



PONTIFICIA UNIVERSIDAD CATOLICA DE VALPARAISO

PROGRAMA DE DOCTORADO EN LINGÜÍSTICA

INSTITUTO DE LITERATURA Y CIENCIAS DEL LENGUAJE

FACULTAD DE FILOSOFÍA Y EDUCACIÓN

**LAS RELACIONES LÉXICO-SEMÁNTICAS EN
ARTÍCULOS DE INVESTIGACIÓN CIENTÍFICA:
UNA APROXIMACIÓN DESDE EL ANÁLISIS
SEMÁNTICO LATENTE**

René Alejandro Venegas Velásquez

Tesis para optar al grado de Doctor en Lingüística -

Becario Proyecto MECESUP PUCV/0004

Profesor Guía:

Dr. Giovanni Parodi Sweis

Valparaíso, 2005

*A Carmen Gloria y Diego
A René, Luzmira y Ricardo
A mis maestros de hoy y de siempre
Gracias por su constante apoyo*

Resumen

En esta investigación, se estudian las similitudes léxico-semánticas entre las variables textuales de las palabras clave, el resumen, la introducción y el contenido de una muestra representativa de artículos de investigación científica, pertenecientes a un corpus de 675 artículos de investigación de tres ámbitos científicos (*ciencias biológicas, ciencias sociales y ciencias exactas*). Para la determinación de las similitudes semánticas se utiliza un método estadístico-computacional denominado Análisis Semántico Latente (LSA). Este método combina el análisis de co-ocurrencias léxicas en los textos y la reducción de dimensionalidad para construir un espacio semántico especializado. Considerando los resultados generales de este análisis, se puede plantear que en la muestra investigada las relaciones léxico-semánticas observadas en los productos textuales de los científicos de diversas disciplinas no muestran mayores diferencias debido, en gran medida, a los procesos de estandarización a los cuales son sometidos los artículos de investigación científica. Además, se identifican patrones de similitud léxico-semántica mayores para las relaciones resumen-introducción y resumen contenido, que para las relaciones palabras clave-resumen, palabras clave-introducción y palabras clave-contenido. Esto implica que el resumen macrosemantiza, mejor que las palabras clave, los contenidos de los artículos de investigación científica. Este resultado es particularmente interesante pues el proceso

computacional concuerda con lo que nocionalmente se esperaría para esta relación de variables textuales. Los resultados que presentamos pueden ser de mucha utilidad para los investigadores que se inician en la escritura de los artículos de investigación científica, así como para la construcción de herramientas computacionales que evalúen las relaciones léxico-semánticas en diferentes tipos textuales.

Palabras Clave: Escritura científica, similitud semántica, análisis semántico latente, artículo de investigación científica.

INTRODUCCIÓN

La investigación que presentamos a continuación se encuentra enmarcada en los estudios del discurso especializado, específicamente en torno al discurso de la ciencia materializado en el artículo de investigación científica. Nuestra preocupación en este ámbito se centra en el modo en que los escritores investigadores materializan los resultados de indagaciones teóricas y/o empíricas en los artículos de investigación científica publicados en revistas que cuentan con procedimientos establecidos de arbitraje y edición. Un foco particular de interés nos resulta el modo de organizar la información y las relaciones de significado que construye el escritor para dar forma a su producto textual. De este modo, aspectos que no escapan a esta investigación son la atención en el desarrollo de una competencia escrita especializada en lengua española y la alfabetización técnica en este género discursivo.

Si bien, el término “discurso especializado” se encuentra, en la actualidad, ampliamente aceptado por los estudiosos del lenguaje, se debe reconocer que su utilización no surgió sino solo hace unos pocos años. En general, la noción de discurso especializado se concibe de forma amplia y globalizadora (Parodi, 2005a), reconociéndose al interior del concepto un *continuum*, en el que se alinean textos que van desde una alta hasta una baja especialidad, de modo que incluye un conjunto variado de tipos de textos.

En términos particulares, el discurso especializado de la ciencia ha sido concebido como un tipo de discurso que expresa credibilidad y prestigio, es decir, un discurso autorizado sobre el tema que trata (Hyland, 2000), teniendo como propósito cristalizar el conocimiento, cumpliendo por ello la función de ser soporte y transmisor del saber disciplinar, a través de estructuras textuales basadas, frecuentemente, en secuencias textuales descriptivas y argumentativas, las que presentan un alto grado de generalización y abstracción semántica (Swales, 1990, 2004; Hyland, 2000; Tapia, Burdiles & Arancibia, 2003).

Acorde con lo anterior, el texto tradicionalmente estudiado en el ámbito del discurso especializado de la ciencia es el artículo de investigación científica. Esta preocupación por el texto se ha focalizado, clásicamente, en el estudio del inglés con propósitos especiales (por ejemplo, Sager, Dungworth & McDonald, 1980; Bazerman, 1988; Swales, 1990; Hyland, 1998, 1999, 2000; Salager-Meyer, 1991, 1992, Martín, 2003). En tanto, para el español los estudios que caracterizan el discurso especializado de la ciencia y los artículos de investigación científica están orientados con fines descriptivo-didácticos y son más bien recientes (por ejemplo, Calsamiglia, 1998; Bolívar, 2000; Ciapuscio, 2000, 2003; Cassany, López & Martí, 2000; Moyano, 2000; Ciapuscio & Otañi, 2002; López, 2002; Mogollón, 2003; Martín, 2003; Gotti, 2003).

En general, estos autores conciben el artículo de investigación científica como un texto escrito, publicado en una revista especializada, que tiene como finalidad informar a la comunidad científica los resultados de un trabajo de investigación realizado mediante la aplicación del método científico, lo que exige una clara estructuración retórica, adhiriendo comúnmente al modelo IMRD (Introducción-Método-Resultado-Discusión) propuesto por Swales (1990). Sin embargo, como el mismo Swales (2004) plantea, esta estructura variará según las características propias de cada disciplina científica.

Al respecto, diversos estudios se han llevado a cabo atendiendo a cada una de las partes de este tipo textual. Así, solo por destacar algunos trabajos, la introducción ha sido profundamente trabajada por Swales (1990, 2004), las introducciones y conclusiones por Gnutzmann y Oldenburg (1991), las conclusiones por Ciapuscio y Otañi (2002), el resumen por Salager-Meyer (1991) y Bolívar (2000), el resumen y las introducciones por Martín (2003), las secciones introducción y discusión por Dudley-Evans (1986).

Ahora bien, todas estas investigaciones han sido llevadas a cabo, fundamentalmente, desde perspectivas lingüístico-textuales, retóricas y sociocognitivas, utilizando, en la mayoría de los casos, muestras ejemplares de textos y criterios preferentemente cualitativos. El enfoque de análisis más común ha sido la comparación interlenguas, con un interés particular en el estudio del inglés para propósitos especiales y didácticos. Por lo mismo, los aportes de estos estudios tienen un sustento empírico bastante limitado. Por otra parte, en general, estos trabajos, si bien han sido comparativos interlenguas, con poco desarrollo en español y mucho en inglés, no han considerado estudios comparativos intercienias en lengua española.

Atendiendo a este desarrollo del campo disciplinar, en este trabajo llevamos a cabo un estudio comparativo intercienia a partir de las relaciones semánticas de la estructura retórica de los artículos de investigación científica escritos en español y disponibles en revistas indexadas que cumplen con los estándares internacionales de publicación científica. De este modo, estudiaremos las relaciones entre las palabras clave, resumen, introducción y contenido, a partir de los valores cuantitativos que pueden ser calculados considerando la similitud léxico-semántica existente entre tales apartados retóricos. Para ello emplearemos el Análisis Semántico Latente (LSA, por su sigla en inglés) (Deerwester, Dumais, Furnas, Landauer & Harshman, 1990; Landauer, Foltz & Laham, 1998).

Cabe hacer notar que la noción de similitud léxico-semántica que utilizamos aquí está fuertemente asociada a los estudios del significado de tipo asociativo colocacional, según el cual, el significado de una palabra o un párrafo se determina por el grado de intercambiabilidad contextual o el grado en el cual una palabra puede ser substituida por otra en un contexto dado. Lo interesante de este concepto de similitud léxico-semántica es que es definible y aplicable computacionalmente. En nuestro caso, el LSA permite, a través de la implementación de la técnica estadística de reducción dimensional, denominada descomposición en valores singulares, representar eficientemente en un espacio semántico multidimensional las relaciones entre palabras y párrafos de grandes corpus textuales no anotados. De este modo se pueden llevar a cabo comparaciones, por medio de valores de similitud léxico-semántica, entre las variables textuales de cualquier texto.

En síntesis, nuestro objetivo de investigación es comparar las relaciones de similitud léxico-semánticas entre las palabras clave, el resumen, la introducción y el contenido de los artículos de investigación científica publicados en revistas científicas indexadas de tres áreas de la ciencia (ciencias exactas, ciencias biológicas y ciencias sociales), utilizando una herramienta computacional de análisis vectorial denominada Análisis Semántico Latente.

Parte importante de esta investigación se concentrará en el desarrollo de un espacio semántico para el español en los términos del LSA. Ello, debido a que no se cuenta con una herramienta disponible para nuestra lengua y, fundamentalmente, se ha trabajado con ella en lengua inglesa.

Ahora bien, con el fin de cumplir el objetivo general de esta investigación doctoral, organizamos el presente trabajo en dos partes. La primera, está constituida por dos capítulos que aportan los antecedentes teóricos necesarios para el desarrollo de la investigación. En el Capítulo 1, luego de una revisión de los conceptos de texto y discurso, se define al texto como una representación lingüística externa material del proceso psicosociodiscursivo, factible de ser analizado a través de métodos computacionales automáticos y semiautomáticos. Luego, caracterizamos al artículo de investigación científica centrándonos, particularmente, en la estructura y función del resumen, las palabras clave y la introducción en este tipo textual. En el Capítulo 2, luego de una exposición de las principales teorías, tipos de significado y sus relaciones, establecemos que el tipo de significado que nos interesa estudiar es el significado de tipo asociativo colocacional, el cual es posible de ser calculado por medio del LSA. Finalizamos este capítulo con el análisis crítico y justificación de este método en la identificación de las relaciones léxico-semánticas de las variables textuales en estudio.

La segunda parte corresponde a la presentación del marco metodológico. Así, en el Capítulo 3, luego de problematizar respecto de la calidad de la escritura científica, presentamos los objetivos e hipótesis de investigación. Además, describimos el corpus y la muestra de artículos de investigación científica utilizados. Finalmente, damos cuenta de la construcción del espacio semántico especializado (ES-ARTICO) según el cual se comprobarán o refutarán las hipótesis generales y específicas de esta investigación. En el Capítulo 4, se presenta el análisis de los resultados, así como las conclusiones obtenidas en

la investigación. Finalizamos el capítulo exponiendo algunas limitaciones detectadas en la investigación, así como las fortalezas y proyecciones.

PRIMERA PARTE
MARCO TEÓRICO

CAPÍTULO 1

“Texto y discurso especializados”

1.1 Objetivo del capítulo

El objetivo de este capítulo es caracterizar nuestro objeto de estudio: el artículo de investigación científica. Para cumplir con este objetivo, nos referiremos, en primer lugar, a los conceptos de “texto” y de “discurso” con el fin de proponer una definición de texto, congruente con el nivel de análisis que llevaremos a cabo. Más adelante, discutiremos los conceptos de “discurso especializado” y el de “texto especializado”, con el objetivo de circunscribir nuestro objeto de investigación a un área particular de estudio que nos permitirá identificarlo con un tipo de discurso particular, el de especialidad, y a una clase textual específica, el artículo de investigación científica. Finalmente, se lleva a cabo una revisión de las partes que conforman el artículo de investigación científica, prestando especial atención al resumen, las palabras clave y la introducción.

1.2 El *continuum* texto/discurso

La primera distinción que realizaremos es la referente a la conceptualización de “texto” y de “discurso”. Ambos conceptos serán entendidos aquí como dos polos de un *continuum* conceptual, el cual surge debido a los múltiples sentidos que han ido adoptando estos conceptos dependiendo de la perspectiva disciplinar y del problema particular en estudio. Esto último ha redundado en un uso polisémico, ambiguo, laxo, incontrolado e incluso contradictorio de estos conceptos (van Dijk, 1980; Bernárdez, 1982; Maingueneau, 1989; Álvarez, 1995, 2001).

El problema de la relación “texto/discurso” surge, tal como afirma van Dijk (1980), debido a que los investigadores alemanes y holandeses, que trabajaron a principios de los años setenta en la fundamentación científica de la lingüística del texto, no disponían en sus

propias lenguas de un concepto transparente de discurso, por lo que optaron por utilizar de manera unánime el término “texto”, que les era más familiar, empleando “discurso” cuando escribían en inglés para referirse a una entidad marcadamente abstracta. Por otra parte, la existencia de estos dos términos en las lenguas romances, se ha explotado para establecer a su vez dos términos técnicos de sentidos diferentes, que dieron lugar a la denominación de la lingüística textual y del análisis del discurso. De esta manera, texto y discurso son dos formas diferentes de mirar el mismo fenómeno, lo que ha redundado en que la división entre el análisis del texto y del discurso se difumine.

Por su parte, y en relación con el incontrolado uso de la palabra discurso, Foucault (1972: 80) plantea:

“Lastly, instead of gradually reducing the rather fluctuating meaning of the word ‘discourse’, I believe that I have in fact added to its meanings: treating it sometimes as the general domain of all statements, sometimes as an individualizable group of statements, and sometimes as a regulated practice that accounts for a certain number of statements; and have I not allowed this same word “discourse”, which should have served as a boundary around the term statement, to vary as I shifted my analysis or its point of application, as the statement itself faded from view?”.

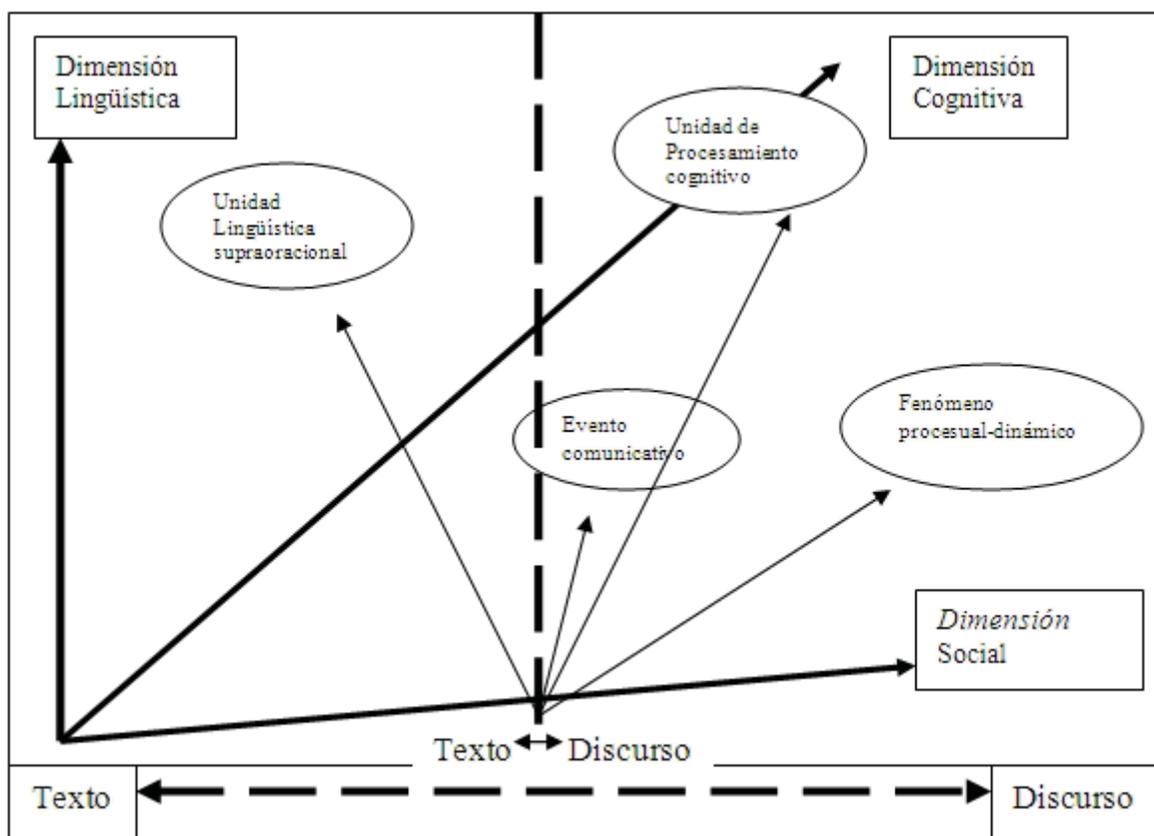
En lo que sigue, llevaremos a cabo un ejercicio de descripción de las dimensiones disciplinares desde las cuales estos conceptos han sido definidos, a pesar de que como plantea Bonilla (1997: 9) tal distinción no es más que “un problema terminológico sin apenas importancia”. No obstante lo anterior, nos planteamos un primer objetivo descriptivo de este *continuum* conceptual con el afán de intentar distinguir teóricamente los conceptos de texto y discurso, de modo que podamos reconocer los criterios y elementos que nos permitan situarnos operacionalmente respecto de tales conceptos, a partir de tres dimensiones básicas que se distinguen en las distintas definiciones de texto y de discurso, a saber: una dimensión lingüística, una dimensión cognitiva y una dimensión social.

A continuación presentamos, de forma muy breve, algunas definiciones categorizadas en cuanto su proximidad a una de las tres dimensiones mencionadas. De este modo, damos a conocer los conceptos de texto, definido en cuanto unidad lingüística transaccional, luego, las definiciones en las que texto y discurso han sido considerados como una unidad

comunicacional, más adelante en la que ambos conceptos han sido tratados como una unidad de procesamiento cognitivo. Posteriormente, presentaremos definiciones en las que estos conceptos han sido entendidos como una unidad sociocognitiva y, finalmente, presentaremos las posturas en las que el discurso es entendido como una práctica social, afecta a sus condiciones de producción.

En la Figura 1 hemos intentado graficar las tres dimensiones que configuran las múltiples posibilidades según las cuales se pueden categorizar las distintas definiciones del *continuum* conceptual texto/discurso.

Figura 1. Dimensiones del *continuum* conceptual de texto/discurso.



Como sabemos, la palabra “texto”, proveniente del latín ‘*textus*’ (Gómez de Silva, 1995) o ‘*textum*’ (Corominas & Pascual, 1986) con el sentido de “tejido construido” y más tarde con el de “lo escrito, composición literaria, estructura”, ha sido objeto de interés científico desde algo más de cincuenta años (Bernárdez, 1982; De Beaugrande & Dressler, 1997). En

efecto, es recién a partir de principio de los años cincuenta, con Harris (1952), que se inicia lo que él denominó “análisis del discurso” y que hoy se concibe como “lingüística textual” (Charadeau & Maingeneau, 2002). Desde esa época comienza el interés científico del texto como objeto de estudio, entendido como una real preocupación por trascender los estudios oracionales. De este modo, la lingüística textual nace debido a la necesidad de considerar el contexto supraoracional en la descripción y explicación de algunos fenómenos gramaticales, en otras palabras, la necesidad de extender la unidad de análisis de la oración al texto. Cabe hacer notar, como plantean De Beaugrande y Dressler (1997), que si bien la lingüística transformacional, basada en principios distribucionalistas, ya había hecho avances en el estudio de los textos, estos avances no habrían sido reconocidos por los lingüistas de la época. De este modo, esta perspectiva lingüística y su centro de interés científico, el texto, comienza a ser reconocida y ampliamente trabajada recién a fines de los años sesenta y principios de los setenta (ver Bernárdez, 1982; De Beaugrande & Dressler, 1997). Entre los exponentes más conocidos se encuentran, según estos autores, los trabajos realizados por Harweg (1968), Stempel (1971), Dressler (1972), Fries (1972), Schmidt (1973), Dressler y Schmidt (1973), Sitta y Brinker (1973), Petöfi y Reiser (1974), Harweg (1974, 1978), Daneš y Viehweger (1976), Coulthard (1977), Gülich y Raible (1977), Dressler (1978), van Dijk (1980), mencionando sólo los más conocidos de la década de los setenta.

En esta primera etapa de la lingüística textual, el objeto “texto” fue estudiado desde dos perspectivas: desde una orientación interna, enfocada en el sistema lingüístico, y desde una externa, centrada en el uso lingüístico del hecho comunicativo (Brinker, 1988). De este modo, se plantea la tarea de estudiar científicamente los textos, particularmente, su estructura y función, privilegiando la “mirada lingüística”, en su sentido amplio (Ciapuscio, 2003). Su objeto de estudio será el texto, el cual se caracteriza por ser complejo, debido a que abarca todos los niveles de análisis lingüísticos posibles, es decir, el nivel léxico, el gramatical, el semántico y el nivel pragmático; así como también, las distintas perspectivas, tanto como producto de interacciones concretas o como proceso en sí, esto es, en lo que se refiere a los complejos procesos de producción y comprensión (Bernárdez, 1982; De Beaugrande & Dressler, 1997; Ciapuscio, 2003).

