras de universidades chilenas revela que las áreas disciplinares concentran la mayor parte (si no todo) el currículo y son escasos los espacios, como parte del currículo obligatorio, que se otorga a asignaturas centradas en la lectura y la escritura del discurso general y/o especializado. Un ejemplo emblemático en Chile lo constituye el proyecto que viene desarrollando la Facultad de Letras de la Pontificia Universidad Católica de Chile, coordinado por la Dra. Marcela Oyanedel, en su vinculación con algunas áreas disciplinares de formación académica y el entrenamiento de profesores de las propias especialidades en cuestiones de lengua escrita. Otro ejemplo, es lo que acontece en Argentina, en la Universidad de Buenos Aires, desde la Facultad de Filosofía y Letras en el proyecto de lectura y escritura en la universidad, coordinado por la Dra. Elvira Narvaja de Arnoux.

Este terreno emergente -a nuestro entender- se encuentra aún escasamente explorado desde estudios de corte lingüístico en nuestro país. Por una parte, este exhibe rasgos lingüísticos y discursivos muy interesantes para el análisis y, por otra, en vista de la urgente necesidad de alfabetización científico-profesional que enfrenta la sociedad actual, ofrece una oportunidad crucial para apoyar el desempeño lingüístico más responsable, competente y autónomo de los estudiantes y de los profesionales del mundo globalizado moderno. Tanto en este como en otros ámbitos, el lenguaje escrito se constituye en eje fundamental para la construcción y el acceso a la información y, mejor aún, al conocimiento altamente especializado. En esta línea, puede estar de más referirse al papel del lenguaje escrito como medio vital de acceso al conocimiento y, en definitiva, hacia la autonomía y mayor libertad de desarrollo personal; no obstante ello, se hace evidente que una parte importante del acceso a la información escrita se articula a través del manejo lingüístico que, entre otros, se alcanza con el adecuado desarrollo de una competencia textual escrita.

Una pregunta de no menor relevancia es aquella que dice relación con la responsabilidad de alfabetizar en el discurso especializado escrito a miembros universitarios en formación de comunidades discursivas disciplinares. Esto quiere decir que nos preocupa quién debe hacerse cargo o quién debería asumir el rol de capacitador de los profesionales en formación, no alfabetizados en lengua escrita especializada para alcanzar una construcción del conocimiento disciplinar y un desempeño óptimo en el mundo laboral. Como se ha dicho, las universidades han venido asumiendo su cuota de responsabilidad; se puede decir que estas instituciones de educación superior se han sensibilizado paulatinamente a los problemas de los alumnos ingresantes y de otros pronto a egresar. Así, en algunas universidades chilenas se han creado cursos en lectura y escritura en forma de asignaturas optativas. Sin lugar a dudas, esto no resuelve los problemas. La mayoría de las veces, los docentes a cargo de estos cursos desconocen el discurso especializado de la disciplina y no cuentan con un conocimiento que logre apoyar certeramente a estos alumnos. En otros casos, solo pequeños grupos de alumnos se motivan a inscribirse en dichos cursos optativos. A pesar de ello, es innegable el mérito de buscar caminos y desarrollar iniciativas pioneras. El problema reside, entre otras muchas variables, en que este conocimiento inicial, en algunos casos, resulta ser muy parcelado y no logra afincarse de manera significativa dentro de la práctica académica, ni menos en la profesional. Tampoco parece posible que así se logre el desarrollo definitivo de habilidades lingüísticas altamente específicas.

Hoy en día, es ya un lugar común que los académicos universitarios nos sigamos sorprendiendo de que nuestros alumnos -en todas las áreas del saber- no comprendan lo que leen. Incluso también, está claro que lo mismo acontece en los niveles de magíster y de doctorado. Lo propio sucede en empresas e instituciones de diversas ramas. Más complejo aún es el escenario en lo que respecta a expresarse adecuadamente por escrito en términos académicos y profesionales. Al mismo tiempo, también es escaso o nulo lo que sabemos acerca de los géneros discursivos escritos que circulan en la academia y en determinados ámbitos profesionales. Es aún menos lo que conocemos acerca de la debida progresión y secuencialidad de materiales lingüísticos que deben apoyar la construcción del conocimiento disciplinar y profesional.

Desde este escenario, en nuestra opinión, es imposible avanzar en tareas como las anteriormente señaladas si no conocemos el material escrito que leen los alumnos de algunas disciplinas universitarias y lo indagamos y contrastamos con los textos escritos que esos mismos profesionales enfrentan en su vida profesional, una vez egresados de la universidad. Resulta vital saber si la formación universitaria entrega sus conocimientos disciplinares y profesionales a través de textos que, por un lado, representan fidedignamente el ámbito y, por otro, presentan vinculación con los tipos de textos y rasgos lingüístico/textuales de los materiales escritos que esos profesionales encontrarán cotidianamente en su vida laboral. También es necesario saber si en la formación universitaria los alumnos tienen conciencia de que algunos materiales escritos formarán parte de su trabajo profesional y deberán aprender no solo a leerlos sino también a escribirlos.

Cabe señalar que en Chile ha habido importantes aproximaciones al ámbito del discurso académico, pero desde el denominado Inglés con Propósitos Específicos. Por ende, los textos estudiados han correspondido a muestras de material escrito en inglés técnico-científico, con rasgos peculiares y en el marco de la enseñanza/aprendizaje de una lengua extranjera (también en francés y alemán, principalmente). En este contexto, los trabajos pioneros y señeros son los realizados por Mercedes Horsella, Anamaría Harvey, Fresia Pérez y Gerda Sindermann (Harvey, 1987, 1994; Harvey & Horsella, 1988; Horsella, Pérez & Sindermann, 1994; Horsella & Pérez, 1996).

Más recientemente, el discurso académico universitario en lengua española ha sido indagado por Harvey, Oyanedel y Núñez a través de un proyecto FONDECYT que les ha permitido adentrarse en un tipo de textos que ellas han denominado □informe□ (Harvey, 2004; Núñez & Espejo, 2004; Harvey, 2005; Núñez, Muñoz & Mihovilovic, 2006).

En este contexto, resulta relevante puntualizar ciertos supuestos fundamentales para esta investigación. Por una parte, la aproximación que tenemos al objeto 'discurso' es decididamente interdisciplinaria de índole psicosociolingüística. Nos interesan los textos en su carácter de unidad lingüística, pero como objetos inmersos en un contexto cognitivo y social, en donde su función está determinada cognitiva y contextualmente. De este modo, para nosotros los textos son unidades lingüísticas pero con un cierre de sentido y en virtud de productores/ hablantes y de lectores oyentes en contextos particulares y con propósitos definidos, con conocimientos previos contruidos desde nuestra cognición humana en contextos sociales específicos y almacenados en nuestra mente. En otras palabras, concebimos los textos como producto de nuestra cognición y contexto y, a la vez, se constituyen en los soportes que -en parte- nos ayudan a construir nuestro mundo y nuestro entorno.

En lo que sigue, puntualizamos otras nociones centrales que enmarcan el proyecto de investigación.

1.2. Discurso especializado: académico y profesional

Como un marco general, en primer lugar, abordamos la noción que consideramos engloba los objetos de estudio de esta investigación, a saber, el discurso especializado (DE). A continuación, nos centraremos en una discusión crítica acerca de los conceptos de discurso académico (DA) y de discurso profesional (DP).

Como bien se sabe, el uso del término "discurso especializado" se encuentra, en la actualidad, ampliamente aceptado por los estudiosos del lenguaje; no obstante, se debe reconocer que su utilización no surgió sino solo hace unos pocos años. Se ha hecho evidente que desde el inicio, este término no ha sido utilizado con el mismo sentido. En un comienzo vino, simplemente, a reemplazar a la noción de "lenguas con propósitos específicos". En la actualidad, en cambio, la noción de DE se enfrenta, por un lado, de modo más específico y, por otro, también se concibe desde una mirada más amplia y globalizadora (Parodi, 2005a). Así, se entiende que el DE incluye un conjunto variado de tipos de textos pero con ciertos rasgos muy prototípicos. Es justamente esta idea de heterogeneidad de textos dentro de una escala de gradación la que rescata Parodi (2005a) al aproximarse a la noción de DE. Así, el DE debe necesariamente entenderse como un *continuum*, en el que se alinean textos a lo largo de una gradiente diversificada que va desde un alto hasta un bajo grado de especialización. Debido a esto mismo, es que concebimos al DE como un hiperónimo de las nociones de DA y DP.

En este sentido, focalizando en lo característico, Parodi (2005a: 26) define el DE a partir de una serie de rasgos caracterizadores:

"como un conjunto de textos que se distinguen y se agrupan por una co-ocurrencia sistemática de rasgos lingüísticos particulares en torno a temáticas específicas no cotidianas en los cuales se exige experiencia previa disciplinar de sus participantes (formación especializada dentro de un dominio conceptual particular de la ciencia y de la tecnología); por otra, son textos que revelan predominantemente una función comunicativa referencial y circulan en contextos situacionales particulares; todo ello implica que sus múltiples rasgos se articulan en singulares sistemas semióticos complejos y no de manera aislada y simple".

Un aspecto que no se rescata directamente de esa definición (pero que sí aparece en los rasgos prototípicos propuestos por Parodi, 2005a) es el empleo de un léxico especializado.

Resulta altamente relevante rescatar este rasgo, dado que se constituye en uno de los principales aspectos señalados por otros autores al caracterizar este tipo de discurso (Cabré, 1993; Burdach, 2000; Cabré, Doménech, Morel & Rodríguez, 2001; Ciapusio, 2003; Cabré & Gómez, 2006).