1.2.1 El texto como unidad lingüística

En cuanto a la orientación interna, la lingüística del texto se comprende a sí misma como una lingüística de la *langue* o de la “competencia” (Saussure, 1983; Chomsky, 1957, 1965). En este contexto, la formación del texto se produce por medio del sistema de reglas y se basa en regularidades generales y sistemáticas que deben ser explicadas por una teoría del texto, anclada en los aspectos teóricos y metodológicos de la corriente oracional, bien estructuralista o bien generativa (Ciapuscio, 2003). Esta mirada redundante en la definición de texto, concebido como: a) un encadenamiento pronominal ininterrumpido (Harweg, 1968), b) un complejo de signos lingüísticos que muestra una sucesión ordenada de oraciones, integrada, finita, continua, y construida de acuerdo con las reglas de la gramática, que el productor pretende que sea semánticamente cerrada, y que proporcionan el desarrollo lineal del desenvolvimiento de un tema a partir de su núcleo temático (Agricola, 1976 citado por Bernárdez, 1982), c) una ordenación de contenidos conceptuales, como reflejos de hechos y fenómenos de la realidad y sucesiones de sonidos o grafemas, realizada de acuerdo a un determinado plan de actividad y realizada mediante las reglas del sistema de la lengua (Viehweger, 1976), d) una sucesión de unidades lingüísticas constitutivas (Brinker, 1979), e) concepto abstracto que se manifiesta o realiza en discursos (van Dijk, 1980). En definiciones algo más actuales, se ha ampliado el concepto considerando aspectos lingüísticos no sólo gramaticales sino también semánticos, algunas definiciones características son: f) una secuencia de oraciones que poseen una macroestructura (van Dijk, 1983), y g) una sucesión de signos lingüísticos coherentes (Brinker, 1988).

Por su parte, el concepto de discurso también ha sido concebido desde una perspectiva eminentemente lingüística, así siguiendo a Maingueneau (1989), se distinguen tres concepciones tradicionales de discurso, entendido como: 1) sinónimo del habla saussureana (Gardiner, 1989), 2) una unidad lingüística de dimensión superior a la oración (un enunciado) (Benveniste, 1966) y 3) el conjunto de las reglas de encadenamiento de las sucesiones de oraciones que componen el enunciado (Harris, 1952).

En definitiva, en todas estas definiciones, tanto de texto como de discurso, se privilegia una orientación interna en la que el objeto se construye en relación con la sucesión de unidades

lingüísticas coherentes, que trascienden la unidad oracional y que, en su conjunto, presentan un cierre semántico.

En función de plantear una diferencia significativa entre los conceptos de texto y discurso, y con el fin de distinguir dos polos en un *continuum*, los significados presentados más arriba los asociaremos a texto, concordando con Charaudeau y Maingueneau (2002) quienes, al establecer la diferencia entre texto y discurso, sostienen que el texto es una realidad concreta, empírica. En este sentido, el texto se entiende como el producto material lingüístico del acto de comunicación entre individuos, producido en un marco social determinado (como veremos más adelante)

1.2.2 Texto/discurso como unidad comunicativa

En la segunda orientación, el foco de interés está centrado en el uso lingüístico, es decir, el texto está ligado a una situación comunicativa y es utilizado para el logro de objetivos específicos de los participantes del acto comunicativo (Ciapuscio, 2003). La fundamentación teórica está dada por la teoría de los actos de habla (Austin, 1982; Searle, 1986), la cual permite concebir al texto como:

“[...] todo componente expresado de un acto comunicativo en un juego de acción comunicativa, que está temáticamente orientado y que se corresponde con una función comunicativa reconocible. Esto es, realiza un potencial ilocucionario reconocible” (Schmidt, 1973, citado por Ciapuscio, 2003: 150).

Algunas definiciones de texto realizadas desde esta perspectiva, todas citadas por Bernárdez (1982: 80-81), son: a) todo aquello que es lenguaje en forma comunicativa o social, es decir, referida al interlocutor (Schmidt, 1971), b) la unidad por medio de la cual se realiza la actividad lingüística en tanto actividad social-comunicativa (Isenberg, 1976), c) producto del acto de habla, como discurso en el que se produce el mensaje y se proyecta una práctica significativa centrada en un aquí y ahora configurado, a su vez, por un “yo” que le da origen (Fonseca & Fonseca, 1977), d) conjunto verbal funcional completo, un “acto de habla”, es habla estructurada, y al mismo tiempo, conjunto comunicativo superior (Kozevniková, 1979).

Cabe aquí establecer, que las orientaciones esbozadas no son, en ningún caso, opuestas, sino complementarias, en el sentido de que la orientación interna, corresponde al complejo de signos lingüísticos que se construye según las reglas del sistema de la lengua; en tanto que la externa corresponde al acto comunicativo, es decir, a la implementación del sistema en contextos comunicativos particulares.

Una visión integradora es la propuesta por Bernárdez (1982: 85), quien concibe al texto como:

“[...] la unidad lingüística comunicativa fundamental, producto de la actividad verbal humana, que posee siempre carácter social; está caracterizado por su cierre semántico y comunicativo, así como por su coherencia profunda y superficial, debida a la intención (comunicativa) del hablante de crear un texto íntegro, y a su estructuración mediante dos conjuntos de reglas: las propias del nivel textual y las del sistema de la lengua”.

Podemos observar que esta definición integra en forma de características atribuidas al texto los criterios básicos definitorios de todo texto. Según lo expone Bernárdez (1982), estos criterios son: carácter comunicativo (actividad), carácter pragmático (intención y situación) y carácter estructurado (reglas textuales y del sistema de la lengua).

En términos más actuales, una de las definiciones más conocida y extendida de texto, según Titscher, Meyer, Wodak y Vetter (2000), es la aportada por De Beaugrande y Dressler (1997). Estos últimos, definen al texto desde una perspectiva comunicativa, en tanto que proponen que el texto, para que sea tal, debe satisfacer ciertas condiciones, acotadas en siete criterios y tres principios regulativos de la comunicación textual. Los criterios pueden ser agrupados en: criterios lingüísticos (coherencia y cohesión), criterios psicolingüísticos (intencionalidad y aceptación), criterios sociolingüísticos (situacionalidad e intertextualidad) y criterio computacional (informatividad). Cabe hacer notar que estos criterios son hoy en día cuestionables, en particular, el criterio lingüístico de la coherencia, dado que como se verá más adelante es un aspecto centrado en procesos cognitivos más que puramente lingüísticos (ver Capítulo 2, apartado 2.5.3). De hecho, el mismo De Beaugrande (2005) modaliza la inclusión de la coherencia y la cohesión en el criterio lingüístico, planteando que estos criterios están prominentemente orientados hacia el texto:

“Of these seven criteria, two seem prominently text-oriented (cohesion and coherence), two prominently psychological (intentionality and acceptability), two prominently social (situationality and intertextuality), and the last, computational (informativity). But close investigation shows that none of the criteria can be appreciated without considering all four factors: language, mind, society, and processing” De Beaugrande (2005: capítulo I).

En cuanto a los principios regulativos de la comunicación textual, De Beaugrande y Dressler (1997) distinguen: el de la eficacia, según el cual quienes intervengan en el intercambio textual obtengan los mejores resultados comunicativos posibles invirtiendo en esa tarea un esfuerzo mínimo; el de la efectividad, relacionado con la intensidad del impacto comunicativo que el texto provoca en sus receptores; y el de la adecuación, entendido como el equilibrio óptimo que se consigue en un texto entre el grado de actualización de los criterios de textualidad, por un lado, y la satisfacción de las demandas comunicativas, por otro. Así, por ejemplo, de acuerdo a esta perspectiva, una señal caminera, un artículo del diario, un argumento y una novela son todos textos que se rigen por las distintas reglas de géneros particulares. Todos los géneros mencionados tienen rasgos lingüísticos características, cumplen con funciones específicas y están limitados a situaciones particulares de producción y recepción (Swales, 1990).

Esta noción se ve reforzada por el mismo De Beaugrande (1997: 10) cuando plantea que el texto es “a communicative event wherein linguistic, cognitive, and social actions converge” entendido como un “multi-system comprising multiple interactive systems” (De Beaugrande, 1997: 11). En esta noción cada sistema conecta unidades de distintos tipos (sonidos, palabras, significados, participantes discursivos, acciones en un plan, etc.), siendo estas unidades a su vez, multifuncionales, ya que, por ejemplo: una palabra funciona como patrón de sonidos, como parte de una frase, como una instrucción para activar significado, etc. De esta manera cada secuencia leída o escuchada es una pequeña porción de materia o energía en la cual ha sido condensada una enorme cantidad de información por parte del escritor o hablante y que espera ser amplificada por el lector u oyente (De Beaugrande, 2005). Notamos claramente, cómo el autor combina rasgos lingüísticos, textuales, sociales, psicolingüísticos y computacionales, haciendo que la definición de texto sea, en definitiva, un constructo teórico multidisciplinar.

Por su parte, Álvarez (2001), desde una perspectiva notoriamente más pragmática, establece que el discurso es un acto de habla y, por tanto, consta de los elementos de todo acto de habla: en primer lugar, un acto locutivo o locucionario, es decir, el acto de decir un dicho (texto) con sentido y referencia; en segundo lugar, un acto ilocutivo o ilocucionario, o el conjunto de actos convencionalmente asociados al acto ilocutivo; finalmente, un acto perlocutivo o perlocucionario, o sea, los efectos en pensamientos, creencias, sentimientos o acciones del interlocutor. El texto, en cambio, es lo dicho, el enunciado y su organización. Para Álvarez (2001), un texto no puede existir aisladamente, pues necesita ser insertado en contextos culturales determinados y en circunstancias específicas, de lo contrario carecería de sentido. Es decir, un texto sólo puede ser parte de un discurso que prevé las condiciones de producción del texto y las condiciones de su consumo.

En todas estas definiciones reconocemos claramente una orientación externa del texto como objeto material, en la que se enfatiza la actividad en la cual está inmerso el texto, es decir, la acción comunicativa compleja. Con relación al concepto de discurso, Maingeneau (1989), distingue que este ha sido definido, desde esta perspectiva, como toda enunciación que supone un hablante y un oyente, y en el primero, la intención de influir de alguna manera en el otro.

En síntesis, estas orientaciones comunicativas de texto/discurso están guiadas por la finalidad o propósito del hablante o productor, por la atención en los interlocutores, y por participación de estos en una particular situación comunicativa.

1.2.3 Texto/discurso como unidad de procesamiento cognitivo

La noción de texto, entendido desde el punto de vista del procesamiento cognitivo, en términos de su comprensión y su producción, ha sido ampliamente abordada por la psicolingüística, aunque rara vez diferenciada del concepto de discurso (por ejemplo, van Dijk & Kintsch, 1983; Peronard & Gómez, 1985; Kintsch, 1988, 1998; Graesser, Singer & Trabasso, 1994; Graesser, Millis & Zwaan, 1997; Peronard, Gómez, Parodi & Núñez, 1997; van den Broek & Gustafson, 1999; Graesser & Wiemer-Hastings, 1999; Sander & Spooren, 2001, Louwerse, 2004). En general, desde esta perspectiva el texto es la instancia

intermedia, codificada lingüísticamente, entre dos representaciones cognitivas: la del productor y la del comprendedor. Al respecto Sanders y Spooren (2001: 3) plantean que:

“In this view, there is a producer who has a cognitive representation of what she [or he] intends to communicate; this is formulated in a linguistic code, called the text, and this text is decoded by the interpreter who can be said to understand a text once he has made a coherent representation of it. This view fits theories that describe the link between the structure of a text as a linguistic object, its cognitive representation and the processes of text production and understanding”.

Como se sabe, esta perspectiva se centra fundamentalmente en el procesamiento cognitivo que los seres humanos hacemos de los textos. Particularmente, en cómo comprendemos y producimos textos a partir de nuestras representaciones mentales. Esta perspectiva es, según van den Broek y Gustafson (1999), característica de una primera generación de los estudios en comprensión, cuyas principales conclusiones son que la representación mental es una construcción realizada por el lector que difiere, y va más allá, de la información presentada en el texto mismo (para mayor desarrollo de los niveles de representación ver 2.5.3). Otra conclusión relevante es que una buena representación del texto siempre es coherente (al respecto ver cohesión y coherencia en 2.5.3). Por último, cabe destacar la idea de que una representación mental coherente de un texto es posible de ser capturada por una red de nodos interconectados, en la cual cada nodo representa un elemento y las conexiones indican las relaciones semánticas que el lector ha identificado entre los elementos del texto (van den Broek & Gustafson, 1999).

Desde la perspectiva de una segunda generación en los estudios de comprensión, van den Broek y Gustafson (1999) destacan la atención al proceso mismo de lectura, donde el centro de interés está en el estudio de las inferencias que se generan durante la lectura. Entre las conclusiones más relevantes de estos estudios está la noción de que los individuos generan inferencias en línea de tipo referencial y causal. Respecto de otros tipos de inferencias en línea, las conclusiones no son generalizadas (ver Graesser, Singer & Trabasso, 1994). En términos más generales, se establece que la generación de inferencias es funcional, esto quiere decir que las inferencias se generan cuando se requiere construir coherencia a partir del texto. Por último, la activación de estas inferencias (referencial y causal, pero también otras) dependen de múltiples factores, entre los cuales destacan: las

restricciones del texto, el conocimiento de mundo del lector (para realizar tal o cual inferencia), los recursos atencionales del lector y los criterios del lector para el establecimiento de la coherencia (van den Broek & Gustafson, 1999).

Una tercera generación en los estudios de la comprensión se ha enfocado, considerando las conclusiones de las otras, en concebir modelos que describan cómo las activaciones en línea conducen hacia una representación coherente en la memoria, bajo la idea de que la lectura de un texto es un cambio de activaciones continuo y dinámico. En estos modelos el proceso de lectura es concebido en ciclos de lectura (oración por oración) en los que los patrones o vectores (de conexiones) van cambiando con cada nueva información proporcionada por el texto (van den Broek & Gustafson, 1999; van den Broek, Young, Tzeng & Linderholm, 1999). Indudablemente se observa aquí un salto en la concepción de las representaciones cognitivas, desde modelos simbólicos a modelos conexionistas. Entre las principales conclusiones obtenidas por estos estudios está la idea de que los recursos atencionales forman una combinación de activaciones distribuidas respecto de múltiples elementos lingüísticos, entre los cuales algunos reciben mayor o menor activación, variando además la fuerza de activación de estos elementos en la construcción de la representación del texto. Un aporte interesante de esta perspectiva es la construcción de una imagen dinámica y fluida del proceso y del producto de la comprensión (van den Broek, Rinden, Fletcher & Thurlow, 1996; van den Broek & Gustafson, 1999; van den Broek, Young, Tzeng & Linderholm, 1999).

Una metáfora muy particular que da cuenta del proceso de comprensión a partir de la representación de un texto es la que plantea Louwerse (2004: 41). Para él, la comprensión es como jugar al rompecabezas:

“Jugar al rompecabezas implica juntar las piezas a fin de formar un todo unificado. Una pieza se enlaza con la pieza siguiente; la siguiente con otra y el rompecabezas se forma a partir de las piezas así relacionadas. En el resultado del proceso de armado, las piezas apenas se notan dado que las piezas interrelacionadas se han convertido en una representación interrelacionada. Pero, ¿cómo es que las piezas apoyan al rompecabezas en el proceso de armado? La analogía del rompecabezas ilustra, en cierta forma, la comprensión de textos, la construcción de una representación mental coherente de los eventos descritos por el texto”.

En esta analogía, los eventos descritos por el texto son las piezas, en tanto que la comprensión es el proceso de armado que permite construir una representación mental coherente a partir de la interrelación de las piezas.

En el ámbito latinoamericano, destacamos los aportes de Peronard y Gómez (1985), Peronard (1997) y Gómez (1997a y b), quienes plantean un modelo de comprensión lingüística que se funda en una concepción antropológica del lenguaje y una visión trascendente y dualista del ser humano (Parodi, 2003). Para estos autores la comprensión “es un acto consciente en que culminan innumerables procesos personales de aprehensión interior de los significados o sentidos de las cosas” (Gómez, 1997a: 99). Lo que en el caso de la comprensión textual se traduce en “el acto o proceso de comprensión de unidades verbales” (Gómez, 1997a: 10), acto en el cual están presentes los siguientes factores: a) un comprendedor (necesariamente humano), b) una unidad verbal o texto, estructurado por su autor con el propósito de transmitir significados o sentidos a otra persona, mediante pistas (léxicas, gramaticales y/o semánticas), fruto de la intencionalidad humana. Por su parte, la unidad verbal, según Gómez (1997a), se caracteriza según distintos aspectos: b.1) una determinada organización física perceptual, oral o gráfica, apta para ser interpretada por quien conozca o descubra la codificación utilizada en su construcción, b.2) ciertos sentidos que el comprendedor ha de captar, interpretar, traducir o completar y b.3) un conjunto de convenciones lingüísticas y culturales indispensables para construir, usar e interpretar textos en sus contextos específicos. Además, Gómez (1997a) agrega que las unidades verbales son tales en la medida en que han sido construidas para ser interpretadas y cumplan con las condiciones mínimas para que ello se concrete, aun cuando no sea fácil hacerlo, esto es cuando no se hayan dado todavía las condiciones contextuales requeridas o su contenido pueda resultar críptico. Por lo anterior, define al texto como una creación humana convencional puramente simbólica que está orientado a significar. Toda su estructura física está destinada a ser utilizada como material significativo.

Para Gómez (1997b), además, se deben reconocer en el acto de comunicación tres instancias textuales estrechamente vinculadas, pero diferentes. De este modo, propone distinguir entre:

- a) Pre-texto: bosquejo o anteproyecto de texto en elaboración. Representa la etapa de gestación. Quien intenta comunicar algo organiza en su mente o anota en sus archivos personales un material todavía no autorizado.
- b) Texto 1: Constructo verbal impreso -texto escrito- que la tradición ha considerado como tal, enfatizando sus rasgos externos, secundarios, pero valiosos en la medida en que favorecen la propuesta o el descubrimiento de los sentidos potenciados en él. Conforman la carga semántica asignada por el autor que ha de ser captada por los intérpretes.
- c) Texto 2: Corresponde a otra etapa de textualización personal del mismo proceso comunicativo, ahora en la interioridad del comprendedor, quien, con recursos propios, ha de actualizar el sentido del texto 1, reconstituyendo su coherencia semántica. Se trata de la instancia asequible sólo al comprendedor, quien deberá, a su vez, reelaborarla, si desea o está obligado a divulgarla entre otros interlocutores

En definitiva, para Peronard y Gómez (1985), el aspecto caracterizador y distintivo de la comprensión lingüística, en tanto acto mental y acto lingüístico es el requisito de la comunicabilidad. Así, para Gómez (1997a: 101-102), el proceso de comprensión se sintetizaría de la siguiente manera:

“Desde la percepción del texto estructurado materialmente hasta su posesión comprensiva, las fuerzas del sujeto comprendedor deben converger de manera que su conciencia vaya descubriendo la red de relaciones significativas que lo manifiestan como unidad.

El resultado alcanzado, comunicable a otros individuos, puede ser reelaborado desde otra perspectiva; puede enriquecer la interioridad del individuo en cuanto conocimiento; puede servir de punto de partida para nuevas creaciones mentales, sobre todo si lo acompañan ingredientes emocionales, afectivos o intencionales persistentes”.

El valor de este trabajo está en la configuración de un modelo integral de comprensión lingüística, el cual incluye componentes lingüísticos, cognitivos, sociales, biológicos y afectivos. Cabe hacer notar, a modo de especificación terminológica, que la definición dada por los autores para pre-texto y texto 2 son lo que nosotros entenderemos como representación psicosociodiscursiva de un texto. Este proceso de representación se realiza

en la mente del sujeto productor/comprendedor y será esencial al momento de producir o comprender un texto (entendido como Texto 1), puesto que el sujeto seleccionará a partir de su competencia discursiva aquellas representaciones psicológicas (afectos y emociones), cognitivas (percepción, atención, memoria) y sociales (interacciones sociales y reconocimiento de prácticas sociales por medio de la experiencia), que le serán útiles, tanto para la interpretación como para la producción del texto.

Por último, para Peronard y Gómez (1985) es importante contar con el sustrato fisiológico del cerebro que será, en definitiva, el encargado de actualizar gran parte de las capacidades o procesos; no obstante, dejan en claro que las estructuras y actividades cerebrales no pueden confundirse con determinados estados o procesos psicológicos de los cuales está consciente el sujeto comprendedor. Dentro de los componentes de sustrato fisiológico, Peronard y Gómez (1985) incluyen inicialmente el fonológico, el gramatical, el ideacional y la memoria. Lo anterior permite inferir, como plantea Parodi (2003), una propuesta de tipo modular y de procesamiento automático a partir de ciertos conocimientos específicos no conscientes o inaccesibles a la conciencia de forma directa, a saber, instancias fonológicas y gramaticales.

1.2.4 Texto/discurso como fenómeno sociocognitivo

Otros autores, prestando mayor atención a los aspectos sociales que a las definiciones anteriores, conceptualizan el objeto texto desde una perspectiva sociocognitiva, en la que se enfatiza la idea de texto como multidimensionalidad del conocimiento (Heinemann & Viehweger, 1991; Antos, 1997), por medio de la cual los interlocutores pueden manifestar sus individualidades en un marco social compartido. Antos (1997) especifica una concepción de texto procesual y dinámica, estableciendo que, por un lado, los textos son sólo “estaciones intermedias” para la creación de otros textos – porque pueden conformar conocimiento siempre solo de manera selectiva -, por otro lado, son punto de partida para el procesamiento receptivo del conocimiento basado textualmente.

Esta idea de “estación intermedia”, potencial o real, que se le adjudica al texto, es central en el ámbito de los textos que comunican ciencia. Así, por ejemplo, la comunicación textual iría desde una situación inicial en la que a partir del artículo científico, se intenta la

validación por parte de los colegas-pares, hasta la comunicación del mismo suceso al público en general, en el caso de los textos de divulgación de la ciencia (Antos, 1997; Gotti, 2003).

Una concepción de texto más actualizada en esta perspectiva sociocognitiva es la que presenta Ciaspucio (2003). La autora concibe al texto como:

“[...] recursos primordiales de constitución de conocimiento, cuyas formas, estructuras, contenidos, funcionalidades, etc, responden a desarrollos socioculturales y, por lo tanto, más allá de los sistemas lingüísticos particulares, pueden exhibir en los niveles más globales (géneros, registros, estilos, etc.) características específicas debidas a la sociedad o comunidad particular en que son constituidos y empleados” (Ciaspucio, 2003: 23).

A partir de esta noción es posible concebir al texto como un producto del conocimiento de una comunidad discursiva. En este sentido, el texto exhibirá las particularidades propias de la comunidad en la cual se crea. Por lo mismo, si extendemos esta noción, cada género textual, registro o estilo dará cuenta de las marcas léxicas, sintácticas o semánticas que caracterizan a una comunidad discursiva particular (ver Swales, 1990; Warta, 1996; López, 2002; Gotti, 2003).

1.2.5 El discurso como práctica social

Como hemos podido notar, a partir de la presentación de los diversos momentos del *continuum*, la noción de texto/discurso se ha ido permeando cada vez más de una orientación cognitivo-social, pudiendo afirmarse que es este rasgo el que permite identificar al discurso como polo opuesto a texto en el *continuum* (ver Figura 1).

Cabe señalar que los autores que entienden el discurso como práctica social, en su mayoría, se basan en los postulados de Bajtín (1982), para quien el discurso es concebido como:

“El producto de la interacción de dos individuos socialmente organizados; añade que aún en el caso de no existir interlocutor <<real>> (como en el discurso interior o pensamiento) éste es reemplazado por la clase social a la cual el locutor real pertenece” (Bajtín, 1982:122)

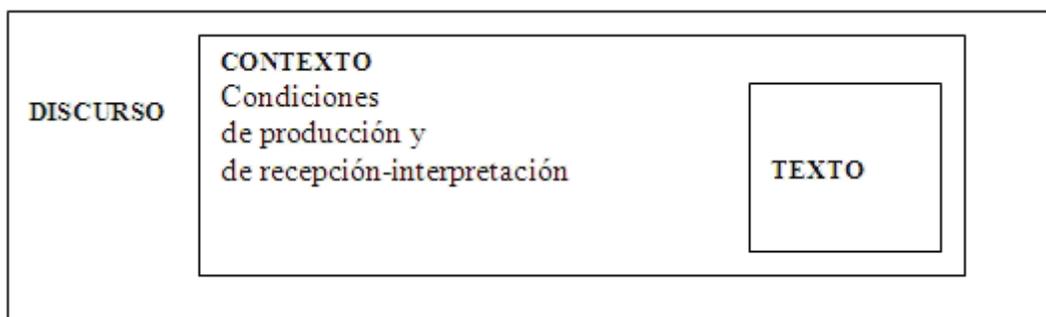
Así pues, la estructura del discurso será determinada por las condiciones reales en las que el discurso se realiza, es decir, sobre todo por la situación social inmediata y el medio social más amplio; por el hecho de que procede de alguien y se dirige a alguien. El discurso, desde esta perspectiva, es en definitiva el producto de la integración del locutor y el auditor, lo cual refleja la naturaleza social del discurso y convergencia en él del individuo y la sociedad.

Desde una perspectiva lingüístico-textual, Adam (1992), en cuanto a la relación texto-discurso, plantea que el discurso debe ser entendido como el texto más las condiciones de producción, en tanto que el texto es el discurso menos las condiciones de producción. Sin embargo, más tarde, el mismo Adam (1999: 39) se encargará de corregir esta idea planteando que:

“La formule [...] n’est pas une formule d’addition et de soustraction du contexte. Cette formule ne doit pas opérer la décontextualisation que je préconisais alors. Il s’agit d’une formule d’inclusion du texte dans le champ plus vast de pratiques discursives qui doivent elles-mêmes être pensées dans la diversité des genres que’elles autorisent et dans la diversité des genres que’elles autorisent et dans leur historicité”.

Como consecuencia de lo anterior, propone una esquematización algo distinta que representaría mejor el lugar del discurso y del texto en el acto discursivo:

Esquema 1. Relación entre discurso, contexto y texto (tomado de Adam, 1999: 39)



Una vez adoptado este nuevo esquema de las relaciones de inclusión existente entre discurso, contexto y texto, Adam (1999: 40) propone las siguientes definiciones que

permitirán distinguir el texto del discurso. En el caso de texto lo define como objeto abstracto y como objeto concreto:

a) “Le texte comme objet abstrait est l’objet d’une théorie générale des agencements d’unités (ce qu’on appellera la texture pour désigner les faits microlinguistiques et la structure pour les faits macrolinguistiques) au sein d’un tout de rang complexité linguistique plus ou moins élevé. Cet objet abstrait était celui des « grammaires de textes », il reste, dans une autre configuration épistémologique, l’objet théorique de la linguistique textuelle”.

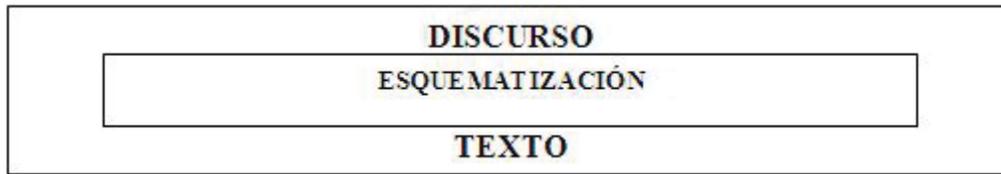
b) “Les textes comme objets concrets, matériels, empiriques. Chaque texte se présente comme un énoncé complet, le résultat toujours singulier d’un acte d’énonciation. C’est, par excellence, l’unité de l’interaction humaine. Chaque texte étudié sera pour nous l’objet d’une analyse textuelle. Opérant sur un événement singulier de parole, l’analyse textuelle ne peut pas faire l’économie de l’articulation du textuel et du discursif”.

Por otra parte, al referirse a discurso, plantea que:

c) “Parler de discours, c’est ouvrir le texte (entendido como objeto concreto), d’une part, sur une situation d’énonciation-interaction toujours singulière et, d’autre part, sur l’interdiscursivité dans laquelle chaque texte est pris –en particulier celle des genres. À l’analyse de/du discours, qui est une théorie générale de la discursivité, nous préférons substituer une analyse DES discours attentive à la diversité des pratiques discursives humaines”.

Cabe hacer notar que Adam (1999) plantea que la separación entre texto y discurso es esencialmente metodológica. Tal separación es fruto de programas de investigación que ponen el acento sobre los diferentes componentes de los productos de la actividad lingüística humana y se confieren, para su quehacer, de metodologías propias. De este modo, y en un afán de demostrar la íntima relación de complementariedad existente entre discurso y texto, propone utilizar una noción intermedia denominada esquematización, para la cual plantea Adam (1999: 102) que “*Penser tout texte comme une schématisation, c’est réunir en un seul concept, l’énonciation comme processus et l’énoncé comme résultat*”. Así, hablar de texto o discurso sería, para este autor, hacer siempre alusión más al resultado de las prácticas discursivas que a las operaciones complejas, inseparablemente psicosociales y verbales, que la han producido. El esquema que propone, añadiendo la noción de esquematización, es el siguiente:

Esquema 2. Esquematización como resultado de las prácticas discursivas (tomado de Adam, 1999: 41)



A partir de este esquema, propone que el campo del análisis del discurso estará centrado en la formación sociodiscursiva (interdiscursos), la interacción sociodiscursiva (paratextos y géneros discursivos) y las acciones perlocutivas del lenguaje. En tanto que el campo de la lingüística textual estará centrado en el estudio de: la textura frástica y transfrástica (gramática y estilo), la estructura composicional (secuencias y planes textuales), la semántica (representaciones discursivas), la enunciación (anclaje situacional) y la orientación argumentativa y actos de discurso (dimensión ilocutoria).

Como observamos, las distinciones que plantea Adam (1999) nos permiten sustentar nuestra idea de la existencia de un *continuum* conceptual, en tanto que distingue la noción de discurso, orientada exclusivamente hacia una dimensión social y la de texto, orientada hacia una dimensión predominantemente lingüística, semántica y estructural. A pesar de lo anterior, la orientación hacia la enunciación y hacia lo argumentativo ilocutorio, que también plantea para el análisis textual, nos parece más propio de la actividad discursiva, ya que ambas orientaciones se manifiestan exclusivamente en el intercambio discursivo entre sujetos afectos a un marco social determinado.

Por otra parte, de manera más explícita aún, Charaudeau y Maingueneau (2002), al establecer la diferencia entre texto y discurso, sostienen que el texto es una realidad concreta, (empírica) correspondiente a un discurso efectivamente emitido. En este sentido, el texto representa el resultado material del acto de comunicación.

Una última perspectiva, brevemente esbozada, es la que representa el polo más social del *continuum* conceptual. Esta es la planteada por Foucault (1999) para quien el discurso "se refiere a los acontecimientos, no al código". Esto significa que el concepto de discurso no tiene relación con un conjunto de signos, sino con el establecimiento de un orden de

elementos en función de los cuales se define un saber y una práctica social. Para este autor, el discurso se circunscribe al campo de la experiencia y al del saber posible, definiendo la manera de ser de los objetos que aparecen en este campo. Así, esta orientación concibe al discurso como una práctica social creadora de un orden específico que, si bien se halla indisolublemente asociado a los otros órdenes de la vida, no está determinado por éstos. Muy por el contrario, es en el orden del discurso donde se elaboran socialmente los modelos de organización e interpretación de la realidad, la historia y la experiencia, de modo que éstas puedan construirse y ser vividas como regularidades estructuradas. Por otra parte, para Foucault (1999) el texto no es considerado en sí mismo, sino como parte de una institución reconocida que define para un área social, económica, geográfica o lingüística dada las condiciones de ejercicio de la función enunciativa de una colectividad.

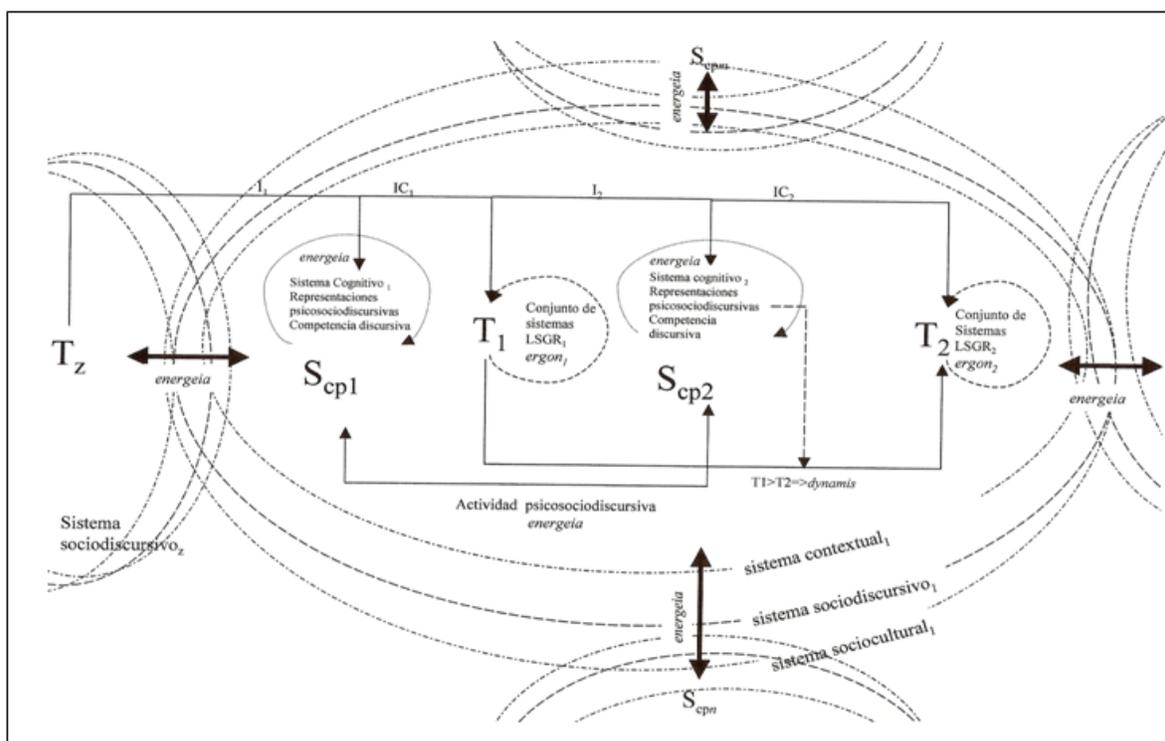
Por último, Iñíguez (1993) refuerza la idea foucaultiana planteando que el discurso es lenguaje como práctica social determinada por estructuras sociales, lo cual significa: primero, que el lenguaje es una parte de la sociedad y no algo externo a ella; segundo, que es un proceso social; y tercero, que es un proceso condicionado social e históricamente, en el mismo sentido que lo son otras partes o procesos no lingüísticos. De este modo, no hay una relación externa "entre" lenguaje y sociedad, sino interna y de dualidad estructural. El lenguaje es una parte de la sociedad, los fenómenos lingüísticos son fenómenos sociales y los fenómenos sociales son (en buena parte) fenómenos lingüísticos.

1.3 Hacia una delimitación metodológica de los conceptos de texto y discurso

Con el objetivo de delimitar metodológicamente los conceptos tratados en el *continuum*, proponemos en este apartado recurrir a la concepción "energética" del lenguaje proveniente de la tradición aristotélica, a través de la propuesta humboldtiana, y actualizada por Coseriu (1986), según la cual el lenguaje es concebido como una actividad humana y puede enfocársela, en cuanto tal, como actividad en sí misma, *energeia*, o en cuanto actividad en potencia, *dynamis*, o en cuanto actividad realizada en sus productos, *ergon*.

Estas distinciones nos son muy útiles para caracterizar las relaciones que hemos mostrado entre las múltiples orientaciones desde la cual se ha definido texto y discurso, ya que es posible categorizar todas las distinciones conceptuales en las cuales se reconoce actividad de lenguaje realizada o potencialmente realizada por los sujetos (*energeia* y *dynamis*) bajo el concepto de “discurso”, incluyendo en este concepto, la dimensión cognitiva del procesamiento lingüístico, la interacción comunicativa por medio del lenguaje verbal y las prácticas sociales. Por otra parte, el producto o *ergon* de la actividad lingüística, en cuanto representación lingüística material externa de esta actividad, es lo que entenderemos bajo el concepto de “texto”.

El esquema 3, que proponemos a continuación, daría cuenta de las múltiples relaciones que se generan en el proceso discursivo entre el discurso (como actividad psicosocial del lenguaje y como actividad potencial de conocimiento) y el texto (como producto lingüístico materializado de los discursos).



Esquema 3. Relaciones en el proceso discursivo

Donde:

- T_z = Todos los textos previos que han sido interpretados por el Sujeto comprendedor/productor primario ($ergon_z$).
 S_{pc1} = Sujeto comprendedor/productor primario.
 S_{pc2} = Sujeto comprendedor/productor secundario.
 S_{pcn} y S_{pcm} = Sujetos comprendedores/productores pertenecientes a otros sistemas sociodiscursivos o socioculturales.
 T_1 = Texto primario, representación lingüística material externa producida por S_{pc1} . ($ergon_1$)
 T_2 = Texto secundario, representación lingüística material externa producida por S_{pc2} . ($ergon_2$)
 $T_1 > T_2 \Rightarrow dynamis$ = Proceso en que un texto se convierte en actividad discursiva potencial para generación de nuevo conocimiento.
 I_1 = Integración de sentido a partir de la interpretación de todos los posibles T_0 realizada por S_{pc1} .
 IC_1 = Intención comunicativa de S_{pc1} .
 I_2 = Integración de sentido a partir de la interpretación de T_1 realizada por S_{pc2} .
 IC_2 = Intención comunicativa de S_{pc2} .
 $LSGR_1$ = Sistemas Léxico, Semántico, Gramatical y Retórico Estructural de T_1 .
 $LSGR_2$ = Sistemas Léxico, Semántico, Gramatical y Retórico Estructural de T_2 .
 \longrightarrow = Relación de dependencia.
 \longleftrightarrow = Relación de interdependencia.
 \dashrightarrow = Sistema cognitivo psicodiscursivo.
 \cdashrightarrow = Relación de inclusión, constitución de un sistema textual compuesto por los subsistemas LSGR.
 \cdashrightarrow = Relación de inclusión, constitución de un sistema contextual sociodiscursivo.
 \cdashrightarrow = Relación de inclusión, constitución de un sistema sociocultural.

En función de lo anterior, planteamos que la relación entre texto y discurso en el proceso discursivo en esta investigación es concebida, siguiendo el Esquema 3, como una actividad psicodiscursiva (*energeia*), en la cual el S_{pc1} lleva cabo un proceso de I_1 , proceso entendido como la captación del sentido (a partir de sus competencias discursivas y sus representaciones psicodiscursivas) de todos los textos orales-escritos que ha experimentado. De este modo, a partir de una IC_1 , este sujeto es capaz de producir un T_1 con un significado particular utilizando los subsistemas LSGR del sistema lingüístico, para ser leído por el S_{pc2} , quien, a su vez, realiza un proceso de captación de sentido a través de la I_2 (utilizando sus competencias discursivas y sus representaciones psicodiscursivas) y comunica lo comprendido a través de un T_2 .

Esta actividad psicodiscursiva (entre S_{pc1} y S_{pc2} por medio de T_1) se desarrolla siempre en un sistema contextual particular del sistema sociodiscursivo. Este sistema contextual se

concibe como la situación temporal y espacial particular y única, en la cual se produce la actividad discursiva entre los sujetos y que está condicionada por el sistema sociodiscursivo en el cual se realiza. El sistema presenta características específicas debidas a la sociedad o comunidad particular en la que los discursos son constituidos y empleados (géneros discursivos); y cuyas formas, estructuras, contenidos, funcionalidades, etc, responden al desarrollo de los sistemas socioculturales.

Todas estas características se ven materializadas en las representaciones lingüísticas externas al individuo, en otras palabras, en los productos lingüísticos generados por los sistemas léxico, semántico, gramatical y retórico estructural (*érgon*). Este producto lingüístico, a su vez, se convierte en actividad en potencia (*dynamis*), entendida como un “estado intermedio” en proceso de construcción de conocimiento. Cabe señalar, además, que todo proceso sociodiscursivo está en interrelación constante con otros sistemas contextuales sociodiscursivos, otros sistemas sociodiscursivos y otros sistemas socioculturales, en los cuales se llevan a cabo diversos procesos discursivos que coexisten temporal y atemporalmente.

De este modo, en esta investigación entenderemos al discurso como una actividad (*energeia*) que se produce tanto a nivel de los procesos psico- y sociodiscursivos como a nivel de actividad caracterizadora del proceso en el que un texto se convierte en actividad discursiva potencial (*dynamis*) para la generación de nuevo conocimiento. En tanto que, por otra parte, el texto es concebido como la representación lingüística material de las representaciones psicosociodiscursivas que realiza un sujeto comprendedor-productor, inmerso en una actividad psicosociodiscursiva afectada por el sistema contextual en el que se lleva a cabo la actividad, sistema que es dependiente de un sistema sociodiscursivo particular el cual, a su vez, es circundado por un sistema sociocultural particular.

Para finalizar este apartado, cabe señalar que el Esquema 3, como representación gráfica del proceso discursivo, permite comprender el texto como producto de una competencia discursiva que faculta al sujeto productor/comprendedor materializar (a través de la construcción de una representación lingüística externa con significado propio) las representaciones psicosociodiscursivas, que, como sujeto, se ha construido a partir de una construcción coherente de sentido a partir de los sistemas contextuales, sociodiscursivos y

socioculturales en los cuales participa. Cabe hacer notar que tal representación es realizada a partir de los múltiples aportes teóricos revisados anteriormente. Por lo anterior, este esquema es deudor de ese conocimiento especializado, aunque necesariamente reelaborado (y quizás cercenado), para los efectos de esta investigación. En síntesis, reconocemos que la polisemia y el uso, muchas veces, laxo e incontrolado de los conceptos de texto y discurso, está dado tanto por los diversos enfoques de investigación como por la complejidad del fenómeno discursivo, el cual es multisistémico y en constante interactividad con diversos multisistemas, lo que difumina los límites de cada uno de estos sistemas.

1.4 El discurso especializado de la ciencia

En el apartado anterior se presentó un esquema que daría cuenta del proceso discursivo, intentando con ello acotar los conceptos de discurso y de texto, entendiéndolos como instancias distintas de este proceso. Ahora bien, en lo que sigue caracterizaremos el discurso especializado de la ciencia, bajo el prisma aportado por el Esquema 3. En este sentido, este tipo de discurso se realiza en un sistema sociocultural determinado (en nuestro caso, el occidental de habla hispana), que se caracteriza por ser parte de un sistema sociodiscursivo particular (en nuestro caso, el de la ciencia) y que se desarrolla en sistemas contextuales específicos (en nuestro caso, la investigación científica). Los sujetos de la actividad psicosociodiscursiva son necesariamente los científicos de alguna área del conocimiento, que construyen textos científicos de distinto tipo (en nuestro caso, artículos de investigación científica).

Sabemos que en la ciencia no existe siempre consenso en la denominación de los objetos de estudio, debido fundamentalmente a la focalización y delimitación que deben realizar los autores al intentar conceptualizar el objeto a estudiar. Normalmente, los abordajes son múltiples en razón de supuestos teóricos divergentes. El concepto de “discurso especializado” no es la excepción. Este ha sido denominado de múltiples maneras, por ejemplo: discurso académico, discurso especial, discurso profesional, discurso técnico, discurso institucional, etc. Alcanzar un relativo orden terminológico y lograr una visión más o menos homogénea tampoco resulta fácil (Ciapuscio, 2000; López, 2002). Ahora

bien, según nuestra perspectiva, podemos afirmar que todas estas caracterizaciones del discurso están realizadas considerando distintos sistemas sociodiscursivos.