Ahora bien, según nuestra opinión, los discursos académicos y profesionales se operacionalizan a través de un conjunto de textos que se pueden organizar a través de un *continuum* en el que se van concatenando desde el discurso escolar general, hacia el académico universitario y el profesional en el medio laboral. Esto se presenta gráficamente a través de la Figura 1:

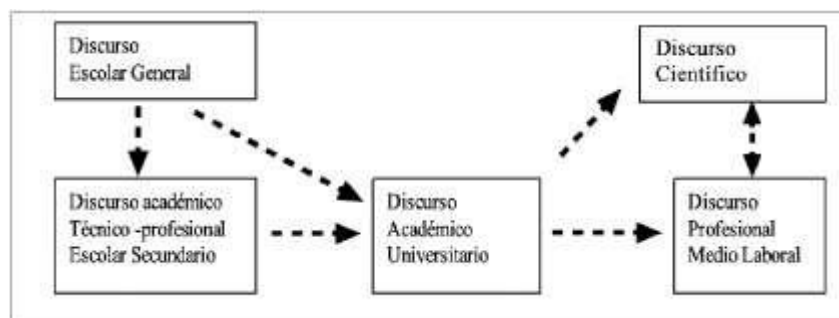


Figura 1. *Continuum* de textos en ámbitos académicos y profesionales.

Como se dijo, a través de la Figura 1, se muestra nuestra concepción del discurso en los ámbitos académicos y profesionales y el *continuum* que siguen de manera secuencial, o más bien, que debieran seguir en un proceso de alfabetización permanente. También allí se da cuenta del discurso científico que -en parte- proviene del DA y que, a su vez, se vincula e interactúa con el D P.

1.2.1. Discurso académico

No cabe duda que quien se aproxima por primera vez al estudio del denominado DA encuentra una diversidad de enfoques y perspectivas que dificultan una comprensión inicial del campo. Por ello, coincidimos con Flowerdew (2002) en que la investigación acerca de qué es el DA es escasa y poco sistemática. Cuando se aborda el estudio de este tipo de discurso resulta muy válido enfrentar las siguientes preguntas: a) ¿existen algunos criterios para definir este tipo de discurso de modo certero?; de ser así: b) ¿de qué tipo son estos criterios? En lo que sigue, realizaremos una revisión del DA desde tres aproximaciones: 1) una comunicativa funcional, 2) otra contextual, y 3) otra textual.

En términos comunicativo-funcionales, el DA se caracteriza por la predominancia de la descripción con propósitos persuasivos y didácticos. Es, además, un tipo de discurso que expresa credibilidad y prestigio, es decir, un discurso autorizado acerca del tema que trata. Al respecto, Hyland (2000: 17) sostiene que este discurso apunta a:

□□an authorized understanding of the world (and how it can be perceived and reported) which acts to reinforce the theoretical convictions of the discipline and its right to validate knowledge□.

Desde esta mirada, el DA tiene como propósito cristalizar y difundir un conocimiento disciplinar y apoyar a la formación de los miembros en formación de esa comunidad de especialistas. Al respecto, Vilá (2000) concibe al DA no como un discurso predominantemente descriptivo sino más bien explicativo, con una función didáctica fundada en demostraciones y justificaciones. Al mismo tiempo, según Tapia, Burdiles y Arancibia (2003), los textos producto de la actividad discursiva académica, se caracterizarían por ser eminentemente referenciales-representativos y por tener como finalidad ser soporte y transmisores del conocimiento.

Como se aprecia, no existe consenso acerca de la función discursiva predominante en el DA, aunque sí se percibe un común denominador en la función general que tiende al uso de secuencias argumentativas-demostrativas e incluso persuasivas. En este sentido, es posible reconocer una estructura basada en secuencias

textuales descriptivas y argumentativas con un alto grado de generalización y abstracción semántica en que prima un propósito divulgativo, didáctico y de formación.

Otro criterio es el contextual. Desde esta mirada, el DA es aquél que se utiliza en contextos académicos o, dicho de otra forma, con propósitos académicos (Kennedy, 2001; Flowerdew, 2002; Dudley-Evans & St. John, 2006). Sabemos, sin embargo, que los propósitos académicos son variados y no siempre fáciles de determinar, razón por la cual, el criterio resulta algo complejo. Esto se debe a que los DA no tienen un claro límite y pueden confundirse o asimilarse con otros tipos de discursos que se encuentran muy cercanos, tales como el discurso técnico-científico, el discurso profesional, el discurso pedagógico o el discurso institucional (López, 2002; Flowerdew, 2004). Desde el punto de vista contextual, Hyland (2000) argumenta que para caracterizar el DA resulta indispensable identificar las interacciones y los participantes involucrados en ellas, es decir, analizar los textos como prácticas sociales. Este enfoque incluye el análisis de los medios donde circulan y se utilizan estos textos (Gunnarsson, 1997). En este sentido, se considera el DA como una manifestación de una comunidad específica (Valle, 1997); así, iniciarse en esta comunidad implica conocer los procedimientos y estilos propios de la comunidad académica (Swales, 1990, 2000, 2004; Berkenkotter, Huckin & Ackerman, 1991), o tal como señalan Hyland y Tse (2004), manejar el metadiscurso académico (Crismore, 1989).

Desde el criterio textual, los textos académicos se caracterizan, según algunos autores (Hyland, 1998a, 2000; Reppen, 2004), por ser altamente elaborados, por utilizar un registro formal de la lengua y por presentar un lenguaje "objetivo" con un léxico preciso y específico (Cassany, Luna & Sanz, 2000). Normalmente, los textos académicos en que se realiza la actividad académica suelen ser: el manual, la guía didáctica, el informe de investigación, la presentación de proyectos, la clase magistral, la monografía, el artículo de investigación, el informe, la reseña, etc. (Hyland, 1998a, 1998b, 2000; Vázquez, 2001; Vilá, 2000; Mogollón, 2003; Parodi & Gramajo, 2003; Marinkovich, 2004). Estos textos tienden a ser caracterizados por un conjunto de rasgos lingüísticos y tramas textuales prototípicas que se acercan, muchas veces, al discurso científico, pero en que tiende a prevalecer un andamiaje didáctico con fines educativos.

Por otro lado, en tanto que manifestación de conocimientos especializados, las prácticas de comunicación académica tienen su canal preferente en el discurso escrito. Este permite la posibilidad de transmisión que necesita la academia para su propósito y progreso, además de que es el más adecuado para el registro estrictamente formal y normado del campo. Esto podría cuestionarse, debido a que es ampliamente conocido que existe una gran cantidad de prácticas orales en el ámbito académico; sin embargo, en muchos casos, estas prácticas se construyen a partir de rasgos que se acercan a la escritura y que, incluso en muchos casos, corresponden a experiencias previas de escritura (Vilá, 2000; Mogollón, 2003). Al respecto, según Mogollón (2003), el que las prácticas orales en el ámbito de la academia compartan los rasgos de la escritura promueve el supuesto de que el lenguaje de la ciencia tiende a lo escrito, porque sería lo que permitiría un discurrir más riguroso, lógico, planificado y objetivo. Por su parte, Hyland (2000) entrega dos argumentos a favor de la importancia de que el discurso disciplinar se cristalice en textos escritos:

"The first reason is that disciplinary discourse is considered to be a rich source of information about the social practices of academics [...]" The second reason given to

academic writing is the fact that what academics principally do is write: they publish articles, books, reviews, conference papers and research notes; they communicate with colleagues by e-mail, reprint requests, and referee evaluations; they communicate with students by handouts, study guides and textbooks; they contribute to electronic lists and to university reports; and they submit applications for grants and equipments□ (Hyland, 2000: 2-3).

Un ejemplo de la prevalencia de lo escrito por sobre lo oral, en el discurso de la academia, es la mayor validez que los académicos le asignan a los resultados de una investigación cuando ella se encuentra publicada. Este hecho se detecta a través de los procesos de citación y de referencias. Normalmente, en estos procedimientos no se encuentran citas de textos orales, ni referencias a resultados científicos presentados en congresos que no hayan sido publicados -al menos- en las respectivas actas del evento. De este modo, en parte, los procesos de citación y de referenciación se constituyen en un criterio textual que caracterizaría al DA. En este sentido, la referencia explícita de antecedentes teóricos y empíricos es un indicador de suma importancia de la relación que existiría entre el texto académico y el contexto de conocimiento en el cual se enmarca, constituyendo al texto en una pieza de vital importancia en la construcción colaborativa de nuevos conocimientos entre los escritores y los lectores (Antos, 1997; Hyland, 2000; Venegas, 2006). De acuerdo a Antos (1997) y Hyland (2000), la inclusión de argumentos a partir de la red de referencias no solo sugiere una apropiada orientación disciplinar, sino que también nos recuerda que las afirmaciones que se realicen responden invariablemente a afirmaciones previas y que a su vez estarán disponibles para nuevas afirmaciones. Justamente en esta línea argumental, Hyland (2000: 21) apunta que:

□The importance of citation as a constitutive element of the modern academic paper can be seen in its increasingly prominent role in the ways writers seek to construct facts through their communicative practices□.