Por otra parte, determinar de forma discreta si un texto se clasifica como de especialidad o de tipo general es, sin duda, un problema teórico y descriptivo (Schröder, 1991; Parodi, 2004). Hoy en día, la postura predominante está en favor de un continuum de textos que se distribuyen de manera progresiva desde un dominio altamente especializado hasta otro extremo mucho más divulgativo y general (Gläser, 1982; Schröder, 1991; Halliday & Martin, 1993; Jeanneret, 1994; Peronard, 1997; Ciapuscio, 1994, 2000; Cabré, 2002; Parodi, 2004), aceptando que la realidad no se circunscribe a la idea de límites discretos sino más bien difusos (Lakoff, 1987). En este sentido, Parodi (2004: 10) plantea que:

“[...] es un hecho que establecer límites precisos entre un tipo de texto y otro es una cuestión de envergadura. Sin importar el foco atencional en uno u otro criterio clasificatorio, siempre existirán casos mixtos o límites; sin embargo, parece ser que el discurso especializado corresponde a una categoría reconocible para cualquier hablante de español”.

Gotti (2003), siguiendo la idea del *continuum*, plantea en relación con la naturaleza multi-dimensional del discurso especializado que no existe homogeneidad entre los diferentes lenguajes especializados. Argumenta que las variaciones disciplinares producen no solo connotaciones léxicas especiales, sino que también a menudo influyen en otras opciones (morfosintácticas, textuales y pragmáticas), teniendo además repercusiones en las peculiaridades epistemológicas, semánticas y funcionales de una variedad de discurso especializado.

De esta manera, las diferencias entre los discursos permiten reconocer diferencias de nivel en el discurso especializado, ya que por ejemplo, la sola presencia de un especialista no es suficiente para asegurar el uso especializado del lenguaje. De hecho, Gotti (2003) distingue al menos tres niveles diferentes en los cuales el experto podría referirse a un tópico relacionado con su profesión. En el primer nivel, el especialista se comunica con otros especialistas para debatir tópicos relacionados con el campo disciplinar, para describir un proyecto de investigación, para entregar resultados, para explicar el uso de ciertos equipos, etc. Si los oyentes comparten una cantidad considerable de conocimiento, el hablante

especialista puede hacer un uso frecuente de terminología especializada, cuyo sentido está garantizado (para nosotros, un sistema sociodiscursivo especializado de la ciencia). En el segundo nivel, el especialista se dirige a no-especialistas para explicar nociones referentes a su disciplina (sistema sociodiscursivo divulgativo-didáctico de la ciencia). En estos casos, por motivos educacionales, el significado del léxico especializado es ilustrado cada vez que este ocurra por primera vez. Textos típicos de esta naturaleza son los textos de estudio y los manuales de instrucción. En el tercer nivel, el especialista provee información de naturaleza técnica, principalmente a través de un vocabulario cotidiano, haciendo uso de la experiencia cotidiana del lego cuando introduce un concepto especializado (sistema sociodiscursivo divulgativo de la ciencia). El propósito aquí es abarcar una audiencia amplia, como la de los diarios y/o artículos de revistas en los cuales se presenta información técnica o científica (Gotti, 2003).

En esta investigación, hemos optado por el término “discurso de especialidad” o “especializado”, centrándonos en el primer nivel planteado por Gotti (2003), pues nos interesa el discurso de la ciencia en un sistema contextual comunicativo entre especialistas que dan cuenta de sus investigaciones científicas, a través de “textos especializados”, entre los cuales nos interesa destacar el artículo de investigación científica (Gläser, 1982, 1993; Halliday & Martin, 1993; Cabré, 1999, 2002; Ciapuscio, 2003; Gotti, 2003). Desde esta óptica, coincidimos con diversos investigadores para quienes el discurso de especialidad se conforma por un conjunto de textos que involucran tópicos que son prototípicos de un área determinada del conocimiento científico y se ajustan a convenciones particulares de índole lingüística, funcional y situacional, entre otras, de las tradiciones de una comunidad discursiva determinada (Brinker, 1988; Schröder, 1991; Geisler, 1994). De modo más específico, el término “discurso especializado” refleja más claramente el uso particular del lenguaje típico de un sistema sociodiscursivo especializado, utilizado por una comunidad determinada, extendiéndose a las áreas de conocimiento y práctica académicas, profesionales, técnicas y ocupacionales. Esta perspectiva incluye tanto al tipo de usuario y el dominio de uso, como a la aplicación especial que se hace del lenguaje en estos escenarios (Gotti, 2003).

1.4.1 Discurso especializado y paradigma científico

El discurso especializado, en nuestro caso, se encuentra en relación con el sistema contextual de la investigación científica y la actividad psicosociodiscursiva en la que se producen y transmiten conocimientos. Por lo tanto, la finalidad del discurso especializado es representar el proceso de investigación científica, esto es: definir un objeto, elaborar un método para examinarlo, construir una teoría que explique el funcionamiento del objeto y comprobar la validez de la teoría. Lo anterior no quiere decir que todos los discursos especializados se constituyan de cuatro partes, pues la complejidad de los objetos obliga a los científicos a examinarlos parcialmente y, por lo mismo, a escribir pequeños reportes sobre aspectos parciales del desarrollo del proyecto científico, por ejemplo, reportes sobre puntos meramente metodológicos, disquisiciones terminológicas necesarias para continuar el proyecto científico mayor o problemas de orden lógico inherentes al desarrollo de toda teoría formal (Elejalde, 1998).

De esta manera, es indudable que nuestra noción de discurso especializado centrado en la actividad científica está teñido por la manera de hacer ciencia bajo un concepto positivista, en tanto existe la pretensión de que el discurso científico sea objetivo, certero y que presente una verdad “incuestionable” (Bunge, 1981). Al respecto, Bunge (1981) plantea que los rasgos propios de este paradigma se pueden consignar en tres aspectos básicos:

- a) Concepción de realidad: la realidad existe fuera del sujeto que la estudia y, por tanto, es separada, externa, objetiva. En esa medida, la realidad es cognoscible y se convierte en objeto de estudio en tanto que es comprobable y verificable.
- b) La verdad y la objetividad: si la realidad es un objeto externo al sujeto que la estudia, éste puede alcanzar la verdad en la medida que elabora descripciones y explicaciones objetivas, fieles y literales de esa realidad. Para ello, no hay más que ser objetivo; es decir, no permitir que la subjetividad intervenga en el estudio, interpretación ni elaboración de hipótesis o leyes. La objetividad es, pues, la búsqueda de la verdad separada del sujeto, quien se convierte en un ser pasivo.
- c) La racionalidad y el método: la racionalidad es el fundamento del método científico. La búsqueda de la verdad objetiva supone que el sujeto investigador actúe de acuerdo a un

plan establecido racional y lógicamente. Esto es, se plantea un problema, se formulan diversas hipótesis, se verifican de acuerdo a la experimentación, recolección, cuantificación y análisis de datos para llegar a conclusiones y establecer reglas generales. De esta manera, los científicos elaboran en sus trabajos de investigación y sus escritos una reconstrucción formal, racional y objetiva de la realidad.

Estos rasgos caracterizadores del paradigma científico positivista, sin lugar a dudas, tendrán su correlato en la forma en que la actividad discursiva se desarrolla y se manifiesta lingüísticamente en los textos. De este modo, tanto el discurso como los textos (productos de estos discursos), originados bajo la perspectiva de este paradigma positivista, en general, se caracterizarán por su aplicabilidad en el mundo real. Los discursos especializados, creados por los investigadores, son altamente referenciales, pues hablan de diversos aspectos de una realidad circundante a la que examinan minuciosamente; su esquema discursivo es rígido y conservador, puesto que su finalidad es mostrar interpretaciones alternativas del mundo construidas de acuerdo a las normas de procedimientos aceptadas en los medios científicos y académicos del mundo. Aunque ni un discurso hablado ni uno escrito son iguales a una demostración lógico-formal, se considera deseable que el discurso especializado no sólo busque la justeza y consistencia lógica de la demostración, sino también que se estructure en función de la intención persuasiva del autor, es decir, de la argumentación (Geisler, 1994; Hyland, 1998; Lo Cascio, 1998).

En términos más específicos, la caracterización de estos discursos especializados ha sido tarea de investigación ampliamente desarrollada por los investigadores del mundo anglosajón, especialmente por los investigadores orientados al estudio del inglés con propósitos especiales (por ejemplo, Sager, Dungworth & McDonald, 1980; Bazerman, 1988; Swales, 1990; Hyland, 1998, 1999, 2000; Salager-Meyer, 1991, 1992, Martín, 2003, Gotti, 2003). Sin embargo, para el español, los estudios que caracterizan el discurso especializado de la actividad científica y académica son más bien recientes (por ejemplo, Calsamiglia, 1998; Bolívar, 2000; Ciapuscio, 2000, 2003; Cassany, López & Martí, 2000; Moyano, 2000; Ciapuscio & Otañi, 2002; López, 2002; Mogollón, 2003; Parodi & Gramajo, 2003; Parodi, 2004, 2005a,b).

En este sentido, una caracterización bastante acotada del discurso especializado de la actividad científica en español es la que presenta Mogollón (2003). Este autor caracteriza lo que él denomina “lenguaje especializado”, en otras palabras, el lenguaje característico de la ciencia, según la presencia de 7 rasgos particulares: a) transmite conocimientos especiales, b) es fundamentalmente escrito, c) usa sistemas de representación no verbales, d) es expresión de racionalidad, e) busca precisión, concisión y claridad, f) busca neutralidad e impersonalidad y g) usa vocabulario de carácter técnico.

A continuación, siguiendo a Mogollón (2003), desarrollaremos cada uno de estos rasgos caracterizadores del “lenguaje especializado”, ejemplificando con un resumen de un artículo de investigación científica los aspectos relevantes de cada una de estas características.

a) Transmite conocimientos especiales: Como hemos ya indicado, es extendida la idea de que lo que hace que un texto sea científico es la temática tratada. Se transmiten conocimientos adquiridos sobre el mundo objetivo, se hace referencia a un objeto, a una realidad que está fuera del sujeto que conoce. Usualmente, se aborda un único tema, para cuyo desarrollo predomina la función referencial (en la medida que lo relevante del proceso comunicativo es el mensaje y no su emisor ni su receptor) y la denotación (puesto que importa lo que la palabra explícita y directamente dice y no lo que el emisor o el receptor quieran pasar por sus filtros personales). El objeto de estudio se puede verificar. Este sentido de inamovilidad del objeto real se logra con un recurso lingüístico muy específico: la nominalización; de manera que lo que es característico o activo, y en esa medida cambiante, se convierte en pasivo y estable y, por tanto, observable. Al ser objeto se vuelve estático y observable, susceptible de estudio. Por la misma razón, hay un predominio del modo indicativo en tanto que modo de expresión en el plano de lo real.

Antes de seguir avanzando en las características propias del discurso especializado de la ciencia, abriremos un breve paréntesis para referirnos a la distinción de rasgos lingüísticos propios de un tipo de discurso.

Un tipo de investigación que se ha preocupado de determinar empírica y cuantitativamente los rasgos propios de un tipo o registro de discurso se ha basado principalmente en análisis

multirasgos y análisis multidimensionales. Este tipo de investigación fue desarrollada originalmente para el estudio analítico de las variaciones entre registros, con el objetivo particular de identificar los patrones lingüísticos sobresalientes y en co-ocurrencia en una lengua. Además, este tipo de análisis permite comparar registros orales y escritos en un espacio lingüístico definido por aquellos patrones de co-ocurrencia (Biber, 1988; Biber, Conrad & Reppen, 1998; Conrad & Biber, 1998, Parodi, 2004, 2005a,b; Louwerse, McCarthy, McNamara & Graesser, 2004). Subyace a esta metodología de investigación el principio de la co-ocurrencia de rasgos lingüísticos y su relación funcional. Esto quiere decir que un registro no se puede caracterizar por un rasgo único, sino que se caracterizará por patrones de co-ocurrencia de rasgos que se interpretan en término de sus funciones situacionales, sociales y cognitivas comunes, pues se asume que la ocurrencia de los mismos rasgos lingüísticos en textos determinados se debe a que cumplen funciones compartidas específicas. En este sentido, Parodi (2004: 19) explica:

“[...] se entiende que un grupo de rasgos co-ocurre frecuentemente en ciertos textos porque ellos son usados para expresar un conjunto de funciones comunicativas. Estos conjuntos de rasgos, determinados estadísticamente por medio de técnicas multivariadas, son lo que se denomina una dimensión; en otras palabras, una dimensión se constituye mediante la interpretación funcional del conjunto de los rasgos co-ocurrentes (factor), agrupados –entre otros- mediante el análisis factorial”.

Es interesante destacar el trabajo realizado por el equipo de investigación lingüística de la Pontificia Universidad Católica de Valparaíso en esta línea de investigación (Parodi & Gramajo, 2003; Marinkovich & Cademártori, 2004; Parodi & Venegas, 2004; Parodi, 2005a,b), ya que es el primer trabajo que se desarrolla con esta metodología para el español utilizado en los textos que leen un grupo de alumnos de secundaria de la educación técnico-profesional.

Parodi (2005a,b), a partir de un estudio multirasgos y multidimensional sobre unos *corpora* de textos orales, literarios y técnico-científicos, plantea que la co-ocurrencia de rasgos morfosintácticos caracterizadores de este último corpus es la compuesta por los rasgos: verbo modal de obligación, subjuntivo, nominalización, frase preposicional como complemento del nombre y participio en función adjetiva. Todos estos rasgos son incluidos en lo que el autor denomina “Dimensión Foco Informacional”, la cual se caracterizaría por

una orientación hacia la informatividad, entendida ésta como la concentración de información en unidades y estructuras lingüísticas compactas, que presentan los datos lo más concisa y precisamente posible. Cabe señalar que el corpus técnico-científico, en este caso, corresponde mayoritariamente a textos del tipo divulgativo didáctico, lo que explica la presencia de los dos primeros rasgos. Por otra parte, cabe hacer notar que en el estudio de Parodi (2005a) el modo indicativo co-ocurre con rasgos como verbos privados, pronombres de primera persona singular, pretérito indefinido, verbos modales de volición, desinencias de primera persona singular; los cuales, entre otros, conforman la denominada “Dimensión Foco Compromiso”. Esta dimensión es interpretada aduciendo que la 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.

Ahora bien, en función de la investigación basada en análisis multidimensional y la presentación de ciertos rasgos característicos del discurso especializado (nominalización y modo indicativo), expuesta por Mogollón (2003), parece limitada a estudios no correlacionales y de trabajo con textos ejemplares. Esto explicaría que en el estudio multidimensional de (2005a) el modo indicativo no sea parte de la “Dimensión Foco Informativa” sino que de la “Dimensión Foco Compromiso”.

Por último, es discutible el estatus del rasgo modo indicativo, como rasgo característico del discurso especializado de la ciencia, ya que la presencia del modo indicativo como propio de una función referencial o de expresión de la realidad aducida por Mogollón (2003) sería compartida por otros tipos de discursos. A nuestro parecer, la presencia de un rasgo o dos no bastan para sustentar la idea de que tales rasgos son propios del discurso especializado. Por lo mismo, se hace necesario probar contrastivamente con otros tipos de discursos si existe correlación estadística significativa entre la nominalización y el modo indicativo en el lenguaje especializado de la ciencia.

Cerramos aquí el paréntesis abierto más arriba para ejemplificar en una muestra de texto de nuestro corpus ARTICO (artículos de investigación científica originales) la presencia de los rasgos mencionados tanto por Mogollón (2003) como por Parodi (2005a).

La muestra corresponde al resumen del artículo de Yáñez y Berndt (2000). Este artículo forma parte del área de las ciencias exactas y está incluido en el corpus ARTICO como BSCQ1-2000 45(4) (ver Anexo1):

“Un sistema de flujo de alta eficiencia en espectrometría atómica (HPF-AS) es implementado mediante la combinación de HPLC y un nebulizador hidráulico de alta temperatura y presión (HT-HHPN). Este sistema se aplicó para la separación de la matriz en muestras de salmueras de la industria de cloro-soda, lográndose la determinación de trazas de metales pesados (mg/L) en estas salmueras sin interferencias de la matriz. La separación de la matriz de NaCl se realiza en una columna de HPLC (C18) después de complejar los metales pesados con el ácido trans-1,2-diaminociclohexano, N,N,N',N'-tatraacético (CDTA). La separación de la matriz y la determinación de los metales puede ser realizada en un tiempo inferior a 1 minuto. Los límites de detección de Cd(II), Co(II), Cu(II), Mn(II) y Zn(II) en FAAS fueron mejorados en más de un orden de magnitud en comparación con la determinación convencional o directa mediante un nebulizador neumático” (Yáñez & Berndt, 2000: 535).

En este resumen nos percatamos de que el hecho observable está consignado al inicio del resumen “Un sistema de flujo de alta eficiencia en espectrometría atómica”. En cuanto a los rasgos lingüísticos notamos que se utiliza una buena cantidad de nominalizaciones deverbales (por ejemplo, combinación, presión, separación, determinación, separación, detección, comparación). Por otra parte, todos los procesos verbales están en modo indicativo. En cuanto la co-ocurrencia de rasgos de la denominada “Dimensión Foco Informacional” (Parodi, 2005a), se nota fundamentalmente la co-ocurrencia de los rasgos nominalización, frase preposicional y participio en función adjetiva, por ejemplo en: “[...] *determinación de trazas de metales pesados [...]*”.

b) Es fundamentalmente escrito: En tanto que manifestación de conocimientos especializados, las prácticas de comunicación en la ciencia tienen su canal natural en el discurso escrito. Éste permite la posibilidad de transmisión que necesita la ciencia para su progreso, además de que es el más adecuado para el registro estrictamente formal y normado del campo. Esto podría cuestionarse, porque se cuenta con una gran cantidad de prácticas orales en el ámbito científico, por ejemplo, las conferencias, las clases magistrales, los informes orales, los discursos de reconocimiento. Sin embargo, se afirma

que estas prácticas tienen rasgos que se acercan a la escritura y que, incluso en muchos casos, corresponden a experiencias previas de escritura.

Según Mogollón (2003), el que las prácticas orales en el ámbito de las ciencias 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 discursar más riguroso, lógico, planificado y objetivo.

Un ejemplo de la prevalencia de lo escrito por sobre lo oral en el lenguaje de la ciencia es la mayor validez que los científicos-escritores le asignan a los resultados de una investigación cuando ésta está publicada. Esto es observable en los procesos de citación y de referencias, en los cuales normalmente no aparecen citas de textos orales, ni referencias a exposiciones en congresos, cuyos resultados no hayan sido consignados en las actas del congreso (si las hubiese).

c) Usa sistemas de representación no verbales: Tanto en la práctica oral como en la escrita del discurso científico se insertan otros códigos además del verbal; encontramos 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, esquemas. Se afirma que estos sistemas no verbales remiten a sistemas verbales; incluso, en la mayoría de los casos, van acompañados de explicitaciones verbales.

Un ejemplo claro del uso de fórmulas es el siguiente enunciado del resumen presentado: *“La separación de la matriz de NaCl se realiza en una columna de HPLC (C18) después de complejar los metales pesados con el ácido trans-1,2-diaminociclohexano, N,N,N',N'-tetracético (CDTA)”*.

d) Es expresión de racionalidad: El conocimiento es de fundamentación lógica. Se plantea un problema, se formulan hipótesis, se experimenta y se alcanza una conclusión cercana a la verdad. La racionalidad dirige, pues, el método científico. De tal forma, se cuantifica, se generaliza y, se ordenan y clasifican los datos lógicamente. De allí que en el discurso científico las oraciones predominantes sean las declarativas.

En el caso del ejemplo se observa la siguiente secuencia lógica:

Se implementa “Un sistema de flujo de alta eficiencia en espectrometría atómica”

Se aplica en “la separación de la matriz en muestras de salmueras de la industria de cloro-soda”

Se logra “la determinación de trazas de metales pesados (mg/L) en estas salmueras sin interferencias de la matriz”.

Lo anterior permite que “La separación de la matriz y la determinación de los metales puede ser realizada en un tiempo inferior a 1 minuto”.

Esto mejora “Los límites de detección de Cd(II), Co(II), Cu(II), Mn(II) y Zn(II) en FAAS”

e) Busca precisión, concisión y claridad: Si el discurso científico busca centrarse en la descripción del objeto de referencia y busca el conocimiento de lo verdadero, se hace necesario que la información se transmita de forma precisa, sin posibilidad de errores en los datos y con exactitud en la selección de información.

En el ejemplo se da cuenta de modo preciso cómo se lleva a cabo la separación de la matriz de NaCl.

“La separación de la matriz de NaCl se realiza en una columna de HPLC (C18) después de complejar los metales pesados con el ácido trans-1,2-diaminociclohexano, N,N,N',N'-tatraacético (CDTA)”.

Del mismo modo, es imperioso que ese conocimiento objetivo se transmita con claridad, evitando ambigüedades e interpretaciones erróneas. Para ello, generalmente se organiza la información con un modelo específico y se apoya en recursos visuales. En este sentido, se utilizan marcadores léxicos que ordenan la exposición de las acciones realizadas, tales como “primero”, “luego” y “por último”. Además, es necesaria 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; y se evita la coordinación y subordinación de proposiciones, la expansión de sintagmas

nominales, las construcciones insertadas. Se prefieren las frases y los párrafos cortos, lo cual se manifiesta en un orden sintáctico tradicional en las oraciones: sujeto+verbo+complementos. Por otra parte, estas oraciones son simples, sin subordinación ni inserción de cláusulas.

En el ejemplo, si bien no hay marcadores léxicos que ordenen la exposición de las acciones, estas se pueden inferir claramente (ver el ejemplo del rasgo d). El uso de los adjetivos es suficiente y necesario en cada momento (por ejemplo, nebulizador hidráulico v/s nebulizador neumático). Se mantiene el orden sintáctico canónico y las oraciones tienen una estructura simple.

f) Busca neutralidad e impersonalidad: Si el conocimiento científico es de fundamentación lógica, la objetividad es inherente a la ciencia. Pretende centrarse en la descripción del objeto de estudio; por tanto, es necesario bloquear la presencia del sujeto que conoce. Indudablemente, esto genera en primer lugar el predominio de construcciones impersonales y el uso de la voz pasiva sin agente explícito. De esta manera, el “se” impersonal borra el sujeto investigador, y la voz pasiva sin el complemento agente encabezado por la preposición “por” presenta los objetos como elementos estáticos que son modificados por un sujeto intangible. También es necesaria la utilización de la tercera persona, dado que es la natural para referirse al objeto de estudio y observación. La impersonalidad implica también, sin lugar a dudas, la ausencia de segundas personas de singular y plural, puesto que se supone que la información no está dirigida a ningún sujeto en particular, sino que es conocimiento universal y verdadero de consumo así mismo impersonal. En general, es negado el uso de la primera persona del singular; y la del plural es usada como indicador de modestia más que de presencia del equipo de trabajo.

En el ejemplo, observamos que tres de las cinco oraciones son pasivas (es implementado, puede ser realizada, fueron mejorados) y las otras dos son impersonales con “se” (se aplicó, se realiza).

g) El vocabulario es de carácter técnico: El conocimiento científico se representa en discursos con una terminología especial o específica del campo disciplinario, que corresponde a una elaboración conceptual extremadamente fijada y precisa. Se habla

incluso de tecnolectos en cuanto que derivación de una lengua natural, pero para uso exclusivo de la comunidad científica que produce y consume el conocimiento especializado. Por ejemplo: “[...]**sistema de flujo** de alta eficiencia en **espectrometría atómica**[...]”

El sentido de los términos, frases y definiciones especializados está contextualizado en el discurso que es propiedad de la comunidad.

Un término denomina unívocamente –con poca, por no decir nula, variación sinonímica- a un determinado concepto, el cual queda definido a partir de sus relaciones con otros conceptos con los que conforma una amplia red semántica previamente elaborada y definida (Por ejemplo, los **metales pesados** (mg/L) se complejan con el **ácido trans-1,2-diaminociclohexano, N,N,N',N'-tatraacético** (CDTA). Los **metales pesados** que se detectan en FAAS son **Cd(II), Co(II), Cu(II), Mn(II) y Zn(II)**).

En muchos casos la morfología de la terminología científica acude a préstamos de otras lenguas, a raíces y desinencias grecolatinas (por ejemplo, espectrometría, hidráulico). También es frecuente la redefinición de vocablos comunes (por ejemplo, sistema de flujo de alta eficiencia), el bautizo de fenómenos o procesos con el nombre de su descubridor o inventor, y el uso de siglas (por ejemplo, las normas HPF-AS, HT-HHPN, CDTA, FAAS) símbolos (por ejemplo, detección de Cd(II), Co(II), Cu(II), Mn(II) y Zn(II)), abreviaturas y fórmulas.

1.4.2 Discurso especializado de la ciencia: Distintas aproximaciones

El estudio de los discursos especializados de la ciencia ha sido abordado desde varias aproximaciones lingüísticas de las cuales destacamos: la aproximación sistémico-funcional, la aproximación lingüístico-textual, la aproximación retórica y el enfoque pragmático y sociocognitivo (López, 2002).

Desde el marco de la lingüística sistémico funcional (LSF), con el fin de determinar criterios organizadores de la caracterización de un texto como de especialidad, se toma en cuenta el principio de la diversidad funcional neofirthiana (Eggins & Martin, 2003) y se

acude a los tres aspectos del contexto situacional o tres variables del registro: Ellas son campo, tenor y modo (Halliday, 1978, 1989; Eggins & Martin, 2003). Por “campo”, se entiende la acción social, lo que está ocurriendo, la naturaleza de la acción social que está sucediendo: qué están haciendo los participantes en lo que el lenguaje figura como un componente esencial. “Tenor”, se considera la estructura de roles: se establece quién participa, la naturaleza de los participantes, su estatus y sus roles. Por último, el “modo” es la organización simbólica: determina cuál es el papel del lenguaje, cómo el lenguaje ayuda en una situación determinada; incluye el canal y también la modalidad retórica, es decir, cómo está organizado el texto en términos de categorías (persuasiva, expositiva, didáctica, etc.)

En el caso del discurso especializado de la actividad científica, las implicancias lingüísticas caracterizadoras globales podrían ser las siguientes:

Campo: corresponde al de la ciencia en el cual existe, fundamentalmente, una orientación hacia la explicitación de métodos y resultados corroborables, que se manifiestan en procesos descriptivos y taxonomías.

Tenor: participan los científicos como productores y consumidores de textos científicos, fundamentalmente demostrativo-persuasivos. Por lo general, no hay interactividad, ya que el discurso es esencialmente monológico y con bajo compromiso afectivo.

Modo: El lenguaje utilizado es abstracto, caracterizado por un estilo nominal, un léxico formal y terminología especializada, con un uso monoreferencial del léxico. Se realiza fundamentalmente por escrito.

Estas características generales son las que luego deberían “materializarse” en los textos de especialidad presentes en el discurso especializado de la ciencia, a través de las distintas estructuraciones retóricas.

Es esta “materialización” del discurso en los textos la que, en definitiva, da unidad a la producción y comunicación de la ciencia, y permite hablar de la existencia de una comunidad discursiva específica, pues se comparten metas comunes (informar y argumentar respecto de un conocimiento investigado), existen mecanismos participatorios,

hay intercambio informacional, se utilizan géneros específicos de comunidad, existe presencia de una terminología altamente especializada y se reconoce un alto nivel de pericia (Swales, 1990).

La aproximación lingüístico-textual aborda la descripción de los rasgos lingüísticos específicos, especialmente léxicos, que caracterizan estos tipos de discurso. En este enfoque se da relevancia al estudio de los registros específicos o tecnolectos, frente al registro general, como variedades del uso de la lengua en contextos profesionales. El estudio de la especificidad de los textos se aborda desde los diferentes niveles lingüísticos: nivel léxico, nivel morfológico y nivel sintáctico (Cabré, 1999; Ciapuscio, 2000). Por otra parte, destacan los estudios multirasgos y multiniveles de grandes corpus de textos especializados, tal como el que se está desarrollando en la Pontificia Universidad Católica de Valparaíso (Parodi & Gramajo, 2003; Parodi, 2004, 2005a, Parodi & Venegas, 2004, Marinkovich & Cademártori, 2004, Sabaj, 2004).

La aproximación retórica tiene en cuenta el proceso específico de producción y comprensión de estos tipos de texto, y no sólo el producto textual. Se busca en estos trabajos detallar cómo se lee y se escribe en el ámbito de los discursos profesionales, qué pasos se han de seguir para lograr los “fines” específicos perseguidos (de ahí el énfasis en la “retórica”) (Gnutzmann & Oldenburg, 1991; Salager-Meyer, 1991; Bolívar, 2000; Ciapuscio & Otañi, 2002; Martín, 2003).

El enfoque pragmático y sociocognitivo privilegia la descripción de los parámetros contextuales como índices caracterizadores de la especificidad lingüístico-discursiva de estos textos. Estos trabajos muestran que los discursos producidos en el ejercicio de una determinada disciplina (por ejemplo, la economía) no pueden analizarse como producciones más o menos homogéneas, con una serie de rasgos lingüístico-discursivos compartidos, sino que debe reconocerse la diversidad discursiva de cada una de las realizaciones específicas, esto es, la heterogeneidad de los géneros discursivos profesionales (Swales 1990; Bhatia, 1993; Gläser, 1993; Berkenkotter & Huckin, 1995; Gunnarsson, 2000; Ciapuscio, 2000; Trosborg, 2000).

1.5 El texto de especialidad en la ciencia

La comunicación científica, entonces, se lleva a cabo por medio del discurso especializado de la ciencia, el cual posee características particulares, y se manifiesta en un conjunto de textos especializados. Estos últimos han sido clasificados de diversas maneras, dependiendo de las funciones que tengan que cumplir; por ejemplo, un texto escrito que comunique el resultado de una investigación de campo no ha de tener la misma organización de contenidos ni el mismo formato, que otro que proponga la solución de un problema particular, específico, localizado temporal y espacialmente (Moyano, 2000).

En el caso particular de los textos científicos, estos están destinados a ser leídos por otros investigadores del ámbito de especialidad que puedan utilizar esa información para el progreso de sus propias actividades y para la construcción del conocimiento científico (recordamos aquí el concepto de *dynamis*). Por lo tanto, la función primera será informar a la comunidad científica de los progresos realizados a través de una investigación. Debemos considerar, sin embargo, que algunos de ellos tienen finalidades subsidiarias: algunos (monografías, tesis y tesinas) son escritos con el fin de acreditar méritos para la obtención de un título de grado, otros, como los informes técnicos o científicos, son destinados a investigadores en su calidad de funcionarios jerárquicos dentro de las instituciones dedicadas a la investigación y cumplen también con la función de acreditar méritos para obtener recursos para un proyecto de investigación, justificar el área en la que están trabajando, etc. En algunos otros casos, como en los ensayos, revisiones, conferencias, debates, los investigadores se propondrán evaluar la información existente y/o asentar posición propia frente a una cuestión teórica (Moyano, 2000).

La comunicación científica se realiza por lo general de forma escrita, pero también a través de textos que han sido escritos para ser leídos o expuestos oralmente en congresos o, textos orales, con mayor o menos grado de planificación. Estas diferencias en el modo de transmitir la información hacen que el emisor elija diferentes recursos para construir su texto de manera de adecuarse al contexto en que se produce la comunicación.

La descripción de los tipos de textos o clases textuales de especialidad, arriba esbozada, se condice plenamente con la definición de género discursivo ampliamente tratado en la

literatura a partir de Bajtin (1982). Desde esta perspectiva, el concepto de género discursivo (o clase de texto) constituye una categoría de análisis fundamental de los textos desde distintas tradiciones de investigación del discurso y de los textos (Bajtín, 1982; Halliday, 1978, 1989; Swales, 1990; Martin, 1997; Martin & Veel, 1998; Martin & Rose, 2003; Ciapuscio, 2000, 2003, 2005; López, 2002). Así, los géneros son concebidos como formas discursivas estereotipadas definidas principalmente por sus características externas, propias del sistema sociocultural, plasmadas en el sistema sociodiscursivo. Estos géneros son reconocidos por los hablantes según los parámetros situacionales en que se suelen producir (sistema contextual sociodiscursivo), muy ligados a las prácticas sociodiscursivas de una comunidad (actividad sociodiscursiva) y por su formato externo, puesto que son materializados en el texto, por medio de los sistemas léxico, semántico, gramatical y retórico estructural.

Una muy útil clasificación de los géneros científicos es la presentada por Moyano (2000), ya que congruentemente con lo anterior, considera los criterios de finalidad, destinatario, contenido y circulación, para dar cuenta de ellos. Por razones de focalización no es posible dar cuenta *in extenso* de cada uno de los géneros, es por ello que a continuación se presenta la matriz que sintetiza la clasificación realizada por la autora.

Tabla 1: Clasificación de géneros científicos (Moyano, 2000: 31)

Finalidad	Destinatarios	Géneros	Contenido	Circulación
Informar	Comunidad Científica	Mural	- Resultados de un trabajo de investigación	Exposición Gráfica Discusión Oral
		Artículo Científico		
		Comunicación Preliminar		
		Comunicación	Técnicas o métodos de investigación - Estudios de caso	
Informar y acreditar méritos	Docentes de grado y postgrado universitario	Abstract	- Resumen de un texto científico	ESCRITA
		Tesis	- Resultados de un trabajo de investigación	
	Monografía	- Investigación documental		
Funcionarios	Informe Científico	Estado de avance de una investigación Descripción de un problema - Propuestas de solución a un problema		
Informar Evaluar Sentar posición	Comunidad Científica	Revisión Bibliográfica	Estados de cuestión - Sugerencias de líneas de investigación	
		Ensayo	Presentación de teorías - Discusión teórica de un problema	
		Ponencia	Formulaciones teóricas Experiencias o estudios de casos - Nueva técnica o método	Exposición oral Escrita para publicar
		Conferencia	Novedades y últimas especulaciones - Sugerencias para investigaciones futuras	
Confrontar		Debate	Aportes a la solución de un problema - Posiciones frente a un tema o problema	ORAL

De la Tabla 1 nos interesa profundizar en la clase textual Artículo Científico, el cual será analizado en el apartado siguiente.

1.5.1 El artículo de investigación científica

El artículo científico, o más específicamente el artículo de investigación científica, es un texto escrito, generalmente publicado en una revista especializada, que tiene como finalidad informar a la comunidad científica los resultados de un trabajo de investigación realizado

mediante la aplicación del método científico, según las características de cada disciplina de la ciencia. Su estructura es bastante rígida (al menos como se presenta en algunas disciplinas empíricas) y expone en general, en el orden que enunciamos, los siguientes apartados: Introducción, Materiales y Métodos, Resultados, Discusión y Conclusiones (Swales, 1990). Estas secciones están precedidas por un título y la mención de los autores y de las instituciones en las que ellos se desempeñan como investigadores, así como por un resumen, destinado a informar sucintamente a los lectores acerca del contenido de todo el artículo para que ellos decidan si les resulta útil la lectura completa del texto (Moyano, 2000).

Se debe hacer notar que el artículo de investigación científica es el producto de un complejo proceso en el cual un manuscrito se va gestando lentamente a través de múltiples borradores, que se construyen a partir de la interacción del autor con sus coautores, colaboradores, colegas, evaluadores y editores (Swales, 2004).

Sin embargo, antes de avanzar con la descripción y análisis más acabado del artículo de investigación científica, es necesario plantear aquí una breve reflexión respecto del propósito que esta clase textual cumple, pues nos parece que no hay unifuncionalidad.

Los resúmenes y las introducciones tienen diferentes propósitos comunicacionales en la escena científica. Desde esta perspectiva, si bien resumen e introducción forman parte de la estructura del artículo de investigación científica son géneros diferentes dentro de un género de investigación común. Siguiendo en esta línea de argumentación, podemos plantear que los elementos de la estructura IMRADC (Introducción, Materiales y Métodos, Resultados, Discusión, Conclusión) en los artículos de investigación científica tienen diferentes propósitos comunicacionales, por lo que ellos pueden ser considerados como géneros distintos dentro de un mismo “Macro-Género” (Swales, 1990). En concordancia con estas ideas, parece necesario introducir la noción de “Macro-Género” y la de “Género-Incrustado” (Warta, 1996).

“Macro-Género” es entendido como un género compuesto de otros géneros. Por ejemplo, el “Macro-Género” denominado artículo de investigación. En tanto que “Género-Incrustado” corresponde a algún género perteneciente al marco del “Macro-Género”. Por ejemplo, el

resumen (concebido como “Género-Incrustado”) en un artículo de investigación (Warta, 1996).

De esta forma, el “Macro-Género” artículo de investigación científica tiene como función, exponer en forma completa un trabajo de investigación realizado mediante la aplicación del método científico, según la modalidad de cada disciplina de la ciencia, con el propósito de contribuir al progreso de los saberes relativos a un tema de interés disciplinar, permitiendo que los lectores especialistas hagan uso de la información contenida en él para construir nuevos conocimientos. Complementando lo anterior, el resumen como “Género-Incrustado” cumpliría la función de compactar, o globalizar, la información contenida en el artículo, con el propósito de orientar al lector en los aspectos fundamentales de la investigación que se presenta. Como más adelante discutiremos, cumple la función de macrosemantizar el significado del artículo.

En síntesis, y retomando nuestra definición de texto planteada más arriba, podemos establecer que el artículo de investigación científica, en tanto producto lingüístico de las representaciones psicosocidiscursivas de uno o varios sujetos productores, corresponde a un tipo de texto especializado (dada la materialización en sus sistemas LSGR de los sistemas sociodiscursivos y socioculturales en los cuales se desempeñan los científicos) que se publica en las revistas de especialidad de una disciplina determinada, en el cual se da cuenta de una investigación particular (en cuanto sistema contextual sociodiscursivo). Como *actividad psicosocidiscursiva* es una actividad especializada científica orientada a la producción y transmisión de conocimientos, y como *actividad en potencia* es un conocimiento puesto a disposición para otros investigadores con el fin de acrecentar el conocimiento sobre un tópico particular de la ciencia.

1.5.1.1 Estructura del artículo de investigación científica

En cuanto a su sistema retórico estructural, en general, se espera que el artículo de investigación científica esté organizado de tal manera que ponga de manifiesto el procedimiento por el cual se haya accedido al conocimiento expresado en las conclusiones: que el autor señale por qué y para qué se realizó el estudio en cuestión, cómo se desarrolló

la investigación y cuál es el aporte concreto que permite ofrecer, además de indicar cómo se relaciona con otros estudios hechos en el mismo campo.

La denominación de estructura planteada aquí es deudora de lo que van Dijk (1983) identificó como superestructura o estructura global de un texto, es decir, el esquema convencional que determina el orden global de sus partes. El respeto por estas convenciones formales aumenta la claridad de un texto y lo vuelve fácilmente comprensible. Si el lector conoce la estructura característica de un género, sabrá qué tipo de contenido y de estilo puede encontrar en cada parte y estará en condiciones de reconocer con mayor facilidad las relaciones entre ellas, porque sabrá que existen y cuáles son.

Para el artículo de investigación científica los acuerdos resultan, en muchos casos, tácitos (Moyano, 2000). En general, los investigadores redactan sobre la base de modelos reales extraídos de la revista donde pretenden publicar, que no siempre resultan modelos adecuados o, al menos, consensuados por los editores o siguiendo las instrucciones de colegas más experimentados en la tarea de escribir. Esto significa un verdadero estorbo para la actividad discursiva científica, puesto que el tiempo con que se cuenta es escaso y los materiales de lectura son abundantes. Los textos difíciles de leer, aquellos donde no hay una estructuración lógica clara, suelen ser desechados por los investigadores. Esto sin contar que si algún investigador insiste en la lectura, lo hará desde sus conocimientos previos sobre el tema, sus propios esquemas y expectativas, de modo que puede alejarse completamente de la intención significativa del escritor, con la gravedad que este hecho puede tener en este tipo particular de comunicación.

Swales (1990, 2004) sostiene que hay características que se repiten suficientemente en los artículos científicos de un extenso rango de disciplinas como para considerar la existencia de un Macro-Género.

De todos modos, los artículos varían de una disciplina a otra en grados de estandarización y estilo: las ciencias conocidas como “duras”, “exactas” o “físicas” siguen un modelo más rígido, mientras que en las ciencias sociales existen grupos que han intentado adaptarse a ese modelo con diferentes grados de éxito, mientras que otros se resisten a establecer reglas fijas para sus textos (Moyano, 2000; Mogollón, 2003, Swales, 2004).

En este sentido, la investigación a desarrollar tomará en cuenta los artículos que pertenezcan a un paradigma de investigación más cercano a una postura positivista, en donde la estandarización retórico-estructural aporte la mayor homogeneidad posible a la selección de los artículos de investigación. Esto último, reflejado en la estructura retórica comúnmente propuesta para los artículos de investigación científica. Estas son las siguientes: Título, Identificación de autores y dirección, Resumen, Introducción, Materiales y Métodos, Resultados, Discusión, Conclusiones, Agradecimientos, Referencias bibliográficas (Swales, 1990; Moyano, 2000; Martín, 2003; Maruhenda, 2003).

En su aspecto más particular, se distingue la estructura retórica IMRDC, siguiendo a Moyano (2000), la cual permite reconstruir los pasos seguidos en la investigación. En primer lugar, en la “Introducción” el investigador describirá el problema o fenómeno que lo llevó a formularse preguntas en relación con sus conocimientos sobre el tema, comunicar los aportes hechos por otros autores a su resolución y planteará los objetivos que se propuso al realizar el trabajo de investigación. Luego en el apartado “Materiales” y “Métodos” (también denominado “Metodología” en algunos casos) consignará todos los elementos utilizados, describirá los procedimientos seguidos, la metodología y las técnicas aplicadas. En “Resultados”, dará cuentas de los datos obtenidos. Más adelante, en “Discusión” expondrá su análisis y su relación con lo hallado por otros autores, las inferencias y deducciones lógicas pertinentes, así como también se adelantará a todas las objeciones que crea pueden hacerse a sus afirmaciones de manera de neutralizarlas. Finalmente, en “Conclusiones” expondrá la respuesta a los objetivos propuestos. Éstas deberán estar sustentadas por los resultados y confirmarán o rechazarán las hipótesis formuladas.

Cabe destacar que Moyano (2000) aboga por la inclusión de las conclusiones en la estructura, llenando el posible vacío que deja la estructura IMRD más difundida por los autores de habla inglesa (ver Swales, 1990). Sin embargo, la estructura IMRD, según Swales (2004), es una estructura más bien flexible, en donde la etiqueta “Discusión” corresponde a todas las posibilidades que tiene un autor de artículos de investigación científica para cerrar su manuscrito (por ejemplo: Discusión, Conclusiones, A modo de conclusión, Palabras finales, Comentarios finales, etc.)

Cabe señalar que esta estructura, si bien es más flexible, aún está fuertemente influenciada por los artículos escritos en ciencias exactas y en inglés. En el ámbito de los artículos escritos en español de otras ciencias (sociales, biológicas, humanas, etc) estas estructuras varían, por ejemplo, el apartado “Discusión” puede alternar con el de “Análisis de los Resultados”, o por una reducción de los últimos pasos retóricos, por ejemplo “Resultados y Discusión” incluyendo en este paso la “Conclusión”.

En definitiva, sabiendo que existen otras posibilidades de estructuración retórica, parece ser que la estructura IMRDC satisface la condición de lograr textos homogéneos para comunicar ciencia. Se hace notar que, en este sentido, las revistas científicas buscan, cada vez más, estandarizar la producción de los artículos con el fin de incorporarse o mantenerse en los indexadores internacionales de artículos científicos. Por último, esta estandarización, en alguna medida, facilitaría tanto la redacción como la lectura del artículo, permitiendo la replicación de la investigación. Además, esta estructura es lo suficientemente flexible para ser adaptada a las exigencias del contenido que se desee transmitir.