En términos más globales, debido a que el DA se orienta a la transmisión de conocimiento, generalmente a través de la definición, la clasificación y la explicación (Wignell, 1998), se persigue que sus rasgos lingüísticos produzcan el efecto de claridad e incluso de objetividad, evitando ambigüedades e interpretaciones erróneas. Por ejemplo, se utiliza organizadores textuales que ordenan la exposición de las acciones realizadas, tales como □primero□, □luego□ y □por último□. Además, se tiende a la economía de palabras, la eliminación de redundancia, la eliminación de repeticiones, la ausencia de adjetivos vacíos para lograr concisión. Se mantiene una sintaxis controlada, en su orden habitual o canónico. Se distingue, además, una proporción más elevada de nominalizaciones deverbales que en el discurso no académico (Ciapuscio, 1992; Halliday, 1993; Lang, 1997; Gotti, 2003; Charaudeau, 2004; Parodi & Venegas, 2004; Cademártori, Parodi & Venegas, 2006). Al mismo tiempo, Simpson (2004) acota que las expresiones formulaicas son significativamente más frecuentes en el DA que en otros discursos orales indagados. Plantea que se reconocen dos funciones para este tipo de expresiones: una relacionada con la organización y estructuración del discurso y otra en relación con la interactividad. Así, Simpson declara que:

□There is a constant interplay between these two overarching characteristics of academic speech, which is by nature an information-rich genre, but in which interaction between the participants is also of paramount importance, and the formulaic expression identified here serve to highlight these dual pragmatic features□ (Simpson, 2004: 60).

Del mismo modo, Biber (2005), focalizado en el DA escrito de un conjunto de disciplinas, distingue distintas categorías de los paquetes lexicales (agrupación de palabras similares a las expresiones formulaicas) y afirma que:

□ Los paquetes referenciales son los más importantes en todas las disciplinas, aun cuando los paquetes posturales son también relativamente importantes. Sin embargo, resulta que estas disciplinas usan estos paquetes con frecuencias diferentes y en diferentes funciones básicas. En términos generales, los paquetes lexicales son más frecuentes en los libros de estudio de las ciencias naturales y menos frecuentes en los de humanidades. Pero las diferencias entre las disciplinas en cuanto a las preferencias por funciones específicas son aun más notables □ (Biber, 2005: 20).

Otro aspecto no considerado hasta ahora en esta revisión son los aspectos no lingüísticos que son frecuentes en este tipo de discurso, tales como: sistemas de representación del pensamiento como fórmulas químicas, ecuaciones físicas, recreaciones virtuales, representaciones matemáticas, símbolos. Además, se presentan disposiciones iconográficas de la información, tales como gráficos, tablas, figuras, esquema, etc. En atención a estos aspectos, Lemke (1998) prefiere hablar del texto como un sistema semiótico híbrido. Ello también ha sido abordado hoy en día como textos multimodales (Krees & van Leeuwen, 1996).

Por último, cabe destacar que en general los rasgos que se han señalado hasta ahora constituyen características no exclusivas del DA, por lo que se puede decir que su descripción acotada es aún poco clara. Como se comprende, el DA tiende a confundirse con el discurso científico (e.g. Bathia, 1993; Gunnarsson, 1997; Bazerman, 1998; Martin, 1998; Martin & Veel, 1998), el pedagógico científico (por ejemplo, Christie, 1998; Rose, 1998; Harvey, 2005) o el de divulgación científica (Bathia, 1993). Además, la mayor parte de los criterios mencionados aquí, en particular los textuales, han sido postulados para textos escritos en inglés, y aún más, en muchos casos solo para artículos de investigación científica.

Por todo ello, si se busca una caracterización más ajustada del denominado DA, se deberá investigar con detalle algunos de sus posibles rasgos prototípicos o, por el contrario, habrá que concluir que dentro del *continuum* de textos especializados se entrecruzan complejamente sin una fácil separación.

Es interesante distinguir, un par de aproximaciones o líneas de acción que pueden explicar algunas de las carencias y de las ambigüedades en cuanto al DA. Por un lado, se detecta una línea de desarrollo enmarcada desde la lingüística aplicada, con intereses pedagógicos y pragmáticos. Por ello, las investigaciones han sido motivadas por preocupaciones educacionales y con objetivos pedagógicos y, por ende, estos estudios (en su mayoría) son de corte descriptivo y se han focalizado en cuestiones aplicadas a los procesos de enseñanza/aprendizaje de lenguas. Mucho de lo producido por esta línea de investigación es lo que, para el caso mayoritario de la lengua inglesa, se recoge en publicaciones como *English for Specific Purposes* y *Journal of English for Academic Purposes*. Por otra parte, existe literatura que ha abordado el DA en términos más teóricos, tratando de avanzar en su caracterización (no necesariamente mediante acciones empíricas), pero aportando decididamente al ámbito. Esta otra línea de acción, no desligada totalmente de lo descriptivo-empírico, ha brindado avances certeros desde la teoría del género, los marcos de la teoría de la enunciación y de la Escuela Lingüística de Sydney, entre otros.

1.2.2. Discurso profesional

Sin lugar a dudas, el problema que enfrenta el DA en la búsqueda de una caracterización estricta constituye un asunto similar al que acontece con el estudio del discurso profesional (DP). Ello sucede debido a que, en algunos trabajos, ambos términos tienden a confundirse y se suele emplear el término DP como inclusivo del DA y vice versa. A pesar de que, en alguna medida, ambos objetos de estudio puedan teóricamente confundirse, en la presente investigación, diferenciaremos operativamente ambos tipos de discurso basándonos en el criterio de recolección ecológica, es decir, el DP será aquel recolectado en contextos de uso y circulación profesional, mientras que el DA será aquel recolectado en contextos de uso y circulación académicos. No negamos que existe un área de intersección entre ambos discursos, es más, como parte de nuestros objetivos de investigación pretendemos identificar y describir aquellos textos que circulan en ambos dominios y que se comparten y son el eje de conexión entre lo académico y lo profesional.

Vale la pena apuntar inicialmente una aproximación a la noción de DP desde los trabajos de Bazerman y Paradis (1991a). De acuerdo a estos autores, el DP se estructura a partir de una dinámica textual, la cual da forma a una profesión. En el volumen clásico editado por Bazerman y Paradis (1991b), se presentan una serie de artículos relacionados con la forma en que las comunidades profesionales construyen, a partir de los textos, una entidad propia. Los DP□s, en ese sentido, son aquellos textos en los cuales se plasma un conocimiento específico del mundo, en el que, a su vez, se constituyen los propósitos de la comunidad profesional (Berkenkotter, Huckin & Ackerman, 1991; Doheny-Farina, 1991; Bathia, 1993; Christie & Martin, 1997).

En palabras de Linell (1998), al tratar el tema de los DP□s, se debe distinguir entre:

- a) un discurso intraprofesional, es decir, entre los integrantes de una misma profesión;
- b) un discurso interprofesional, entre integrantes de profesiones distintas, y
- c) un discurso profesional-lego, entre los participantes de una comunidad profesional y un público lego.

La distinción planteada por este autor es relevante para los propósitos de esta investigación en cuanto el objeto de estudio a tratar se limitará solo a los discursos profesionales entendidos como en la letra (a), es decir, aquellos discursos que circulan al interior de un campo profesional determinado y que circulan entre los miembros de esa comunidad (Gunnarsson, 2004).

Con el fin de comprender el alcance y el desarrollo de las aproximaciones al estudio del D P, así como también, para visualizar las diferencias entre la presente investigación y los trabajos ya realizados, resultan esclarecedoras las categorías de análisis que propone López (2002: 212):

Tabla 1. Categorías para el análisis del discurso profesional (López, 2002).

Categoría	Aspectos involucrados
-----------	-----------------------

Género discursivo	Actividad profesional, mundo cognitivo, comunidad discursiva, objetivo pragmático, interlocutores específicos
Macro y superestructura	Selección de información y organización prototípica
Microestructuras retóricas o funciones discursivas	Relaciones retóricas recurrentes
Rasgos lingüísticos	Léxico, morfología, sintaxis
Características gráficas	Elementos de estructuración externa o paratextuales (títulos, subtítulos) gráficos, formalización (códigos no verbales)

Un primer comentario respecto de la Tabla 1 dice relación con que no todas estas categorías han sido tratadas en las investigaciones disponibles con el mismo énfasis. Resulta interesante comprobar que la mayoría de los estudios acerca del DP se focalizan en el análisis del discurso, entendido como una práctica social (primera categoría), es decir, adoptan un análisis del género discursivo como una manifestación de una comunidad discursiva. En esta línea, se encuentran, por citar solo algunos, los trabajos de Bazerman y Paradis (1991b), Gunnarsson, Linell y Nordberg (1997a), Connor y Upton (2004a) y Vine (2004).

Como es fácil comprobar, la macroestructura y la superestructura no han sido un foco muy desarrollado dentro del análisis del D P. A pesar de ello, se encuentran algunos estudios que aplican estas categorías al lenguaje administrativo (MAP, 1995). López (2002), por su parte, aplica algunas microestructuras retóricas al análisis de un texto de política económica, pero tal como la misma autora reconoce, no existen estudios sistemáticos respecto de cómo se distribuyen estas microestructuras retóricas en cada ámbito profesional.

En el criterio de los rasgos lingüísticos, destaca el nivel léxico como el que ha concentrado la mayor parte de los estudios; de manera especial, el análisis terminológico asociado a cada una de las profesiones (Ciapuscio, 2003). En el nivel morfológico, un tema altamente investigado ha sido el del rol que desempeñan las nominalizaciones (Chafe, 1982, 1985; Biber, 1986; Ciapuscio, 1992; Halliday, 1993; Lang, 1997; Parodi & Venegas, 2004; García, Hall & Marín, 2005; Cademártori, Parodi & Venegas, 2006). Por su parte, López (2002) contribuye con una descripción de una serie de rasgos sintácticos característicos del discurso profesional, tales como sintagmas nominales complejos y un grupo de adjetivos deverbales.