Diversos estudios se han llevado a cabo en función de cada una de las partes del artículo de investigación científica. Por ejemplo, la introducción ha sido profundamente trabajada por Swales (1990), las introducciones y conclusiones por Gnutzmann y Oldenburg (1991), las conclusiones por Ciapuscio y Otañi (2002), el resumen por Salager-Meyer (1991) y Bolívar (2000), el resumen y las introducciones por Martín (2003), las secciones introducción y discusión por Dudley-Evans (1986). Todas estas investigaciones han sido llevadas a cabo fundamentalmente desde perspectivas lingüístico-textuales, retóricas y sociocognitivas y, la mayoría de ellas, desde un enfoque comparado interlenguas en muestras ejemplares de textos.

En lo que sigue nos referiremos, con algo más de detalle, a tres aspectos del artículo de investigación científica, estos son: el resumen, las palabras clave y la introducción. Nuestro interés particular por estos aspectos está orientado por las relaciones léxico-semánticas que estos presentan entre sí y con el contenido semántico total del artículo. En este sentido, se puede plantear, como se argumentará más adelante, que el contenido del artículo, el resumen y las palabras clave son niveles de semantización de abstracción creciente. Esto

quiere decir, que el contenido del artículo es macrosemantizado en el resumen y que las palabras clave macrosemantizan, a su vez, al resumen y al contenido del artículo.

1.5.2 El resumen en el artículo de investigación científica

El resumen es un texto breve que sirve para que el lector identifique de forma rápida y precisa el contenido básico del artículo. Por lo general, se ubica entre el título y la introducción. En él, el autor debe reflejar de forma fiel el contenido del artículo, con lo cual conviene redactarlo al finalizar la elaboración del documento. Hay que destacar que el resumen representa en la actualidad un papel muy importante, pues existen sistemas de recuperación de información bibliográfica que al realizar una búsqueda de artículos, además del título y los autores, muestran también el resumen (Maruhenda, 2003).

El resumen tiene la misma estructura lógica que el artículo, aunque no lleva subtítulos y carece de discusión, así como de citas bibliográficas, cuadros, tablas o figuras. En él, en una extensión no superior a las 250-300 palabras, se tiene que poner en evidencia el rigor científico del trabajo de investigación, por lo que expone brevemente los objetivos, la información imprescindible acerca de los materiales y los métodos de investigación utilizados, las conclusiones –en valores numéricos cuando se trate de información cuantitativa–, que las sustenten. Todo en el orden lógico IMRC (Moyano, 2000).

Cabe hacer notar aquí que se mantiene la diferencia retórico estructural entre la propuesta de Swales (1990) y Moyano (2000) en cuanto a que la autora recomienda la construcción del resumen siguiendo la misma estructura lógica del artículo, sin incluir el apartado *discusión* (Moyano, 2000). En tanto Swales (1990) sugiere que el resumen comúnmente sigue el patrón IMRD en donde se plantea una introducción general, seguida por la declaración del método, de los resultados y se plantea la discusión. Nos parece, de modo muy particular, que la postura de Moyano (2000) es más acertada y más cercana a nuestra experiencia como lectores de artículos de investigación científica en español.

Existe consenso en que el resumen no debe resultar una mera transposición de fragmentos del texto original, sino una reelaboración de lo que éste propone globalmente. El resumen, desde el punto de vista semántico, corresponde a una globalización (condensación de la

información en unidades menores) y a una conceptualización de la red de contenidos del texto. En este sentido, proponemos denominar a este proceso de globalización como “macrosemantización”, ya que el resumen es la textualización de un significado que representa de modo abstracto el significado total del contenido del artículo.

Dentro del marco de la informática, se define al resumen como la indicación concisa del contenido temático de un documento, cuya realización excluye por regla general cualquier valoración crítica y cuyo propósito es facilitar la determinación de la relevancia del texto original para un eventual lector (López, 1997).

Según van Dijk y Kintsch (1983), el hecho de que el resumen se encuentre al inicio de la lectura ayuda al lector a formarse una hipótesis sobre el tópico del discurso o del episodio, de tal manera que las oraciones siguientes pueden interpretarse de manera arriba-abajo para tales macroproposiciones. Cuando en su confección se sigue este principio, de arriba-abajo, el producto que se obtiene resulta en ocasiones hasta más claro y coherente que el propio trabajo sometido al proceso de análisis y síntesis, cuya información fundamental se mantiene sin que se haya alterado su contenido. Tal afirmación se explica por sí misma, pues el resumen como producto no es más que el resultado de una abstracción, en la que se sintetiza la información que ofrece el documento de origen manteniendo sus partes esenciales (López, 1997).

Algunos autores se han dedicado a calcular el porcentaje de condensación adecuado en la correlación resumen-trabajo original, y han llegado a la conclusión de que el factor de reducción debe guardar una proporción que oscile en el rango de 1:10 a 1:12 (Borko & Bernier, 1975 citado en López, 1997).

Otros expertos (Gnutzmann & Oldenburg, 1991; Salager-Meyer, 1991; Bolívar, 2000; Martín, 2003) han demostrado que el contenido de un documento se analiza a partir de un nivel microestructural, formado por las estructuras lineales del texto que se expresan de un modo directo; y de un nivel macroestructural, basado en el texto como un todo. Asimismo, se sostiene que las superestructuras son formas convencionales y esquemáticas de las macroestructuras, que se identifican con el esquema utilizado en la confección de trabajos científicos y que deben estar presentes en la preparación de sus respectivos resúmenes.

Una forma de estudio más actualizada, aunque menos conocida, es la que permite estudiar el resumen desde una perspectiva más computacional. En este sentido destacan los trabajos realizados por los investigadores del Instituto de Ciencias Cognitivas de la Universidad de Colorado, Boulder, Estados Unidos. Ellos han desarrollado estudios teóricos y empíricos orientados fundamentalmente a probar que los modelos vectoriales, en particular el Análisis Semántico Latente (LSA, por su sigla en inglés)¹, puede dar cuenta del procesamiento de los significados que realizan los seres humanos. El LSA es una técnica matemático-estadística que sirve para la extracción y representación de relaciones de significado entre palabras y párrafos, lo que se realiza a partir de una gran cantidad de textos. Esta técnica utiliza la Descomposición en Valores Singulares (SVD, por su sigla en inglés), la cual es una forma de análisis factorial para condensar una gran matriz de frecuencias de palabras por contexto en una mucho más pequeña, pero todavía grande, en la cual se consideran entre 100 y 500 dimensiones representativas de la matriz original (Deerwester, Dumais, Furnas, Landauer & Harshman, 1990). La selección de un número correcto de dimensiones parece ser crucial, ya que los mejores valores (normalmente 400 dimensiones) producen una simulación hasta cuatro veces más confiable de los juicios humanos que las medidas de co-ocurrencia.

Las similitudes entre los vectores resultantes para palabras y contextos, así como las mediciones por coseno de sus ángulos, han mostrado para el inglés, de varias maneras, ser muy parecidas a los juicios de similitud semántica realizados por humanos y a las actuaciones humanas basadas en tales similitudes. En cuanto al resumen, con LSA se han desarrollado trabajos que permiten evaluar la calidad de los resúmenes de los estudiantes de sexto grado y proporcionar retroalimentación respecto del proceso de escritura. Para ello se realiza la comparación semántica de los textos proporcionados por los alumnos y los textos originales (Kintsch, Steinhart, Stahl, LSA Research Group, Matthews & Lamb, 2000). Para estos autores, el proceso de síntesis (*summarization*) es concebido como una estrategia que puede ser usada para determinar cuánto comprende o no comprende un estudiante a partir

¹ Un análisis en profundidad del Análisis Semántico Latente se realiza en el Capítulo 2 de esta investigación.

de lo que lee y cuánto aprende o no del texto leído. El desarrollo de un sistema de evaluación y retroalimentación basado en LSA no sólo revela la existencia de quiebres en la comprensión durante la lectura, sino que ayuda a identificar las causas y ubicaciones de estos quiebres comprensivos. Escribir un resumen, como se sabe, involucra construcción de significados, aun cuando se focalice en la base textual. En un mayor grado de abstracción, la toma de notas y las esquematizaciones requieren de la construcción de una representación mental que permita unir los elementos informativos de los textos con elementos del conocimiento previo (Kintsch et al., 2000).

En el caso de los estudiantes de sexto grado, ellos adquieren nuevo conocimiento que les será útil y de larga duración, ya que construyen representaciones del conocimiento bien elaboradas, lo que les permite una adecuada aplicación de este conocimiento, así como una buena recuperación. Finalmente, los resúmenes son una herramienta comunicativa, un medio que permite compartir el conocimiento de unos con los otros (Kintsch et al., 2000).

En virtud de las anteriores consideraciones, viene al caso precisar que el resumen ha de cumplir no sólo la función de proporcionar elementos que estimulen la consulta del documento original, sino, más aún, debe facilitar la obtención de un primer nivel de asimilación del problema que se aborda y propiciar un precedente informativo sólido. En términos más cognitivos, el resumen debe entregar al lector, en este caso un lector especializado, los elementos necesarios para construir una representación del modelo de situación, que le permita comprender los procedimientos básicos y resultados que se consignarán en el artículo.

Esto exige del autor ser capaz de textualizar macroproposiciones, por ejemplo, las generadas a partir de la investigación (en ocasiones pretextualizadas en notas e informes de investigación) en el artículo de investigación, en un proceso que involucra la aplicación estratégica de las macrorreglas que han sido definidas en la teoría del discurso (van Dijk & Kintsch, 1983). Estas macrorreglas son las siguientes:

- 1.- Eliminación: dada una secuencia de proposiciones, se elimina toda proposición que no sea una condición de interpretación para otra proposición en la secuencia (por ej. una presuposición)

2.- Generalización: Dada una secuencia de proposiciones, se sustituye la secuencia por una proposición que contiene cada una de las proposiciones de la secuencia.

3.- Construcción: Dada una secuencia de proposiciones, se reemplaza por una proposición que está contenida por un grupo de proposiciones de la secuencia.

La aplicación de estas reglas en el proceso de construcción del resumen es lo que nosotros hemos venido denominado macrosemantización, entendido como el proceso de compactación de significado en niveles de textualización cada vez más abstractos y que son llevadas a cabo, muy probablemente, a través del uso estratégico de las macrorreglas presentadas arriba.

El plantear el uso de las macrorreglas para el procesamiento cognitivo del significado, no está exento de problemas. Por una parte, porque cabría suponer que toda persona que siga las macrorreglas estará en condiciones de obtener un resultado, si no igual, muy semejante a otra, lo cual deja sin considerar aspectos que sabemos son importantes en la construcción de significado, como por ejemplo: Las estrategias de interpretación, las diferencias de conocimiento, las creencias, las opiniones, las actitudes, los intereses o metas particulares, etc.

Al respecto, van Dijk y Kintsch (1983) plantean, que desde una perspectiva metodológica, un modelo cognitivo no especificará cómo los individuos comprenden el discurso, pero sí intentará formular los principios generales. De igual modo, una comunicación verbal adecuada es posible sólo si los usuarios de lenguaje tienen significados y conocimiento en común. Por consiguiente, un modelo cognitivo adecuado de macroestructuras debe especificar los principios generales seguidos por los usuarios en la comprensión global del significado del discurso, y mostrar cómo las diferencias individuales presuponen suficiente información común para establecer una comunicación adecuada. Al respecto afirman que:

“Whatever the individual cognitive make-up of language user, he or she can not assign completely arbitrary or disparate macrostructure to a discourse. The speaker`s intentions as expressed by the text or reasonably adscribed to the speaker by the hearer on the grounds will play a normative role in the derivation of macrostructures” (van Dijk & Kintsch, 1983: 193).

Si bien la discusión en cuanto a esta problemática no está zanjada, asumimos que a un nivel semántico lingüístico la utilización de estas macrorreglas permitirá a distintos individuos construir un resumen, como resultado de un proceso de macrosemantización, de manera relativamente similar en términos de su significado global, sobre todo en un discurso especializado donde los miembros de la comunidad discursiva utilizarían estrategias de interpretación relativamente comunes, donde el conocimiento requerido para la integración del significado global es más bien compartido y donde las creencias, las opiniones, las actitudes y los intereses responden al mayor o menor apego a los paradigmas científicos conocidos.

Un paralelo a este procedimiento es la estandarización de normas que establecen el estilo y contenido de los resúmenes según las orientaciones de la American National Standard for Writing Abstracts (ver López, 1997). Con estas reglas, lo que se pretende no es tanto realizar la operación de extraer información del texto original cuanto de condensar lo más posible, manteniendo las características esenciales del texto, sin que se distorsionen los puntos de referencia cognitivos que favorecen la coherencia textual originaria, como inferencias, implicaciones y presuposiciones.

1.5.2.1 Clasificación de los resúmenes

En cuanto al estudio de una tipología del resumen, Ratteray (1985) propone la existencia de al menos siete clases de resúmenes. Estos han sido clasificados en resúmenes secuenciales, es decir, aquellos que mantiene el orden original de la información presentada y en resúmenes que alteran esta secuencia para lograr objetivos específicos, estos son conocidos como resúmenes sintéticos.

Entre los primeros se encuentran: El resumen de artículo científico, el Précis, la minuta y el compendio. Entre los segundos se encuentran: El compendio locacional, el compendio reestructurado y la reseña.

Las diferencias entre estos tipos de resúmenes estriban en el modo más o menos extractado en que indican el contenido del documento original, en la forma implícita o explícita en que lo reflejan y en el objetivo que cumplen (López, 1997; Moyano, 2000; Maruhenda, 2003).

Cada uno de los tipos de resúmenes mencionados cuenta con subclasificaciones, sin embargo, aquí sólo daremos cuenta del resumen del artículo científico, perteneciente a la clase de los resúmenes secuenciales:

a) El resumen indicativo o descriptivo ofrece los enunciados fundamentales del trabajo original. Con él se puede conocer de qué trata el documento a que hace referencia, pero no adquirir conocimientos en cuanto a resultados concretos de las reflexiones implícitas o de los estudios expuestos en éste. Se considera que este tipo de resumen tiene buena acogida en las comunicaciones a conferencias y en los informes técnicos publicados o inéditos (Ratteray, 1985; López, 1997; Moyano, 2000).

b) El resumen informativo contiene de manera generalizada todos los enunciados importantes del documento primario, incluyendo el material que lo ilustra, los datos relativos a los métodos de investigación, los porcentajes, el empleo de equipos y la esfera de aplicación. Se dice que la información que brinda es proporcional al tipo y al estilo del documento base y que su extensión puede oscilar entre las 200 y las 250 palabras. El resumen informativo es muy apropiado para los artículos originales, dada su peculiaridad de sintetizar cada uno de los principales aspectos tratados en sus textos (Ratteray, 1985; López, 1997; Moyano, 2000).

c) El resumen indicativo-informativo es una forma de síntesis que mezcla las dos modalidades anteriores, con las que el lector puede obtener una información más exhaustiva sobre determinados aspectos tratados en el texto original, mientras que de otros sólo se hace mención. En otras palabras, este tipo de resumen, también apropiado para los artículos originales, presenta en forma condensada los principales argumentos, datos y conclusiones, a la vez que indica brevemente el alcance y contenido del documento (López, 1997).

d) El resumen estructurado, muy utilizado últimamente por las revistas médicas de mayor prestigio internacional, adquiere cada vez mayor importancia, sobre todo cuando el documento original que sintetiza aborda asuntos de interés clínico, ya que permite juzgar con más facilidad la validez y posibilidad de aplicación inmediata de los hallazgos en él registrados (Salager-Meyer, 1991; López, 1997). A diferencia de las modalidades de

resumen descritas anteriormente, esta presenta de modo explícito los componentes que integran las distintas partes del trabajo que le sirve de base, en virtud de lo cual proporciona una información mucho más exhaustiva. Aunque tiene detractores debido a su limitación potencial en el estilo y la creatividad que impone a los autores o a la disminución de la legibilidad que provocan, se ha demostrado sus ventajas en cuanto a que es más informativo y a que su formato normalizado ayuda a los lectores a seleccionar los artículos apropiados con mayor rapidez, permite que las búsquedas automatizadas sean más precisas y facilita el arbitraje de los artículos que aspiran a ser publicados (López, 1997).

Por último, entre los resúmenes sintéticos cabe destacar, para el ámbito de la ciencia, la anotación y la reseña:

a) La anotación, cuya característica fundamental es precisamente su brevedad, además de que en ella casi nunca aparece siquiera una palabra de la presente en el título del trabajo sintetizado. Su función se reduce a ofrecer el tema central o el objeto del documento primario sin entrar en otros detalles, por lo mismo raramente excede los 200 caracteres. Por regla general, esta forma de síntesis se utiliza en los índices analíticos de las revistas científicas (López, 1997).

b) La reseña es básicamente una presentación de la opinión que se formula sobre otra pieza escrita; justificada con hechos e incidentes específicos del trabajo en cuestión. Su propósito principal es dejar saber si vale la pena leer la obra que está en discusión. Independientemente si el lector ha leído la obra, el autor de la reseña debe incluir alguna información sobre el contenido, sin olvidar nunca que el objetivo de la reseña es una presentación de la opinión de quién la lleva a cabo (Ratteray, 1985).

Como hemos visto hasta aquí, la función del resumen, sobre todo la del informativo, es de vital importancia para acceder al contenido global del artículo de investigación científica, dado que sintetiza cada uno de los aspectos principales tratados en el contenido del artículo. Es por ello que se convierte en una vertiente central de nuestra atención en esta investigación.

1.5.3 Palabras clave

Muchas revistas piden a los autores que agreguen, luego del resumen, un listado de palabras clave a su artículo. Sin embargo, seleccionar tales palabras clave no es tarea fácil, pues ellas deben condensar de una manera aún más abstracta que el resumen los significados más relevantes del artículo. Al respecto, Hartley y Kostoff (2003) indican que estas, junto al resumen y al título, constituyen rasgos que describen el contenido en grados variables de detalle y de abstracción.

Existen artículos en los cuales es notoria la dificultad de abstraer el contenido en apenas cinco o seis palabras. Esta dificultad a veces se resuelve a través de la inclusión de frases de dos o más palabras, generalmente nominales. Es por ello que hay autores que han optado por llamarlas frases clave (Turney, 1997, 1999).

Las palabras o frases clave pueden cumplir principalmente dos objetivos: por una parte, permiten resumir el contenido de un artículo de investigación. En este caso, las palabras clave son una forma altamente abstracta de resumir los significados más relevantes del contenido de un documento. Ellas le permiten al lector determinar rápidamente si el artículo dado está en su campo de interés. Y, por otra, permiten indexar artículos según una temática determinada. Ellas facilitan una búsqueda rápida de un artículo relevante para el lector, cuando hay una necesidad específica (Turney, 1999).

Para Hartley y Kostoff (2003), las palabras clave indican los conceptos principales y delimitan el campo de interés de la investigación. De esta manera, las palabras clave se caracterizan por permitir al lector decidir cuando un artículo contiene o no material relevante para su interés, estas proveen a los lectores de un grupo conveniente de términos para ser usados en búsquedas en la internet para localizar otros materiales en el tópico, ayudan a los editores/indizadores a agrupar materiales relacionados, permiten a los editores/investigadores intercambiar documentos en los temas de una disciplina y vinculan tópicos específicos de preocupación con tópicos en metaniveles más altos. Esto último, referido a las listas de palabras clave que pueden identificar un área de investigación en particular. Por ejemplo, si se hace una lista de las palabras clave de los artículos de investigación científica de una disciplina particular, las más altas frecuencias indicarían

cuáles son los tópicos más comúnmente tratados en esa disciplina, lo que indicarían niveles de abstracción más altos y darían cuenta de los metatópicos del área (Hartley & Kostoff, 2003).

Por lo general, el estudio que se hace de las palabras clave está en relación con los estudios que tienen por interés la extracción de información desde los textos, a través de herramientas computacionales que funcionan, principalmente, con métodos matemático-estadísticos y con modelos algorítmicos. Estos estudios se basan en el principio de obtener documentos relevantes ante la necesidad de un usuario de conseguir cierta información. Por ejemplo, en el caso de un artículo científico que debe ser indexado y no posee palabras clave o resumen. Los distintos métodos o modelos permitirán, a través de distintas técnicas, entregar, bien un listado de las palabras más relevantes del texto o bien un texto que sintetice la información del documento de origen (ver Turney, 1997, 1999).

Entre los estudios que se hacen con los métodos matemáticos-estadísticos, lo que se realiza es la identificación de palabras clave por medio de la comparación de patrones de frecuencia de aparición de una palabra en dos *corpora*. Para estos estudios, las palabras identificadas como clave no son necesariamente las más frecuentes en el corpus de estudio, sino aquellas que poseen una frecuencia significativa al compararlas con las del corpus de referencia. Existen varios métodos para determinar la relevancia estadística en la diferencia de frecuencia de aparición de un elemento léxico, de los cuales el chi-cuadrado (o prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas) y el *Log-Likelihood* (o logaritmo de la probabilidad de los resultados observados dadas las estimaciones de los parámetros) son los más usados (Lebart, Salem, & Bécue, 2000; Pérez, 2002).

Las palabras clave que seleccionan estos procedimientos de comparación de ambos *corpora* sirven para dar una buena indicación de la temática de los textos. Estos métodos han demostrado ser útiles para delimitar áreas conceptuales importantes en el ámbito de especialidad o metatópicos (Pérez, 2002; Hartley & Kostoff, 2003) y, v por lo tanto, pueden ser muy útiles para los estudios terminológicos en distintas áreas de especialidad.