Aunque las características gráficas de los discursos profesionales no ha sido un asunto abordado de manera específica, existen algunos trabajos que lo enfrentan. Janssen y Jaspers (2004) estudian algunas orientaciones en el diseño de la correspondencia para clientes. Carliner y Boswood (2004) proponen algunas directrices de diseño gráfico para investigar las comunicaciones en línea al interior de un ámbito laboral y Sless (2004) hace una propuesta para el diseño de documentos públicos. Por su parte, Kress y van Leeuwen (1996), Kress, Jewitt, Ogborn y Tsatsarelis (2001) y Kress (2004) sientan las bases teóricas para el estudio de la interacción entre gráfica y discurso, en el ámbito de la multimodalidad.

Ahora bien, apartándonos de estas cinco categorías, existen algunos otros aspectos más generales que es necesario destacar respecto del análisis del D P. En primer

lugar, cabe señalar que el estudio de este tipo de discursos ha estado asociado al creciente interés que ha surgido acerca de la comunicación al interior de las organizaciones, preferentemente, gubernamentales y de negocios (Trosborg, 1997, 2000; Renkema, 2003; McCarthy & Handford, 2004; Warren, 2004). En segundo, las investigaciones acerca del DP han tendido a preferir ciertas profesiones denominadas clásicas y con alto impacto en la sociedad, a saber, abogado y médico (Gunnarsson, Linell & Nordberg, 1997b; Gallardo, 2005; Ferrari & Gallardo, 2006). En ellas se han producido abundante investigación empírica y se ha caracterizado estas prácticas discursivas desde diversas ópticas. En esta misma línea, Bhatia, Langton y Lung (2004) y Bhatia (2004) describen una metodología para el estudio del discurso legal y aportan hallazgos empíricos relevantes. Por su parte, interesantemente, Cheng (2004) y Connor y Upton (2004a) se escapan de este foco más tradicional: el primero aborda el estudio de la interacción de los practicantes de hotelería, mientras que Connor y Upton (2004b) realizan un análisis de las propuestas de investigación con financiamiento gubernamental. En general, entonces, debido a razones prácticas o al prestigio social de ciertas profesiones, el estudio del PD ha estado preferentemente solo en algunas profesiones o instituciones (los negocios, las leyes, la medicina y entidades de gobierno). En parte por lo expuesto más arriba, en el proyecto de investigación que describimos en este artículo hemos decidido centrarnos en otras profesiones menos o casi nada indagadas.

2. La investigación

2.1. Proyecto en marcha

Dado el marco anterior, en esta investigación nos proponemos realizar un estudio descriptivo-comparativo de orden lingüístico-textual a partir de los textos que son leídos en el ámbito académico y en el profesional tanto en las áreas de las Ciencias Básicas y de la Ingeniería como en el de las Ciencias Sociales y Humanas. Para ello, estudiaremos, con una metodología basada en la lingüística de corpus (Leech, 1991; Sinclair, 1991; Stubbs, 1996, 2006; McEnery & Wilson, 1996; Kennedy, 1998; Simpson & Swales, 2001; Tognini-Bonelli, 2001; Teubert, 2005; Parodi, 2006), desde el análisis multidimensional y multirasgos, un corpus académico y otro profesional. El primero circula en el ambiente universitario y el segundo se encuentra en entornos laborales en que estos mismos profesionales, egresados de cuatro carreras universitarias, se desempeñan. Buscamos que los textos que componen los corpora sean recogidos siguiendo criterios de carácter altamente ecológico y de representatividad. Esto quiere decir que recolectaremos cerca del 100% de los textos que circulan en cada ámbito, lo que constituye un corpus robusto y amplio de textos completos.

El ámbito académico estará definido por cuatro carreras impartidas en la Pontificia Universidad Católica de Valparaíso, Chile: Química Industrial, Ingeniería en Construcción, Trabajo Social y Psicología. Por su parte, el nivel profesional, estará conformado por un conjunto representativo de textos que circulan en las esferas profesionales en las que se desempeñan los cuatro tipos de egresados de estas mismas carreras.

Un rasgo innovador y vanguardista de esta propuesta en el marco nacional y latinoamericano, lo constituye el empleo de herramientas computacionales para el tratamiento de corpora digitales. En efecto, el uso de nuestra herramienta e

interfaz El Grial (www.elgrial.cl) posibilitará el marcaje y etiquetaje morfosintáctico automático de todos los textos, así como su interrogación vía Internet. Además, se utilizará un método poco explotado para el español, a saber, el análisis contrastivo multirasgos y multiregistros tanto en los niveles académicos entre ámbitos de las Ciencias Básicas y de la Ingeniería y de las Ciencias Sociales y Humanidades, como también entre cada disciplina estudiada en el nivel profesional. Todo ello se realizará tanto desde el enfoque de rasgos lingüístico/textual como desde las clasificaciones tipológicas y a partir de grandes muestras de textos, superando los estudios de índole ejemplar, o de trozos de textos y de selección *ad hoc*.

Los focos de la investigación permiten recabar datos en diversos frentes de acción: caracterización tanto del discurso académico como del profesional a partir de rasgos lingüístico/ textuales, conocimiento de los tipos de textos que circulan en el ámbito académico y en el profesional, disponibilidad de los textos en formato digital con etiquetaje morfosintáctico por medio de la aplicación de herramientas computacionales, identificación y cálculo de frecuencias de ocurrencia de un conjunto de rasgos relevantes y con funciones comunicativas específicas, etc. Al contrastar ambos tipos de discurso (académico y profesional), esperamos reconocer un *continuum* entre lo académico y lo profesional, aportando así información para el establecimiento estratégico de criterios de selección de los recursos textuales, que permitan enfrentar de mejor manera los requerimientos, tanto de la formación académica universitaria como del desempeño profesional. Y, por supuesto, impactar en el largo plazo en los procesos de comprensión y producción de textos técnicos especializados y, en definitiva, lograr así aportar a los procesos de construcción y apropiación de conocimientos disciplinares.

2.2. Constitución del Corpus PUCV-2006

Como se dijo anteriormente, pretendemos recolectar cerca del 100% del material escrito que se lee obligatoriamente o se consulta -en todas las asignaturas que conforman el plan obligatorio de especialidades- por los alumnos de cuatro carreras universitarias (dos de Ciencias Básicas y de la Ingeniería y dos de Ciencias Sociales y Humanas) y por los profesionales de cuatro áreas laborales vinculadas directamente a esas disciplinas y título profesional. A continuación, en la Tabla 2 se muestran las 4 áreas académicas y profesionales y su división en áreas científicas:

Tabla 2. Áreas académicas y profesionales.

	Áreas Académicas y Profesionales
Ámbito de las Ciencias Básicas y de la Ingeniería	Ingeniería en Construcción
	Química Industrial
Ámbito de las Ciencias Sociales y Humanas	Trabajo Social
	Psicología

La elección de estas cuatro carreras universitarias obedece a diversos criterios:

1) buscamos explorar áreas diferentes a las clásicamente indagadas en idioma inglés y en la mayoría de los trabajos disponibles en español, tales como leyes, medicina, economía y negocios.

2) nos interesa contrastar, desde diversas ópticas, los textos que circulan en el medio disciplinar de formación académica universitaria con los del ámbito profesional.

3) también es de nuestro interés el contraste, en diversos puntos, entre disciplinas de Ciencias Básicas y de la Ingeniería y de las Ciencias Sociales y Humanas.

Esta investigación, a diferencia de otras en esta línea, destaca por su carácter ecológico y de muestra no intencionada en un sentido estricto; además, de trabajar con un porcentaje cercano al 100% del material disponible y recolectado y no a partir de textos ejemplares o muestras aleatorias. Esto quiere decir que recolectaremos el total de los textos que efectivamente leen y circulan en los contextos tanto académico universitario como laboral-profesional de los cuatro ámbitos seleccionados, lo que, en un sentido, implica una doble cuestión de aproximación metodológica (ecología y representatividad). La investigación adquiere un importante carácter naturalístico, dado que no hemos intervenido ni introducido variables de ningún tipo en la recolección del material mismo (salvo, como se dijo, las de ámbito y nivel). No obstante lo anterior, por supuesto, que sí existe un sesgo y una intención general en la recolección del corpus, a saber, la de las Ciencias Sociales y Humanas y de las Ciencias Básicas y de la Ingeniería.

Ahora bien, la metodología más específica de esta investigación se divide en diferentes etapas según el estadio y foco de la misma:

a) Recolección del material: nivel académico

Con el fin de conformar el corpus del nivel académico se recolectarán los textos que leen los alumnos en las asignaturas de especialidad de las respectivas carreras en estudio. Para acceder y recoger efectivamente el material escrito, por un lado, se accederá a los programas de las asignaturas y se recopilará los textos que se indiquen en las bibliografías obligatorias y complementarias. Por otro, se entrevistará y se encuestará a los profesores de cada asignatura de especialidad para pedirles acceso a los textos que ellos proporcionan a sus alumnos como lectura obligatoria y/o complementaria en sus clases, pero que no está en las bibliografías de sus respectivos programas.

b) Recolección del material: nivel profesional

Con el fin de recopilar el corpus del ámbito profesional en las empresas y/o instituciones, procederemos a contactar a un número significativo de ex-alumnos de las carreras seleccionadas en el nivel académico con el fin de establecer quiénes serán los sujetos que corresponden al perfil del egresado determinado por la carrera. Esto quiere decir que identificaremos a aquellos sujetos que se desempeñan laboralmente dentro del ámbito que la carrera en cuestión proyecta para ese profesional. Una vez que hayamos determinado cuáles son los roles profesionales prototípicos de cada título profesional, procederemos a contactar y entrevistar a estos profesionales. El objetivo último será conocer sus prácticas de comunicación escrita en su ambiente laboral y recolectar la mayor cantidad posible de material escrito prototípico que estos sujetos leen en el desempeño de sus funciones.

c) Procedimientos: Corpus Académico

- ☐ Construcción de una base de datos con las mallas curriculares de las cuatro carreras (incluidos los programas de todas las asignaturas obligatorias).
- ☐ Construcción de una base de datos con las referencias bibliográficas obligatorias de todas las asignaturas obligatorias.
- ☐ Recolección de materiales complementarios que los profesores entregan a través de centrales de apuntes y centros de fotocopiado.
- ☐ Construcción de una encuesta para todos los profesores de cada una de las cuatro carreras. En ella, también se solicitó a los profesores la entrega de material complementario.

d) Procedimientos de análisis y construcción del corpus digital

A partir de la base de datos con las referencias bibliográficas se trabajó paralelamente en varios frentes de acción:

- ☐ Revisión en Internet de la disponibilidad de los títulos existentes en formato digital, con el fin de evitar así el escaneo y digitalización de los mismos en formato plano.
- ☐ Recolección de los textos desde las bibliotecas correspondientes.
- ☐ Fotocopia de cada texto con el fin de contar con una base de datos en formato papel, tanto para su estudio como tal como soporte para el proceso de digitalización.
- ☐ Constitución y entrenamiento de un equipo de diez personas para realizar las labores de escaneo y conformación de textos digitales en formato plano.
- ☐ Selección e implementación en nuestro Laboratorio de Ciencias del Lenguaje de los programas computacionales requeridos.

Una vez construido el corpus digital, se procedió a:

- ☐ La codificación de cada texto con el fin de identificar rápidamente su área disciplinar, el tipo textual y su número correlativo dentro del corpus.
- ☐ El marcaje estructural (proceso de etiquetaje morfosintáctico, ver Parodi, 2006), a través de la interfaz El Grial (www.elgrial.cl).
- ☐ El ingreso de cada texto marcado e identificado a la plataforma computacional en línea El Grial con el nombre de Corpus El Grial PUCV-2006, con el fin de posibilitar su disponibilidad en la red y su posterior análisis.

3. Resultados preliminares

En lo que viene, entregaremos dos tipos de resultados. Por una parte, daremos a conocer la constitución definitiva, en términos cuantitativos, del Corpus Académico

PUCV-2006, así como su distribución disciplinar en las cuatro carreras universitarias en estudio. Como parte de esta visión numérica, también mostraremos el estado de avance realizado hasta el momento en términos porcentuales del corpus digitalizado y operativo en el sistema en línea El Grial. Por otra, proponemos una primera aproximación tipológica de los textos que conforman el total del Corpus Académico PUCV-2006. De este modo, se aportan definiciones de nueve tipos textuales detectados y se cuantifican globalmente su ocurrencia al interior del corpus total.

Tal como se indicó, en la Tabla 3, podemos observar la cifra total alcanzada a través de los procedimientos indicados más arriba de recolección del Corpus Académico PUCV-2006.

Tabla 3. Constitución del Corpus Académico PUCV-2006: número de textos.

Ámbito científico	Carrera universitaria	Número de textos
Ciencias Básicas y de la Ingeniería	Química Industrial	76
	Ingeniería en Construcción	92
Ciencias Sociales y Humanas	Trabajo Social	188
	Psicología	296
Total de textos		652

Según se aprecia, se logró recoger 652 textos en las cuatro carreras en indagación. Lo primero que llama la atención de estas cifras es su alto grado de heterogeneidad, vale decir, existe un interesante aumento progresivo de la cantidad de textos desde las carreras de las Ciencias Básicas y de la Ingeniería hacia las de las Ciencias Sociales y Humanas. Aún más, esta tendencia se acrecienta entre Química Industrial e Ingeniería en Construcción tal como sucede entre Trabajo Social y Psicología. Una primera interpretación preliminar podría conducir a pensar que los alumnos de las carreras de Ciencias Sociales y Humanas leen muchos más textos que los alumnos de estas otras dos carreras de Ciencias Básicas y de la Ingeniería. Siendo más arriesgados, se podría incluso decir que los alumnos de Psicología leerían hasta cuatro veces más que los de Química Industrial. No obstante ello, es necesario tener presente un dato muy relevante, este es, la extensión de cada texto en cuestión, hecho que puede brindar una nueva mirada a las interpretaciones expuestas más arriba.

En todo caso, sin considerar la extensión como un factor específico, es un hecho que circula una mayor cantidad de textos en un dominio disciplinar que en otro. De este modo, sí resulta cierto apuntar que en Química Industrial se detecta una cantidad de textos muy inferior a la de Psicología, hecho que puede redundar en una serie de factores lingüísticos y académicos, tales como una posible menor exposición por parte de los alumnos a una amplia variedad tipológica, concentración en un tipo de discurso posiblemente muy prototípico, mayor acercamiento a un tipo de razonamiento discursivo que otro, etc.

Ahora bien, con el fin de dilucidar el asunto del factor extensión por número de palabras y de dimensionar el tamaño del corpus en cuestión, en la Tabla 4,

entregamos las cifras a partir de este dato. Vale la pena destacar que la tabla ha sido construida de modo diverso a la anterior, o sea, se ha hecho primar el orden creciente del número de palabras, lo que ha producido un cambio entre las carreras del área de Ciencias Básicas y de la Ingeniería (se consigna en primer lugar a Ingeniería y luego a Química).

Tabla 4. Corpus Académico PUCV-2006: número de palabras.

Ámbito Científico	Carrera universitaria (ámbito disciplinar)	Número de palabras
Ciencias Básicas y de la Ingeniería	Ingeniería en Construcción	14.653.760
	Química Industrial	22.103.620
Ciencias Sociales y Humanas	Trabajo Social	29.394.904
	Psicología	30.136.590
Total de textos		96.288.874

Las cifras, resumidas en la Tabla 4, constituyen un verdadero hito. No existe registro disponible en lengua española de un corpus académico escrito de estas características que alcance tal dimensión con un grado tan alto de representatividad y circunscripción temática y ecológica. Un corpus con un tamaño cercano a los 100 millones de palabras en formato digital, tematizado, tipologizado y para consulta en línea se transforma en una herramienta fundamental para el desarrollo de investigaciones de vanguardia desde los enfoques de la lingüística de corpus en lengua española. Así, el Corpus Académico PUCV-2006 se suma a los ya existentes en nuestro sitio web www.elgrial.cl, con lo que se fortalece el material disponible diversificado por registros para consultas e indagaciones múltiples.

Ahora bien, volviendo sobre el asunto del número de textos en relación a la extensión de los mismos, los datos aportados en la Tabla 4 revelan que sí existe una tendencia en el incremento progresivo registrado a partir del número de textos, el tipo de carrera y la adscripción disciplinar de la misma. Esta tendencia, en términos de palabras, no resulta ser tan determinante como lo demuestran los datos de la Tabla 3, pues entre Ingeniería y Psicología se detecta un aumento no superior al 50%. Este aumento es relativamente menor que el de 4 veces entre Química y Psicología en términos de número de textos. De cualquier modo, sea en términos de número de textos o número de palabras, según las cifras de las tablas 3 y 4, se hace evidente la existencia de un mayor número de material de lectura en las dos carreras de las Ciencias Sociales y Humanas que entre las otras dos de las Ciencias Básicas y de la Ingeniería. En Trabajo Social y Psicología se registra un número casi igual de palabras, aunque se consigna una diferencia en la cantidad de textos. Es relevante registrar que, según nuestras pesquisas, los alumnos de Química Industrial e Ingeniería en Construcción de la PUCV de Chile no están expuestos a la misma cantidad de material de lectura durante el transcurso de su carrera universitaria, hecho que debe redundar en diversas implicancias. Evidentemente, estas cifras deben ser analizadas con profundidad y a la luz de mayores informaciones que esperamos aportar, en parte, en lo que sigue de este artículo y, en parte, en el desarrollo del proyecto de investigación.

Justamente con el objetivo de profundizar en el análisis del material escrito recolectado, procedimos a su estudio y su clasificación en tipos de textos. Para ello, empleamos criterios lingüísticos, comunicativo-funcionales y textuales-discursivos.

En líneas generales, seguimos los principios orientadores de una taxonomía compleja en los términos que operan Parodi y Gramajo (2003). A continuación, en la Tabla 5, se nombran los nueve tipos de textos detectados y se consignan las cifras brutas de ocurrencia a partir del total de textos del Corpus Académico PUCV-2006.

Tabla 5. Distribución por tipo de textos.