Por otra parte, los modelos algorítmicos de extracción de información están basados en una interpretación “extrema” de los principios de la semántica composicional (Jurafsky & Martin, 2000). En estos sistemas, el significado de los documentos reside solamente en las palabras que están contenidas en ellos. Es decir, el orden y la constitución de las palabras que forman las oraciones, que a su vez constituyen el texto, no tienen importancia en la determinación de su significado. Debido a que ignoran la información sintáctica, este tipo de aproximación es comúnmente denominada como métodos de “depósito de palabras” (*bag of words*) (Jurafsky & Martin, 2000).

De modo muy general, es posible reconocer dos grandes modelos para realizar la tarea de recuperación, estos son: los modelos clásicos, también llamados variantes *ad hoc*, y los modelos estructurados o de *filtrado de documentos*. A los modelos clásicos pertenecen tres submodelos importantes que son la base de muchos otros: modelo booleano, vectorial y probabilístico. Y entre los modelos estructurados se encuentran las listas no coincidentes y el modelo de nodos próximos (Maldonado, 2002).

La diferencia fundamental entre estos modelos es que el modelo clásico trabaja con una colección de documentos estática, es decir, una base de datos cerrada (por ejemplo: en nuestra investigación usaremos un número fijo de artículos científicos, no se irán actualizando). A esta colección o base de datos se le pueden hacer consultas, por ejemplo, respecto de un artículo en particular. Esta consulta activará una orden que buscará aquellos documentos (entendidos como textos completos, párrafos o palabras únicas) que parecen ser relevantes dentro de dichas colecciones y los jerarquizará de acuerdo al grado de relevancia que tienen según la consulta realizada. Por el contrario, en el modelo estructurado, las consultas permanecen estáticas y es la colección de documentos o base de dato la que cambia. Esto significa que al ingresar información nueva a la base de datos, si esta información es relevante según el proceso de filtrado para alguna consulta realizada previamente por un usuario, la respuesta a la consulta llegará al sistema (computador) vía electrónica, sin necesidad de hacer una nueva consulta. Para que el proceso de filtrado de documentos funcione, es necesario tener almacenadas las características y preferencias del usuario, para que estas se comparen con los documentos y así se tome la decisión de qué le será más interesante al usuario (Maldonado, 2002).

En nuestro caso, nos interesará dentro de los modelos clásicos, el modelo vectorial, y dentro este último, el modelo de indexación semántica latente o, como es conocido en psicolingüística computacional, Análisis Semántico Latente. Esto fundamentalmente, porque trabajaremos con una colección de documentos estáticos y porque nos interesa realizar búsquedas que nos permitan cuantificar las relaciones semánticas existentes entre el significado de una palabra clave y el significado del documento. Otro motivo es que estos modelos no requieren un marcaje estructural de ningún tipo.

En el capítulo 2, en el apartado 2.7, referente a las representaciones computacionales del significado, explicaremos en detalle el funcionamiento de este modelo de análisis semántico, dando cuenta de manera amplia y crítica de las fortalezas y las debilidades del modelo.

1.5.4 La introducción

Como ya mencionamos más arriba, la introducción del artículo de investigación científica tiene por propósito describir el problema o fenómeno que llevó al investigador a formularse preguntas en relación con sus conocimientos sobre el tema, comunicar los aportes hechos por otros autores a su resolución y plantear los objetivos que se propuso mediante la investigación.

Un estudio, ya clásico, respecto de la introducción del artículo de investigación científica en inglés es el realizado por Swales (1990) y levemente modificado en Swales (2004). Este autor, desde la teoría del análisis del género, propone un modelo denominado CARS (del inglés *Create a Research Space*), el cual se compone de tres categorías generales denominadas “Movidas” y subcategorías, a veces optativas entre sí, denominadas “Pasos”. Para Swales (1990, 2004) la definición de “Movida” es la siguiente: “A ‘move’ in genre analysis is a discursal or rhetorical unit that performs a coherent communicative function in a written or spoken discourse” (Swales, 2004: 229).

A continuación, en el Esquema 4 presentamos el modelo clásico de introducción para artículos de investigación en inglés propuesto por Swales (1990: 141).

Esquema 4. Versión clásica del Modelo CARS de Swales (1990) propuesto para la introducción.

Move 1: Establishing a territory	
Step 1	Claiming centrality and/or
Step 2	Making topic generalizations and/or
Step 3	Reviewing items of previous research
Move 2: Establishing a niche	
Step 1a	Counter-claiming or
Step 1b	Indicating a gap or
Step 1c	Question-raising or
Step 1d	Continuing a tradition
Move 3: Occupying the niche	
Step 1a	Outlining purposes or
Step 1b	Announcing present research
Step 2	Announcing principle findings
Step 3	Indicating Research article structure

Según Swales (1990), la “Movida 1” debe proveer los antecedentes que se conocen respecto de los tópicos principales o problemas que serán discutidos en el artículo. Dado que el texto se orienta de lo general a lo específico, el investigador-autor debe convencer al lector de sus afirmaciones de una manera secuencial, desde lo más importante a lo menos importante. En la “Movida 2” se debe identificar el tópico principal o área del problema que será discutido, desde una perspectiva orientada por los objetivos de la investigación.

De esta manera, una vez identificado el problema o tópico en relación con los antecedentes teóricos en la primera movida, el autor debe cuestionar o desafiar las afirmaciones asentadas en el conocimiento disciplinar. En la “Movida 3” se deben indicar los contenidos, estructura y/o objetivos del artículo en relación con la información teórica existente y el tópico principal. Además, se debe identificar explícitamente el foco del artículo.

Por último, Swales (1990) aporta algunas pistas léxicas o rasgos lingüísticos característicos de cada “Movida” y “Paso”, las cuales presentamos traducidas y adaptadas al español.

Tabla 2. Pistas léxicas de la introducción, según el modelo clásico de Swales (1990).

MOVIDA	PASO	Ejemplo de pista o rasgo léxico
MOVIDA 1 “Estableciendo el territorio”	Paso 1: Pretendiendo centralidad	Recientemente, ha habido un amplio interés en..
	Paso 2: Haciendo generalizaciones sobre el tópico	Un procedimiento estándar para la evaluación ha sido...
	Paso 3: Mostrando ítemes de investigaciones previas	Verbos como mostrar, demostrar, establecer, etc.
MOVIDA 2 “Estableciendo el nicho”	Paso 1a: Oposición	Cuantificadores negativos o cuasi negativos (no, poco)
	Paso 1b: Indicando el vacío	Negaciones léxicas (verbos como falta de, carecer de) o adjetivos como erróneo, incompleto, etc.
	Paso 1c: Preguntas emergentes	Preguntas, expresión de necesidad/deseo/interés, conclusiones lógicas, comentarios contrastivos, etc.
	Paso 1d: Continuando con la tradición	
MOVIDA 3 “Ocupando el nicho”	Paso 1a: Delineando los propósitos	El presente artículo, aquí se analizará, se informa sobre,.
	Paso 1b: Anunciando la investigación	
	Paso 1c: Anunciando los resultados principales	El propósito de este artículo es...
	Paso 1d: Indicando la estructura del artículo	El artículo se estructura, en los siguientes apartados, etc..

Cabe señalar que, con el transcurso de los años, este modelo ha sido criticado por diferentes autores, por ejemplo Chu (1996: 11) quien plantea:

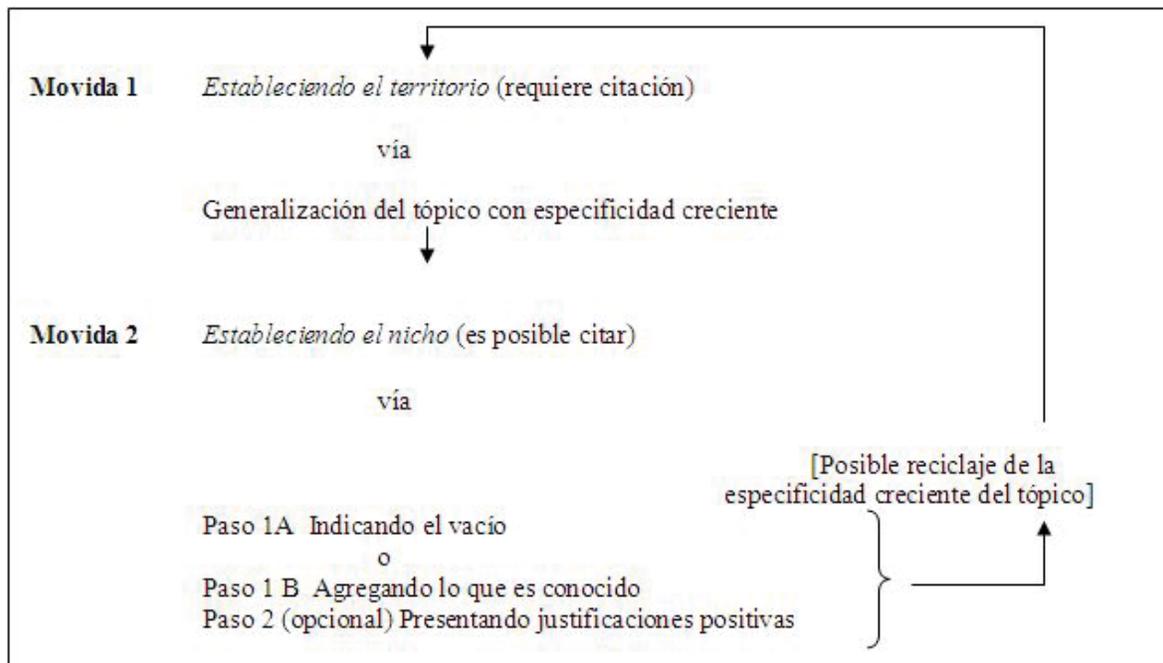
“He [Swales] also maintained that words such *suffer, is limited to, time consuming, expensive* and *not sufficiently accurate* are indications of gaps of some kind. However, these lexical items can also be interpreted as inherent problems of the previous methods rather than gaps to fill, and thus can be seen as counter-claims of various sort” .

Este tipo de crítica al modelo hace que Swales (2004) proponga una reducción de los cuatro pasos a dos pasos en la “Movida 2”, acogiendo así el potencial cíclico o iterativo de la secuencia que va de la “Movida 1” a la “Movida 2”, aspecto que, según Swales (2004), muchos investigadores han encontrado importante de rescatar, sobre todo en las introducciones largas.

Otro elemento que debe ser modificado, según Swales (2004), responde al propuesto por Samraj (2002), quien muestra en sus investigaciones que la indicación de un “vacío” es inmediatamente seguida por lo que la autora denomina “presentando justificaciones

positivas”. En la siguiente Figura 3 se presenta el Modelo CARS Revisado para las “Movidas 1 y 2”.

Figura 3. “Movidas 1 y 2” del Modelo CARS reestructuradas por Swales (2004: 230)



En cuanto a la “Movida 3” Swales (2004) propone hacer algunos cambios en el modelo clásico, considerando los trabajos de Chu (1996), Anthony (1999), Dressen y Swales (2000) quienes plantean que la separación entre el paso de apertura (delineando el propósito/ anunciando la investigación) y los últimos no siempre es fácil. La reestructuración para esta tercera “Movida” (ocupando el nicho) es más compleja y elaborada que la originalmente propuesta. De este modo, la nueva estructura propuesta por Swales (2004) es presentada a continuación en la Figura 4.

Figura 4. Reestructuración de la Movida 3 según Swales (2004:).

Movida 3 *Presentando el trabajo (es posible citar)*

vía

Paso 1 (obligatorio) Anunciando descriptiva o propositivamente la investigación

Paso 2 * (opcional) Presentando preguntas de investigación o hipótesis

Paso 3 (opcional) Clarificaciones definicionales

Paso 4 (opcional) Resumiendo los métodos

Paso 5 (PISF**) Anunciando los principales productos

Paso 6 (PISF) Estableciendo el valor de la investigación

Paso 7 (PISF) Delineando la estructura del artículo

- * Los pasos dos al cuatro no solo son opcionales sino que también son menos fijos en su orden de ocurrencias que otros pasos.
- ** PISF (del inglés, probable en algunas áreas, pero improbable en otras)

En suma Swales (2004), considerando algunas de las investigaciones basadas con su Modelo CARS para el estudio de las introducciones en inglés durante más una década, renueva su modelo ajustándose a los nuevos datos y proponiendo un modelo que parece más flexible y preciso que el modelo clásico.

Cabe señalar que este tipo de estudios, centrados en la identificación de las estructuras retóricas de los textos producidos en el ámbito científico, proliferó en las lenguas anglosajonas durante la década de los años noventa (por ejemplo, Gnutzmann & Oldenburg, 1991; Paltridge, 1997) y se extendió más tarde a otras lenguas, entre ellas el español (por ejemplo, Salager-Meyer, 1991, 1992; Moyano, 2000; Ciapuscio & Otañi, 2002)

En particular nos interesa destacar la propuesta de Moyano (2000), para la estructuración retórica de la introducción en los artículos de investigación científica. Según esta autora, la introducción debe incluir los siguientes cinco contenidos: a) definición del problema, b) antecedentes, c) justificación, d) objetivos y e) limitaciones y alcances del trabajo.

a) Definición del problema: El investigador inicia un estudio porque hay un problema que despierta su interés. Se plantea interrogantes que no puede responder, puesto que todavía no existe un saber especializado al respecto. En el momento de la redacción el investigador debe definir el problema lo más claramente posible y destacar no sólo la naturaleza de este sino también su importancia.

b) Antecedentes: En la introducción se presenta una revisión breve de otros estudios que hayan tratado el problema que se intenta solucionar. Esta revisión tiene una doble finalidad, por una parte, sirve para actualizar los conocimientos del lector acerca de la cuestión. Por otra, es una demostración de que el investigador no ignora el estado de avance de la investigación en el tema, permitiéndole insertar sus afirmaciones en el contexto de producción de la disciplina. Esta revisión debe ser concisa, ya que, en general, el artículo de investigación no debiera superar las 25 páginas aproximadamente, aunque esto depende del tipo de revista que publicará el artículo. Los antecedentes llevarán a la justificación, puesto que la introducción es una argumentación destinada a establecer la necesidad del estudio y, por lo tanto, el valor de dar a conocer sus resultados. El investigador pondrá en diálogo los trabajos de los diferentes autores sobre el mismo tema, entre los que encontrará posiblemente posiciones contrapuestas, y establecerá la suya frente a ellos, expondrá sus propias hipótesis.

c) Justificación: Dado todo el marco de referencia y los antecedentes, el autor debe expresar su propósito al añadir más información a la existente. Esto es, destacar el aporte del trabajo en términos de adición de conocimiento al tema de estudio en la disciplina.

d) Objetivos: Al finalizar el tramo de la justificación, el emisor del artículo tendrá que señalar en forma breve, precisa y destacada cuál es la información nueva que pretende aportar al tratamiento del tema. Los objetivos de un trabajo de investigación expresan una o pocas preguntas de investigación que serán respondidas en las conclusiones del estudio sobre las bases de los resultados obtenidos. Los objetivos deben ser, por lo tanto, concretos y medibles.

e) Limitaciones y alcances del trabajo: El autor debe en este tramo señalar en qué contexto o para qué condiciones son válidas las afirmaciones que vaya a realizar, dado que, en

general, las conclusiones de una investigación no tienen validez universal. Por lo anterior, es conveniente destacar aquellos aspectos que no se hayan podido demostrar en el curso de su investigación. También es importante señalar, en este, cuáles son los alcances y proyecciones tanto teóricas como prácticas que pueda tener su investigación. Por último, es útil aprovechar esta parte del artículo para definir los términos técnicos menos conocidos y las abreviaturas y siglas que serán usadas en adelante, ya que esto aporta claridad al texto en general.

El interés por dar cuenta de algunos estudios respecto de la introducción del artículo de investigación científica (en inglés y en español) se fundamenta, en esta investigación, debido a que es un apartado obligado del género artículo de investigación, independiente de la disciplina científica de la que se trate (y de todo tipo de texto). Es por ello, un común denominador que nos servirá para cuantificar las similitudes léxico-semánticas entre el resumen y la introducción y las palabras clave y la introducción, permitiéndonos establecer un patrón de comportamiento semántico para las distintas áreas de la ciencia tratadas en esta investigación (ciencias exactas, ciencias biológicas y ciencias sociales); de este modo, podemos no solo establecer comparaciones entre las similitudes léxico-semánticas, sino que también y, en una medida exploratoria, podremos llevar a cabo comparaciones en el nivel de las relaciones retórico-estructurales, utilizando a diferencia de los estudios mencionados, patrones de cuantificación de similitud semántica.

CAPÍTULO 2

“Estudios del significado y su representación”

2.1 Objetivo del capítulo

El objetivo de este capítulo es revisar crítica y selectivamente las bases semánticas sobre las cuales se asienta la herramienta computacional de análisis semántico a utilizar en este estudio: *El Análisis Semántico Latente* (LSA, por su sigla en inglés). Esta herramienta será la que nos permitirá llevar a cabo los análisis cuantitativos de las relaciones semánticas entre las palabras clave, el resumen, la introducción y el contenido de los artículos de investigación científica, permitiendo de esta manera crear una representación matemático-estadística de estas relaciones.

Para llevar a cabo este objetivo, en un primer momento, comentaremos brevemente algunos de los hitos más importantes en la historia de la semántica. Seguidamente, se pasará a revisar los principales enfoques teóricos respecto del estudio del significado. Más adelante, se estudiarán algunos conceptos de “significado” y se mostrarán diversos estudios respecto de las relaciones asociativas del significado, tanto a nivel del léxico, del enunciado, como del texto. Luego, se presentarán los dos enfoques cognitivos que, a partir de las relaciones de significado, permiten representar conocimiento, estos son el enfoque simbólico y el enfoque conexionista. Esto último se justifica, pues son estos enfoques los que de una u otra manera nutren las representaciones computacionales del significado, entre las cuales destaca el modelo vectorial LSA. Este modelo es capaz de representar cuantitativamente, a partir de medidas de similitud semántica, las relaciones semánticas que se presentan entre palabras, entre palabras y unidades mayores (oraciones o párrafos) y entre estas unidades mayores en un texto dado.

2.2 Visión histórica de los estudios del significado

Los estudios respecto de la significación han estado presentes desde épocas muy remotas en la preocupación por conocer y entender el lenguaje. Un claro ejemplo de esto es lo aportado por el gramático hindú Panini en el Siglo IV antes de Cristo, quien extendió sus observaciones sobre el sánscrito más allá de los principios normativos que imponía el estudio del Veda y se preocupó del estudio y descripción del valor y del empleo de las palabras, así como de los aspectos fonéticos y gramaticales del sánscrito. Todo esto fue descubierto recién en el Siglo XVIII por los gramáticos comparatistas (Leroy, 1969). Otro aporte que da cuenta de lo antiguo de esta preocupación por la significación es el de Platón (1981) quien, en el Crátilo, expone algunas tesis que estarán a la base de las preocupaciones sobre el lenguaje humano; en especial, en la introducción y en las conclusiones de ese trabajo, en donde se preocupa por la relación entre significante y significado, por la arbitrariedad del signo lingüístico y por el valor social del lenguaje.

Con Aristóteles (1966) las preocupaciones por el lenguaje se ven orientadas fundamentalmente hacia la gramática, con un cariz fundamentalmente normativo, lo que es seguido más tarde por los latinos. Entre estos destaca Sexto Empírico, quien hacia el año 180-200 d.C. en su libro “*Adversus mathematicus*” explica que el significado, el significante y el objeto se unen entre sí de tres maneras. El significante es la imagen fónica, el significado es la cosa misma expresada por la imagen fónica, y ambos, significante y significado, son percibidos simultáneamente. Con el advenimiento de la lingüística moderna esta idea será tratada con especial énfasis (Fernández, Hervás & Báez, 1979).

Siglos más tarde, Petrus Heliae (S. XII), el iniciador de la lógica lingüística, al tratar de responder a las preguntas: ¿qué significan las palabras? y ¿cómo significan las palabras?, da inicio a numerosos tratados de *modis significandi*, formulándose la teoría en términos inequívocos: “*la significación es la forma de la palabra o al contrario: el ser designado no está en la forma fonética que es el signo del ser designado*” (Fernández et al., 1979: 15).

Durante los siglos siguientes, la preocupación por la significación fue muy esporádica y, fundamentalmente, orientada hacia estudios de carácter etimológico. De esta manera, como plantea Fernández et al. (1979: 16) “*Hay que llegar hasta fines del siglo XIX, siglo de la*

filología, para que los estudios semánticos cobren eco propio". Así es como August Friedrich Pott (1974), seguidor de Wilhelm von Humboldt, escribe acerca de la importancia de reparar en lo que se llama la diversidad de significación de una palabra, ya que tal significación no se encuentra en la palabra misma, sino en su relación en el interior de la frase, abriendo el camino hacia los estudios semánticos propiamente tales. Aparece en esta época la figura del filósofo alemán Reisig (1839), quien, en sus "*Vorlesungen über Lateinische Sprachwissenschaft*" (seminarios sobre lingüística latina)², desarrolló un nuevo concepto de gramática en la que la *semiologie* o estudio del significado formaba parte. Según Fernández et al. (1979), Reisig (1839) concebía a la semasiología como una disciplina histórica cuya finalidad debía ser el establecimiento de los principios que rigen la evolución del significado. Sin embargo, la introducción del término *semántica* en la ciencia del lenguaje será propuesto algunos años más tarde por Michel Bréal en su artículo de 1883 "*Les lois intellectuelles du langage*", en donde justificaba el neologismo de la siguiente manera:

"L'étude où nous invitons le lecteur à nous suivre est d'espèce si nouvelle qu'elle n'a même pas encore reçu de nom. En effet, c'est sur le corps et sur la forme des mots que la plupart des linguistes ont exercé leur sagacité: les lois qui président à la transformation des sens, au choix d'expressions nouvelles, à la naissance et à la mort des locutions, ont été laissées dans l'ombre ou n'ont été indiquées qu'en passant. Comme cette étude, aussi bien que la phonétique et la morphologie, mérite d'avoir son nom, nous l'appellerons la SÉMANTIQUE (du verbe [*sêmainô*]) c'est-à-dire la science des significations." (Bréal, 1883: 1).