Tipología Corpus PUCV-2006	N° de textos
Artículo de investigación científica	29
Conferencia	1
Diccionario	2
Guía didáctica	58
Informe	19
Manual	259
Norma	16
Test	2
Texto disciplinar	266
Total	652

La organización de la información en la Tabla 5 ha seguido el principio del orden alfabético por los nombres otorgados a los tipos de textos. Decidimos buscar nombres simples, cotidianos y de fácil acceso y transparencia en su uso. Como se aprecia, a partir de esta tabla, emerge un panorama muy heterogéneo pero con claras concentraciones. En principio, se puede decir que son dos los tipos textuales más recurrentes: Manual (M) y Texto Disciplinar (TD). Ello entrega un panorama global inicial en que se conjuga, por una parte, el saber disciplinar propiamente tal a través del TD que aborda conocimiento especializado, a veces, de alto grado de complejidad; por otra, encontramos el M que, aunque está fuertemente orientado por el conocimiento disciplinar, cuenta con un carácter didáctico variable en que se busca, en la mayoría de los casos, difundir saberes por medio de diversos recursos educativos (gráficos, tablas, diagramas, etc.) y llevar a los lectores al desarrollo de ejercicios y aplicación de conocimientos. No deja de llamar la atención que un corpus de esta naturaleza presente, comparativamente, escasa ocurrencia de, por ejemplo, Guía Didáctica (GD). Aunque ellas ocupan el tercer lugar en jerarquía numérica, estimamos que su cantidad es limitada, dado que se ha recolectado el material de cuatro carreras universitarias y a partir de currículos académicos con cuatro a cinco años de duración.

También resulta interesante constatar la escasa aparición del Artículo de Investigación Científica (AIC). Podría pensarse que este medio de transmisión de conocimiento especializado ocuparía un lugar y espacio más relevante, al menos, dentro de las carreras de Ciencias Básicas y de la Ingeniería. La ocurrencia de solo 29 AIC en el total del corpus resulta limitada si se piensa que este es un medio vigente de información actualizada y de vanguardia.

Como se dijo, los tipos de textos en este corpus aparecen claramente concentrados en torno a dos prototipos. Ellos revelan dos puntos del *continuum* de discursos, pero no dos puntos extremos. Tanto el M como el TD están orientados hacia el eje de mayor especialización, con una clara oscilación hacia el punto medio de difusión

didáctica, mas no de manera radical. La GD que representa el punto extremo de la gradiente, en el eje supuesto como de mayor preocupación educativa, solo muestra una aparición muy reducida. Ahora bien, con el fin de hacer más transparentes nuestros tipos de textos y de que se comprenda mejor nuestro punto de clasificación, a continuación, entregamos una breve definición operativa de cada uno de ellos:

1. Artículo de Investigación Científica: texto que presenta una revisión teórica o estudio empírico publicado en una revista de especialidad.

2. Conferencia: texto leído en público por un especialista en un tema específico.

3. Diccionario: texto en el que se recoge y explica de forma ordenada conceptos o procedimientos de una disciplina o materia determinada.

4. Guía Didáctica: texto breve de carácter didáctico que contiene un conjunto de explicaciones e indicaciones acerca de un tema específico, presentando generalmente ejercicios o actividades sugeridas.

5. Informe: presentación de una situación, problema o caso en que se pueden exponer las variables que intervienen y/o conclusiones o comentarios.

6. Manual: texto disciplinar de carácter didáctico que puede contener ejemplos, ejercicios, tablas y recursos multimodales que facilitan la comprensión de su contenido.

7. Norma: texto que expone una reglamentación de conductas y/o procedimientos.

8. Test: conjunto de pruebas mediante las que se valoran en términos cuantitativos las características psicológicas de un sujeto.

9. Texto Disciplinar: texto que entrega conocimiento especializado sobre uno o varios temas de una disciplina particular.

Como se aprecia, esta breve definición permite reconocer cada tipo en su marco de acción cotidiano y distinguirlos unos de otros. Al mismo tiempo y de modo nocional, resultan, a partir de sus etiquetas, relativamente de fáciles de reconocer por sus usuarios en la práctica académica diaria.

Por último, en la Tabla 6, entregamos las cifras del corpus actualmente digitalizado. En otras palabras, la conformación del corpus entregado en la Tabla 4 que arrojaba casi 100 millones de palabras mostraba el corpus total en formato papel. A continuación, se calcula el corpus ya procesado digitalmente hasta este momento, el que alcanza un porcentaje representativo del 70%.

Tabla 6. Corpus digital Académico PUCV-2006: 70% del total.

	Porcentaje del Corpus Total	Número de palabras
Corpus recolectado en papel	100%	96.288.874

Corpus digital	70%	67.402.212
-----------------------	-----	------------

Según los datos de la Tabla 6, el 70% del corpus alcanza una cifra superior a 67 millones de palabras. Ello corresponde a un total de 457 textos. Este corpus ya se encuentra en formato digital (con marcaje morfosintáctico) y disponible a través del sitio web www.elgrial.cl. Para identificarlo, se utiliza la sigla Corpus PUCV-2006 y se puede también acceder a las diferentes clasificaciones tipológicas.

COMENTARIOS FINALES Y PROYECCIONES

De acuerdo a lo apuntado en la Introducción de este artículo, y a partir del marco del proyecto en marcha y de los resultados preliminares, nos parece que tanto las implicancias científico-investigativas *per se* como las pedagógicas pueden ser múltiples y variadas. Contar con datos efectivos acerca de los tipos de discurso escrito utilizados al interior de las aulas de formación académica universitaria como en su práctica laboral y profesional, así como una caracterización de sus tipos y rasgos será de suyo una contribución al campo.

En lo más específico, contar con un corpus académico de estas proporciones y disponible gratuitamente en línea (www.elgrial.cl) constituye un hito dentro de las investigaciones en lengua española. No tan solo por el tamaño de este corpus académico del español, sino también por su alto grado de representatividad y valor ecológico. Un conjunto de textos especializados académicos, seleccionados desde criterios bien definidos y compartiendo un propósito único y que alcanza un cifra cercana a los 100 millones de palabras constituye ciertamente un aporte para el desarrollo de la investigación en Iberoamérica y en el mundo. No existe registro que dé cuenta de algo similar, con excepción de los corpora de recolección heterogénea y difusa en su origen así como de los recolectados a través de Internet, cuyas limitantes resultan ser de variada índole.

Respecto de la recolección del corpus académico PUCV-2006 y de los primeros análisis de su constitución, cabe llamar la atención acerca de las relevantes distinciones detectadas entre las áreas de las Ciencias Sociales y Humanas y de las Ciencias Básicas y de la Ingeniería. Los datos aportados permiten, preliminarmente, relevar diferencias tanto en el número de textos escritos que circulan como medio de transmisión y apropiación de conocimientos disciplinares como en la variedad de textos en que se difunden los saberes especializados. Así, se hace evidente que Psicología y Trabajo Social tienden a emplear una mayor cantidad de textos con una relativa mayor extensión (en número de palabras) en su formación universitaria mientras que, comparativamente, Química Industrial e Ingeniería en Construcción presentan una mayor escasez de volúmenes y de extensión menor en cuanto al número de palabras.

En cuanto a los tipos textuales detectados, se aprecia una interesante heterogeneidad que confirmaría la idea de un *continuum* de tipos de textos entre los polos de mayor complejidad disciplinar y una mayor preocupación didáctica de corte más general y divulgativo. En este sentido, se identifican claramente dos tipos textuales: Manual y Texto Disciplinar. Ambos muestran una especial atención al conocimiento disciplinar en formato relativamente especializado. Sin duda, de manera especial, llama la atención que se haya recolectado, comparativamente, un número escaso de Guías Didácticas. Aunque ellas ocupan el tercer lugar en

jerarquía numérica, estimamos que su cantidad es limitada, dado que se ha recolectado el material de cuatro carreras universitarias. Este asunto será explorado en detalle en las

próximas fases de proyecto y se buscará describir por carrera las ocurrencias de los tipos de textos con el fin de establecer patrones más finos de uso y posibles vinculaciones con el tipo de conocimiento especializado que se va conformando en los alumnos.

Ahora bien, entre las varias tareas pendientes, se encuentra la construcción de una matriz de rasgos que permitan la exploración y explotación del corpus, con el fin de alcanzar objetivos de descripción profunda. Al mismo tiempo, pretendemos realizar diversos tipos de comparaciones como también establecer otras indagaciones, a saber, realizar estudios comparativos entre discurso escrito y discurso oral, especializado y más divulgativo, etc.

En parte, para cumplir lo señalado, en este momento se están realizando las acciones para incorporar a nuestras bases de datos en línea un nuevo corpus oral del español de Chile. Gracias a la gentil colaboración de la Dra. Gladys Cepeda de la Universidad Austral de Valdivia, Valdivia, Chile, podremos contar con gran parte del Corpus del Habla de Valdivia, recolectado por ella a través de numerosos proyectos de investigación en la zona sur de Chile (Cepeda, 2001, 2002). Este material actualmente está siendo recopilado y sistematizado por personal técnico de nuestro equipo de la PUCV y será traído prontamente a Valparaíso e incorporado a través de los procedimientos informáticos pertinentes a nuestro sitio e interfaz de consulta El Grial. De este modo, estará gratuitamente a disposición de la comunidad científica internacional. Con ello, avanzamos en el cumplimiento de uno de los objetivos de la creación de este sitio web, cual es ofrecer a los investigadores una herramienta en línea y sin costo alguno y de algunos compartidos con la **Escuela Lingüística de Valparaíso** (www.linguistica.cl).

REFERENCIAS BIBLIOGRÁFICAS

Antos, G. (1997). Texte als Konstitutionsformen von Wissens. Thesen zu einer evolutionstheoretischen Begründung der Textlinguistik. En G. Antos & H. Tietz (Eds.), *Die Zukunft der Textlinguistik. Traditionen - Transformationen - Trends* (pp. 43-63). Tübingen: Niemeyer. [[Links](#)]

Bhatia, V. (1993). *Analysing genre: Language use in professional settings*. London: Longman. [[Links](#)]

Bhatia, V. (2004). *Worlds of written discourse. A genre-based view*. London: Continuum. [[Links](#)]

Bhatia, V., Langton, N. & Lung, J. (2004). Legal discourse: Opportunities and threats for corpus linguistics. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 203-231). Amsterdam: Benjamins. [[Links](#)]

Bazerman, Ch. (1998). Emerging perspectives on the many dimensions of scientific discourse. En J. Martin & R. Veel (Eds.), *Reading science. Critical and functional*

perspectives on discourses of science (pp. 15-28). London: Routledge. [[Links](#)]

Bazerman, Ch. & Paradis, J. (1991a). Introduction. En Ch. Bazerman & J. Paradis (Eds.) *Textual dynamics of the professions* (pp. 3-10). Wisconsin: The University of Wisconsin. [[Links](#)]

Bazerman, Ch. & Paradis, J. (Eds.) (1991b). *Textual dynamics of the professions*. Wisconsin: The University of Wisconsin. [[Links](#)]

Berkenkotter, C., Huckin, T. & Ackerman, J. (1991). Social context and socially constructed texts: The initiation of a graduate student into a writing research community. En C. Bazerman & J. Paradis (Eds.), *Textual dynamics of the professions* (pp. 191-215). Wisconsin: The University of Wisconsin. [[Links](#)]

Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language*, 62, 384-414. [[Links](#)]

Biber, D. (1988). *Variation across speech and writing*. Cambridge: CUP. [[Links](#)]

Biber, D. (1994). Using register-diversified corpora for general language studies. En S. Armstrong (Ed.), *Using large corpora* (pp. 180-201). Cambridge: The MIT Press. [[Links](#)]

Biber, D. (2005). Paquetes léxicos en textos de estudio universitario: Variación entre disciplinas académicas. *Revista Signos*, 38(57), 19-30. [[Links](#)]

Biber, D. (2006). *University Language. A corpus-based of spoken and written registers*. Amsterdam: Benjamins. [[Links](#)]

Burdach, A. (2000). El léxico científico técnico: Un recurso publicitario persuasivo. *Onomazein*, 5, 189-208. [[Links](#)]

Cabré, T. (1993). *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Antártica/Em-puréis. [[Links](#)]

Cabré, M., Doménech, M., Morel, J. & Rodríguez, C. (2001). Las características del conocimiento especializado y la relación con el conocimiento general. En M. Cabré & J. Feliú (Eds.), *La terminología técnica y científica* (pp. 173-186). Barcelona: Instituto Universitario de Lingüística Aplicada. [[Links](#)]

Cabré, T. & Gómez, J. (2006). *La enseñanza de los lenguajes de especialidad. La simulación global*. Madrid: Gredos. [[Links](#)]

Cademártori, Y. , Parodi, G. & Venegas, R. (2006). El discurso escrito y especializado: Caracterización y funciones de las nominalizaciones en los manuales técnicos. *Literatura y Lingüística*, (17), 243-265. [[Links](#)]

Carliner, S. & Boswood, T. (2004). Genre: A useful construct for researching online communication for the workplace. *Information Design Journal + Document Design*, 12(2), 124-135. [[Links](#)]

Cassany, D., Luna, M. & Sanz, G. (2000). *Enseñar lengua*. Barcelona: Editorial Graó. [[Links](#)]

Cepeda, G. (2001). Estudio descriptivo del español de Valdivia, Chile. *Estudios Filológicos*, 36, 81-97. [[Links](#)]

Cepeda, G. (2002). Entonación, actitud y modalidad. *Estudios Filológicos*, 37, 7-27. [[Links](#)]

Chafe, W. (1982). Integration and involvement in speaking, writing and oral literature. En D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-53). Norwood, N.J.: Ablex. [[Links](#)]

Chafe, W. (1985). Linguistic differences produced by differences between speaking and writing. En D. Olson, N. Torrence & A. Hidiard (Eds.), *Literature, language and learning: The nature and consequences of reading and writing* (pp. 105-123). Cambridge: Cambridge University Press. [[Links](#)]

Charaudeau, P. (2004). La problemática de los géneros. De la situación a la construcción textual. *Revista Signos*, 37(56), 23 -39. [[Links](#)]

Cheng, W. (2004). Did you took from the minibar? What is the practical relevance of a corpus-driven language study to practitioners in Hong Kong's hotel industry? En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 141-166). Amsterdam: Benjamins. [[Links](#)]

Christie, F. (1998). Science and apprenticeship. The pedagogic discourse. En J. Martin & R. Veel (Eds.), *Reading Science. Critical and functional perspectives on discourse of science* (pp. 152-180). London: Routledge. [[Links](#)]

Christie, F. & Martin, J. (Eds.) (1997). *Genre and institutions. Social Processes in the workplace and the school*. London: Continuum. [[Links](#)]

Ciapuscio, G. (1992). Impersonalidad y desagenticación en la divulgación científica. *Lingüística española actual*, 2, 183-205. [[Links](#)]

Ciapuscio, G. (2003). *Textos especializados y terminología*. Barcelona: Instituto Universitario de Lingüística Aplicada. [[Links](#)]

Connor, U. & Upton, T. (Eds.) (2004a). *Discourse in the professions. Perspectives from corpus linguistics*. Amsterdam: Benjamins. [[Links](#)]

Connor, U. & Upton, T. (2004b). The genre of grant proposals: A corpus linguistics analysis. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 235-255). Amsterdam: Benjamins. [[Links](#)]

Crismore, A. (1989). *Talking with readers. Metadiscourse as rhetorical act*. New York: Lang. [[Links](#)]

Doheny-Farina, S. (1991). Creating a text / creating a company: The role of a text in the rise and decline of a new organization. En C. Bazerman & J. Paradis (Eds.), *Textual dynamics of the professions* (pp. 306-335). Wisconsin: The University of Wisconsin. [[Links](#)]

Dudley-Evans, T. & St. John, M. (2006). *Developments in English for Academic Purposes. A mul-tidisciplinary approach*. Cambridge: Cambridge University Press. [[Links](#)]

Ferrari, L. & Gallardo, S. (2006). El estudio diacrónico de la evaluación en las introducciones de artículos científicos de medicina. *Revista Signos*, 39(61), 161-180. [[Links](#)]

Flowerdew, J. (Ed.) (2002). *Academic discourse*. Cambridge: Cambridge University Press. [[Links](#)]

Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional language. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 11-33). Amsterdam: Benjamins. [[Links](#)]

Gallardo, S. (2005). *Los médicos recomiendan. Un estudio de las notas periodísticas sobre salud*. Buenos Aires: Eudeba. [[Links](#)]

García, M., Hall, B. & Marín, M. (2005). Ambigüedad, abstracción y polifonía del discurso académico: Interpretación de las nominalizaciones. *Revista Signos*, 38(57), 49-60. [[Links](#)]

Gotti, M. (2003). *Specialized discourse. Linguistic features and changing conventions*. Bern: Lang. [[Links](#)]

Gunnarsson, B. (1997). On the sociohistorical construction of scientific discourse. En B. Gunnarsson, P. Linell & B. Nordberg (Eds.), *The construction of professional discourse* (pp. 99-126). Essex: Longman. [[Links](#)]

Gunnarsson, B. (2004). The multilayered structured of enterprise discourse. *Information Design Journal + Document Design*, 12(1), 36-48. [[Links](#)]

Gunnarsson, B., Linell, P. & Nordberg B. (Eds.) (1997a). *The construction of professional discourse*. Essex: Longman. [[Links](#)]

Gunnarsson, B., Linell, P. & Nordberg, B. (1997b). Introduction. En B. Gunnarsson, P. Linell & B. Nordberg (Eds.), *The construction of professional discourse* (pp. 1-12). Essex: Longman. [[Links](#)]

Halliday, M. (1993). On language and physical science. En M. Halliday & J. Martin (Eds.), *Writing science. Literacy and discursive power* (pp. 54-68). Pittsburgh: University of Pittsburgh Press. [[Links](#)]

Harvey, A. (1987). *A macrostructure analysis of 20 computer-data-based abstracts in the field of engineering*. Trabajo presentado en el Seminario Nacional de Lenguas con Fines Específicos, Pontificia Universidad Católica de Chile, Santiago. [[Links](#)]

Harvey, A. (1994) *La interlocución en informes de investigación*. En Actas 9º Encuentro Sociedad Nacional de Profesores de Lenguas Extranjeras, La Serena, 195-203. [[Links](#)]

Harvey, A. (2004). *La evaluación en el discurso de informes escritos por estudiantes universitarios*. Ponencia presentada en el III Encuentro Nacional de Análisis del Discurso, Universidad Austral de Chile, Valdivia. [[Links](#)]

Harvey, A. (2005). La evaluación en el discurso de informes escritos por estudiantes universitarios chilenos. En M. Pilleux (Ed.), *Contextos del discurso* (pp. 215-228). Valdivia: Universidad Austral de Chile. [[Links](#)]

Harvey, A. & Horsella, M. (1988). Exploring and exploting the structure of computational abstracts. *The Specialist*, 9, 229-247. [[Links](#)]

Horsella, M. & Pérez, F. (1996). *Las nominalizaciones en el discurso científico: Una visión discrónica*. En Actas X Encuentro Sociedad Nacional de Profesores de Lenguas Extranjeras, Iquique, 205-220. [[Links](#)]

Horsella, M., Pérez, F. & Sindermann, G. (1994). *La síntesis lingüística en la ciencia*. En Actas 9º Encuentro Sociedad Nacional de Profesores de Lenguas Extranjeras, La Serena, 205-212. [[Links](#)]

Hyland, K. (1998a). *Hedging in scientific research articles*. Amsterdam: Benjamins. [[Links](#)]

Hyland, K. (1998b). *Other floors. Other voices. A textography of a small building*. Mahwah, N.J.: Erlbaum. [[Links](#)]

Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London: Longman. [[Links](#)]

Hyland, K. & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25, 156 - 177. [[Links](#)]

Jansen, D. & Jaspers, J. (2004). Client-oriented correspondance: How do you actually do it. *Information Design Journal + Document Design*, 12(3), 187-189. [[Links](#)]

Kennedy, G. (1998). *An introduction to corpus linguistics*. New York: Longman. [[Links](#)]

Kennedy, J. (2001). Language use, language planning and EAP. En J. Flowerdew & I. Peacock (Eds.), *Research perspectives on English for Academic Purposes* (pp. 23-41). Cambridge: Cambridge University Press. [[Links](#)]

Kress, G. (2004). Reading images: Multimodality, representation and new media. *Information Design Journal + Document Design*, 12(2), 110-119. [[Links](#)]

Kress, G. & van Leeuwen, T. (1996). *Reading images. The grammar of visual design*. London: Routledge. [[Links](#)]

Kress, G., Jewitt, C., Ogborn, J. & Tsatsarelis, C. (2001). *Multimodal teaching and learning*. London: Continuum. [[Links](#)]

- Lang, M. (1997). *Formación de palabras en español*. Madrid: Cátedra. [[Links](#)]
- Leech, G. (1991). The state of the art in corpus linguistics. En K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics. Studies in honor of Jan Svartvik* (pp. 8-29). London: Longman. [[Links](#)]
- Lemke, J. (1998). Multiplying meaning. Visual and verbal semiotics in scientific text. En J. Martin & R. Veel (Eds.), *Reading science. Critical and functional perspectives on discourses of science* (pp. 3-14). London: Routledge. [[Links](#)]
- Linell, P. (1998). Discourse across boundaries: On recontextualisation and the blending of voices in professional discourse. *Text*, 18(2), 143-157. [[Links](#)]
- López, C. (2002). Aproximación al análisis de los discursos profesionales. *Revista Signos*, 35 (51-52), 195-215. [[Links](#)]
- MAP (1995). *Manual de documentos administrativos*. Madrid: Tecnos. [[Links](#)]
- Marinkovich, J. (2004). Los textos especializados en la educación media técnico profesional. Hacia una clasificación en el ámbito de la divulgación científica. En A. Harvey (Comp.), *En torno al discurso. Contribuciones de América Latina* (pp. 129-134). Santiago: Ediciones Universidad Católica de Chile. [[Links](#)]
- Martin, J. (1998). Discourse of science: Recontextualisation, genesis, intertextuality and hegemony. En J. Martin & R. Veel (Eds.), *Reading Science: Critical and functional perspectives on discourses of science* (pp. 3-14). London: Routledge. [[Links](#)]
- Martin, J. & Veel, R. (Eds.) (1998). *Reading science. Critical and functional perspectives on discourses of science*. London: Routledge. [[Links](#)]
- McCarthy, M. & Handford, M. (2004). Invisible to us: A preliminary corpus-based study of spoken business English. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 167-202). Amsterdam: Benjamins. [[Links](#)]
- McEnery, T. & Wilson, A. (1996). *Corpus linguistics*. Edinburgh: Edinburgh University Press. [[Links](#)]
- Mogollón, G. (2003). Paradigma científico y lenguaje especializado. *Revista de la Facultad de Ingeniería de la U.C.V.*, 18, 5-14. [[Links](#)]
- Núñez, P. & Espejo, C. (2004). Estudio exploratorio acerca de la conceptualización del informe escrito en el ámbito académico. En A. Harvey (Comp.), *En torno al discurso. Contribuciones de América Latina* (pp. 135-148). Santiago: Ediciones Universidad Católica de Chile. [[Links](#)]
- Núñez, P. , Muñoz, A. & Mihovilovic, E. (2006). Las funciones de los marcadores de reformulación en el discurso académico en formación. *Revista Signos*, 39(62), 471-492. [[Links](#)]
- Parodi, G. (Ed.) (2005a). *Discurso especializado e instituciones formadoras*. Valparaíso: Ediciones Universitarias de Valparaíso. [[Links](#)]

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

Parodi, G. (2006). Discurso especializado y lengua escrita: Foco y variación. *Estudios Filológicos*, 41, 165-204. [[Links](#)]

Parodi, G. (Ed.) (2007). *Working with Spanish corpora*. London: Continuum. [[Links](#)]

Parodi, G. & Gramajo, A. (2003) Los tipos textuales del corpus PUCV-2003: Una aproximación multiniveles. *Revista Signos*, 36(54), 207-223. [[Links](#)]

Parodi, G. & Venegas, R. (2004). BUCÓLICO: aplicación computacional para el análisis de textos. Hacia un análisis de rasgos de la informatividad. *Lingüística y Literatura*, 15, 223-251. [[Links](#)]

Renkema, J. (2003). En torno a la solución del problema de Orwell en la comunicación gubernamental: Investigación experimental de la estructura de la información en los sitios Web. *Revista Signos*, 36(53), 103-120. [[Links](#)]

Reppen, R. (2004). Academic language: An exploration of university classroom and textbook language. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 65-87). Amsterdam: Benjamins. [[Links](#)]

Rose, D. (1998). Science discourse an industrial hierarchy. En J. Martin & R. Veel (Eds.), *Reading science. Critical and functional perspectives on discourses of science* (pp. 236-265). London: Routledge. [[Links](#)]

Simpson, R. (2004). Stylistic features of academic speech: The role of formulaic expressions. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 37-64). Amsterdam: Benjamins. [[Links](#)]

Simpson, R. & J. Swales. (2001). Introduction. North American perspectives on Corpus Linguistics at the Millennium. En R. Simpson & J. Swales (Eds.), *Corpus Linguistics in North America. Selections from the 1999 Symposium* (pp. 1-14). Ann Arbor: The University of Michigan Press. [[Links](#)]

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press. [[Links](#)]

Sless, D. (2004). Designing public documents. *Information Design Journal + Document Design*, 12(1), 24-35. [[Links](#)]

Stubbs, M. (1996). *Text and Corpus Analysis*. Oxford: Blackwell. [[Links](#)]

Stubbs, M. (2006). Corpus analysis: The state of the art and three types of unanswered questions. En S. Hunston & G. Thompson (Eds.), *System and corpus: Exploring connections* (pp. 15-36). London: Equinox. [[Links](#)]

Swales, J. (1990). *Genre analysis. English in academic and research settings*. Cambridge: Cambridge University Press. [[Links](#)]

Swales, J. (2004). *Research genres. Explorations and applications*. Cambridge: Cambridge University Press. [[Links](#)]

Tapia, M., Burdiles, G. & Arancibia, B. (2003). Aplicación de una pauta diseñada para evaluar informes académicos universitarios. *Revista Signos*, 36(54), 249-257. [[Links](#)]

Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10(1), 1-13. [[Links](#)]

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: Benjamins. [[Links](#)]

Trosborg, A. (Ed.) (1997). *Text, Typology and Translation*. Amsterdam: Benjamins. [[Links](#)]

Trosborg, A. (Ed.). (2000). *Analising professional genres*. Amsterdam: Benjamins. [[Links](#)]

Valle, E. (1997). A scientific community and its texts: A historical discourse study. En B. Gunnarsson, P. Linell & B. Nordberg (Eds.), *The construction of professional discourse* (pp. 76-98). Essex: Longman. [[Links](#)]

Vásquez, G. (2001). *Guía didáctica del discurso académico escrito. ¿Cómo se escribe una monografía?* Madrid: Edinumen. [[Links](#)]

Venegas, R. (2006). La similitud léxico-semántica en artículos de investigación científica en español: Una aproximación desde el Análisis Semántico Latente. *Revista Signos*, 39(60), 75-106. [[Links](#)]

Vilá, M. (2000). *La enseñanza y el aprendizaje del discurso oral formal: Una secuencia didáctica sobre las explicaciones orales en clases*. Tesis Doctoral no publicada Universidad Autónoma de Barcelona, España. [[Links](#)]

Vine, B. (2004). *Getting things done at work. The discourse of power in workplace interaction*. Amsterdam: Benjamins. [[Links](#)]

Warren, M. (2004). So what have you been working on recently? Compiling a specialized corpus of spoken business English. En U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 115-140). Amsterdam: Benjamins. [[Links](#)]

Wignell, P. (1998). Technicality and abstraction in social science. En J. Martin & R. Veel (Eds.), *Reading science. Critical and functional perspectives on discourses of science* (pp. 297-326). London: Routledge. [[Links](#)]

Williams, I. (1998). Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics*, 3, 151-171. [[Links](#)]

Recibido: 2-VI-2006, **Aceptado:** 17-X-2006

Dirección para Correspondencia: Giovanni Parodi (gparodi@ucv.cl). Tel.: (56-32) 2273390. Pontificia Universidad Católica de Valparaíso. Av. Brasil 2830, piso 9, Valparaíso, Chile.

* Proyecto FONDECYT 1060440.