De esta manera, Bréal (1897), consecuentemente con su época, quiso designar con el nombre de semántica al estudio de las leyes que rigen los cambios de significado, la elección de expresiones nuevas y el nacimiento y muerte de las locuciones. Este estudio se convertirá en la ciencia de las significaciones o la ciencia del sentido o, mejor dicho, en las diversas teorías lingüísticas del sentido.

Bréal (1897) lo plantea de la siguiente manera en el prólogo de su Libro "*Essai de Sémantique*":

² Todas las traducciones desde el alemán son nuestras.

“J’ai pris à dessein mes exemples dans les langues les plus généralement connues : il sera facile d’en augmenter le nombre ; il sera facile aussi d’en apporter de régions moins explorées. Les lois que j’ai essayé d’indiquer étant plutôt d’ordre psychologique, je ne doute pas qu’elles ne se vérifient hors de la famille indo-européenne. Ce que j’ai voulu faire, c’est de tracer quelques grandes lignes, de marquer quelques divisions et comme un plan provisoire sur un domaine non encore exploité, et qui réclame le travail combiné de plusieurs générations de linguistes. Je prie donc le lecteur de regarder ce livre comme une simple Introduction à la science que j’ai proposé d’appeler la Sémantique” (Breál, 1897: 9).

Tanto Reisig (1839) como Breál (1883, 1897) consideraron a la semántica como un estudio puramente histórico del lenguaje en el cual se intenta explicar las evoluciones de significado en forma de leyes, acorde con la preocupación neogramática de construir la lingüística sobre el modelo de las ciencias exactas (Fernández et al., 1979; Leroy, 1969). En este sentido, la iniciativa de Breál (1897) tuvo el mérito de cristalizar, en forma elegante y lúcida, las ideas de análisis y clasificación de los cambios de significación de acuerdo a ciertos principios lógicos (Fernández et al., 1979) y constituir a la semántica como una ciencia cuya materia la conformaba el estudio de la causa y estructura de los procesos de cambio en los significados de las palabras (Schaff, 1972).

Comienza aquí una nueva etapa en la historia de la semántica, etapa en que la semántica solo tendrá éxito entre el público aficionado. Los lingüistas seguirían negándole el carácter científico, porque consideraban que el léxico era un repertorio de significados sumados sin relación estricta. A esta crítica, se le sumaba una evidente posición antimentalista (Baldinger, 1970).

La nueva ciencia alcanzó, luego de muchas discusiones sobre el lugar que le correspondía entre las disciplinas lingüísticas, mayor asentamiento cuando el filólogo danés Nyrop (1913) le dedicó el cuarto volumen de su “*Grammaire historique de la langue française*”. Sin embargo, el intento fue todavía algo prematuro, porque la semántica no disponía aún de las técnicas necesarias para identificar las tendencias distintivas de una lengua particular.

A partir de Ferdinand de Saussure (1983), la lingüística sufre un cambio paradigmático que va a afectar a todos los planos de la lengua, incluyendo a la semántica que se va a enfocar

desde una perspectiva netamente sincrónica. A pesar de que como plantean Bally y Sechehaye (1983: 61):

“[...] el maestro trabaja en profundidad y no se extiende por la superficie más que donde sus principios hallan aplicación, especialmente patente, o donde topan con alguna teoría que lo podía comprometer”.

Lo anterior sin duda da valor a la obra, mas esto mismo:

“[...] explica que ciertas disciplinas estén apenas desfloradas; la semántica, por ejemplo“ (Bally & Sechehaye, 1983: 61).

No obstante Saussure (1983) sienta las bases de los estudios del significado acogidos bajo el alero de la lingüística, asumida ya como ciencia, a través de su teoría del signo lingüístico en la que el “significado” y las “significaciones” son elementos primordiales y de carácter social, ya que es la comunidad hablante la que establece unos valores de significación para cada significante, y sólo con el asentimiento de esa misma comunidad se consuman los cambios, ampliaciones y restricciones de los significados. También fue importante la distinción que estableció entre “significado” y “sentido”, por la cual el sentido es un elemento de la significación y se obtiene por la contraposición de unos signos a otros en la frase, porque como plantea Saussure (1983:188):

“La lengua es un sistema cuyos términos se condicionan recíprocamente y el valor o sentido de unos es el resultado de la presencia simultánea de otros”.

Además, plantea que los elementos del lenguaje no adquieren valor más que oponiéndose a otros, no confundiéndose con otros; así no es su cualidad propia y positiva lo que los caracteriza, sino antes bien su calidad de oposición y su valor diferencial. Por último, con Saussure (1983) queda expreso el tránsito de los estudios de la palabra aislada hacia el análisis de la palabra en el sistema de la lengua.

El primero en aplicar esta nueva concepción de semántica fue Jos Trier (1931), quien expone su teoría de los campos lingüísticos o campos semánticos, la cual, tendiendo un puente entre el plano de los conceptos y el de la expresión, aparece como una especie de *sprachliche Zwischenwelt* (mundo lingüístico intermedio) y ayuda a comprender mejor esas relaciones entre significante y significado (Leroy, 1969). Para Trier (1931) son “campos”

las realidades lingüísticas vivas existentes entre palabras individuales y el vocabulario total; son partes de un todo y se parecen a las palabras en que las unidades menores se articulan en cierta unidad superior. Trier (1931) concibe el “campo léxico” como un sistema organizado en el que todo se equilibra, en que los elementos se delimitan recíprocamente y obtienen su valor según el puesto que ocupan dentro del campo. Encontramos aquí la concepción saussureana de la delimitación recíproca de los valores, donde todo cambio de valor de un concepto entraña una modificación de los valores de los conceptos vecinos siendo las palabras, especialmente las del vocabulario intelectual, las que reflejan esas modificaciones del contenido de los conceptos (Leroy, 1969).

Sobre la base de esta teoría de los campos semánticos, Bally (1940), más tarde, desarrolla la noción de *campo asociativo*, noción que es aplicada al vocabulario y no a la estructura, y, además, limitada a las relaciones de orden semántico. Con este trabajo, el discípulo de Saussure se limitaba a hacer más explícita la idea de “constelaciones” o “relaciones asociativas”. Asociaciones que fueron organizadas en el *Curso*, dependiendo de si el elemento común era la radical, el sufijo, o bien si la asociación reposaba en la analogía de los significados o bien en la afinidad de las imágenes acústicas (Saussure, 1983).

A partir del año 1950 aumenta considerablemente el interés por el estudio de la semántica y con ello se inicia lo que Baldinger (1970) denominó la tercera etapa: la de los intentos científicos sobre semántica.

El primer autor que destaca en este periodo es Stephen Ullmann (1967) quien, inspirándose en las consideraciones de Bally (1940), establece inicialmente la distinción entre cambios debidos al conservadurismo lingüístico y los debidos a la innovación lingüística; verificando que en estos últimos puede haber transferencia del nombre o transferencia del sentido y que tal transferencia se efectúa sea por similitud o por contigüidad; en consecuencia, construye un esquema que tiene en cuenta todos los posibles tipos de asociaciones. De esta manera, Ullmann (1967) permanece fiel a la antinomia saussureana “sincronía/diacronía” y considera evidente el contraste entre estas dos perspectivas. A partir de esta antinomia, se aboca a la tarea de edificar una semántica “pancrónica” y se fija la tarea de determinar lo que es común a todas las lenguas y a todas las épocas, utilizando el análisis binario de “pansicronía y pandiacronía” (Leroy, 1969).

En definitiva, como lo plantea Baldinger (1970), Ullmann (1967) es quien desde una semántica tradicional y semiestructuralista establece un cuerpo de doctrinas sincrónicas alrededor del famoso triángulo de *nombre-concepto-cosa*, cuya forma exterior se remonta a un ensayo de Ogden y Richards (1984), que se reimprime de manera constantes desde 1923, “The Meaning of the Meaning” (el significado del significado). En la versión de Ullmann (1967), el triángulo da cuenta, acorde con la postura saussureana, de las relaciones existentes entre el significado, el significante y la realidad. De esta manera plantea que el significante (o imagen acústica) simboliza un significado (concepto), el que a su vez se refiere a la cosa o realidad y que ésta es representada, por medio de una relación convencional, por el significante.

Sin duda, es a partir de mediados de los años 50 que la semántica cobra cada vez mayor interés, encontrando su punto culmine en Europa hacia 1957, año en el que se celebró el Congreso de Oslo del Círculo Lingüístico de Praga (CLP). En este evento se trató la cuestión sobre ¿hasta qué punto puede considerarse que los significados de las palabras forman una estructura? y se dio a conocer, nada menos que por intermedio de Louis Hjelmslev (1974), el programa de una semántica estructural, en el cual se establecía que “[...]introducir la noción de estructura en el estudio de los hechos semánticos era introducir la noción de valor al lado de significado” (en Leroy, 1969: 181).

Hasta aquí dejaremos esta síntesis histórica del desarrollo de la semántica, con el convencimiento de que hay mucho aún por decir. En lo que sigue, nos abocaremos a una revisión de los enfoques teóricos más importantes durante el siglo XX respecto del estudio del significado. A partir de esta revisión, iremos esbozando el marco lingüístico-cognitivo del concepto de significado que nos interesará destacar.

2.3 Aproximaciones teóricas al estudio del significado en el siglo XX

Como se sabe, las concepciones del significado son muchas y muy diversas (Leech, 1977; Fodor, 1977, Lyons, 1980). Para todas ellas, el significado es la asociación entre la expresión lingüística y otro elemento; las diferencias estriban, precisamente, en el tipo de

elemento con que se relaciona la expresión lingüística. Según Escandell (2004) los distintos enfoques que dan cuenta de esta relación se pueden agrupar en tres grandes categorías: *referencialistas y denotacionalistas, representacionalistas y mentalistas, convencionalistas y pragmático/sociales*. Esto se condice con los tres planteamientos en teoría del significado propuestos desde la filosofía del significado. Estos son: semantista, intencionalista y pragmática (Corredor, 1999).

2.3.1 Las teorías referencialistas y denotacionalistas

Estas teorías enfatizan la relación entre las expresiones lingüísticas y el mundo; es decir, entre las expresiones lingüísticas y los objetos, las entidades y las situaciones del mundo que se describen por medio de la lengua. La teoría referencial del significado, planteada por Frege (1972) y retocada por Russell (1973), plantea la hipótesis de que el significado de una palabra es el objeto denotado por esa palabra. El fundamento principal de este enfoque se sustenta en las nociones de referencia y referente. La referencia es la relación que se crea en un momento concreto entre una expresión lingüística y la entidad o situación del mundo a la que se conecta; puede verse también como el acto que lleva a cabo un hablante para aludir a una entidad o situación por medio de una expresión lingüística. El referente, por su parte, es la entidad o situación particular seleccionada por dicha expresión lingüística (Corredor, 1999; Escandell, 2004).

Los enfoques referencialistas comprenden diferentes teorías. Algunas afirman que el significado de una expresión es aquello a lo que la expresión se refiere, es decir, identifican el significado con el referente. Este tipo de teorías obtiene buenos resultados con algunas expresiones, como los nombres propios. Sin embargo, presenta una serie de limitaciones. En primer lugar, la teoría parece predecir, equivocadamente, que las expresiones que no tengan referente no tendrán significado. Así las expresiones que carecen de contenido descriptivo como: las que poseen contenido gramatical (ver ejemplo 1), las que tienen contenido expresivo (ver 2), las que, sin carecer exactamente de contenido descriptivo, no sirven para describir estados de cosas preexistentes, porque es precisamente la utilización de la expresión en las condiciones adecuadas la que crea nuevos estados de cosas. A esta categoría pertenecen los llamados predicados realizativos (ver ejemplo 3). También se

encuentran en esta situación las expresiones que no tienen un referente en la realidad como: las que aluden a entidades abstractas (por ejemplo, 4) y, por último, las que aluden a seres o acontecimientos imaginarios, como en el ejemplo 5.

- 1) “Ahora bien” o “por lo tanto”.
- 2) “¡ajá!” o “¡caramba!”.
- 3) “Yo os declaro marido y mujer” o “queda inaugurada esta sala”.
- 4) “La tranquilidad que se respiraba allí” o “ el deseo de verla de nuevo”.
- 5) “El unicornio pastaba junto al río”.

Por otra parte, la teoría parece predecir que dos expresiones que tengan el mismo referente tendrán el mismo significado. A estas expresiones que identifican el mismo referente se les llama correferenciales. Sin embargo, sabemos que el contexto de comunicación juega un rol preponderante en la elección de las expresiones; así, expresiones como en 6 y 7, si bien refieren a una misma persona, no necesariamente significan lo mismo.

- 6) “El bombardero de La Reina”.
- 7) “El tenista Fernando González”.

Debido a estas restricciones, es que la teoría referencialista ha debido revisar sus postulados y se han planteado soluciones como que el referente no necesariamente tiene una entidad física. Así, la referencia puede ser establecida para mundos imaginarios. Por otra parte, frente a la idea de las expresiones con un mismo referente, se han postulado distinciones que permiten refinar el modo de abordar las relaciones entre una expresión y su referente (Escandell, 2004). La primera distinción que se debe realizar es entre expresiones que tienen un referente constante y único, que no varía de situación a situación (por ejemplo, “el océano pacífico”) y las que tienen un referente variable, es decir, un referente que cambia según las circunstancias en las que se emplee la expresión, por ejemplo en 8. La variación reside no sólo en que, por ejemplo, existen muchos “presidentes de los congresos del mundo”, sino también en que cada congreso tendrá, según lo establezca la ley, distintos presidentes cada cierto periodo de tiempo.

8) “El presidente del congreso”.

En cuanto a las expresiones correferenciales, Frege (1972) propone la distinción entre los conceptos de “sentido” y “referencia”. El sentido viene determinado por las relaciones que contrae la expresión con otras expresiones dentro del sistema. El sentido de una expresión es el modo en que dicha expresión presenta al referente, por ello recibe el nombre de “modo de presentación” (Escandell, 2004). Por ejemplo, en una situación dada, las expresiones 9 y 10 pueden ser correferenciales; sin embargo, la manera de cada una de presentar o introducir el referente es muy distinta.

9) “La mujer de lentes”.

10) “La primera de la fila”.

Determinar el sentido de una expresión puede ser relativamente fácil cuando el sistema al que pertenece es relativamente limitado y está explícitamente definido; sin embargo, resulta algo más complicado cuando no se está en condiciones de tomar otros modos de presentación diferentes, tal y como ocurre a menudo en el uso espontáneo de la lengua. De modo que la noción no es fácil de aplicar en los términos en que fue propuesta originalmente. Por ello, a partir de esta noción de sentido, se han propuesto reelaboraciones de tipo cognitivo. En ellas, el sentido es la parte constante del significado, la que no varía cuando los referentes cambian. Conocer el sentido de una expresión es precisamente lo que permite identificar el referente adecuado en cada conjunto diferente de circunstancias. Al respecto, Carnap (1958) propone la división de la semántica en teoría de la extensión y teoría de la intensión. La primera estudiaría la relación de las palabras y frases con las cosas (“denotación”, “extensión”); la segunda se ocuparía del significado o sentido de las palabras y de las frases (“connotación”, “comprensión”, “intensión”).

La doctrina de las relaciones entre la extensión y la comprensión se encuentra clásicamente expuesta en la *logique* de Port Royal (Garrido, 1977) de este modo cuanto mayor es la extensión de un predicado menor es su comprensión y viceversa (por ejemplo, “animal” es más extenso que “hombre” y a la vez más reducido en comprensión). La diferencia entre extensión e intensión se aprecia fácilmente analizando el uso de los predicados (nombres comunes). Así, por ejemplo, la extensión del predicado “azul” está determinada por la clase

de objetos que son azules. Pero también cabe decir que el color azul tiene una serie de características propias, como la de ocupar un determinado lugar en el espectro cromático: tales características constituirían el significado intensional del predicado “azul”.

Al considerar a los predicados desde el punto de vista extensional, se dice que aluden, o mejor, que denotan clases o conjuntos. Pero cuando se los contempla desde el ángulo intensional se dice que designan propiedades o notas de los objetos (Garrido, 1977).

En términos más específicos en la denotación, ya no se considera el conjunto de referentes, sino el vínculo que existe entre dicho conjunto y la expresión. Esto es, la relación que se establece entre una expresión y su conjunto extensional. Es una relación constante que no varía de contexto a contexto y de situación a situación, mientras que la referencia es una relación que se crea en el acto mismo de utilizar una expresión para aludir a una entidad concreta, la denotación es una propiedad intrínseca de las expresiones (Garrido, 1977).

Cabe, por último, la necesidad de establecer que esta teoría pretende dar cuenta del vínculo entre la lengua y la realidad, es decir, defiende la idea de que los seres humanos entre sus capacidades asociadas al conocimiento semántico pueden establecer una relación entre una expresión lingüística y un determinado estado de cosas del mundo. Esta capacidad está ligada al significado de las expresiones (Corredor, 1999).

2.3.2 Las teorías representacionales y mentalistas

Estas teorías se caracterizan por sostener que el significado de una expresión es la representación mental que asociamos con ella. Por medio de la lengua podemos, efectivamente, referirnos al mundo, pero lo hacemos siempre a través de representaciones cognitivas, las que nos permiten establecer el vínculo entre la lengua y la realidad. El identificar el significado con una representación o imagen mental nos permite visualizar, obviamente, el significado de las expresiones que designan seres y objetos de la realidad; pero también nos permite representarnos entidades imaginarias (por ejemplo, “unicornios pastando junto al río”).

Pese a lo atractivo de este enfoque, las limitaciones que se le reconocen son las siguientes:

- 1) Hay expresiones para las cuales no resulta sencillo producir una imagen mental, por ejemplo, las expresiones con contenido gramatical (i.e. ¿qué imagen corresponde a “también” o “sin embargo?”), las expresiones de contenido abstracto (i.e. ¿cuál es la imagen para “la quietud del lugar”?) y las expresiones de contenido expresivo (i.e. “¡por la chita!”). ¿Qué sucede con estas expresiones?, ¿acaso no tienen significado, porque no se puede realizar una imagen mental de las expresiones? Obviamente, lo tienen (Escandell, 2004).

Por otro lado, las imágenes mentales son representaciones privadas. Cada individuo puede representarse una expresión a su manera. Es por ello que el significado de una expresión variará considerablemente de un individuo a otro.

Para salvar las dificultades expuestas, se han propuesto otros enfoques. El más conocido es el que postula que las representaciones mentales no deben identificarse con imágenes, sino con conceptos (Garrido, 1977; van Dijk, 1983; Corredor, 1999; Escandell, 2004). Un concepto es una representación mental que recoge sólo los rasgos esenciales y constitutivos de una entidad. Gracias a estos rasgos, podemos identificar entidades y relacionar sus propiedades con las de otras entidades semejantes. Los conceptos son, por tanto, mucho más abstractos que las imágenes mentales. El concepto asociado a una expresión puede identificarse con lo que se denomina comprensión o intensión, como ya fue mencionado en el apartado anterior. Este es el conjunto de rasgos que comparten los elementos pertenecientes a un conjunto. Esta noción de concepto permite la estabilidad cognitiva e intersubjetiva que asegura la categorización y comunicación entre individuos diferentes (Garrido, 1977; van Dijk, 1983; Corredor, 1999; Escandell, 2004). De esta manera, el acercamiento al significado puede realizarse por extensión (listando uno por uno los ejemplares que componen la expresión) o por intensión (expresando las propiedades características que los unen).

De este modo, al eliminarse la necesidad de que las representaciones mentales tengan componentes visuales, se eliminan en parte los problemas relacionados con aquellos significados que difícilmente se asocian con imágenes visuales. Y, por otra parte, al manejarse una noción abstracta que contiene solo rasgos esenciales se corrigen los

problemas derivados de la variabilidad inherente a las imágenes mentales. La imagen que cada uno de nosotros asocia a una expresión puede ser muy diferente, los conceptos en cambio, tienen la propiedad de estabilizar cognitiva e intersubjetivamente los significados por medio de la categorización, lo cual permite la comunicación entre individuos diferentes.

Para poder trabajar con conceptos debemos poder definir de manera precisa y explícita el contenido de cada concepto. Esto requiere establecer los rasgos necesarios y suficientes. Así, por ejemplo, para definir “oro” habría que entregar criterios que permitan distinguir el oro de los otros metales. Estos criterios existen y son bien conocidos por los expertos, pero la gran mayoría de los hablantes no son especialistas que conozcan estas propiedades definitorias, sin embargo, normalmente no se equivocan en el uso del término. Por lo tanto, se nos plantea la cuestión de cuáles son los rasgos propios suficientes para definir un concepto. ¿Los que conoce y usa el experto?, ¿los que utilizan los hablantes no expertos? Como respuesta a estas interrogantes se pueden reconocer dos tipos de conceptos: los “especializados”, con una identificación acabada de rasgos integradores, que determinan la univocidad semántica y los “comunes”, con rasgos identificadores, apenas suficientes de un concepto. Esta cuestión aún se encuentra en constante debate, y son los trabajos realizados en áreas como la terminología y la lexicología los que se han preocupado más detalladamente de esta problemática (Pérez, 1999; Albentosa & Moya, 2000, Cabré & Feliu, 2001).

2.3.3 Las teorías convencionalistas y pragmático-sociales

Las teorías convencionalistas y pragmático-sociales identifican el significado de una expresión con el uso que de ella se hace en la actividad lingüística. Este enfoque pone el énfasis en las relaciones entre las expresiones y los usuarios.

Las teorías pragmatistas del lenguaje, como las denomina Corredor (1999), comparten una explicación pragmática del significado en términos del uso que los hablantes hacen del lenguaje. Estas teorías coinciden en enfatizar la función práctica del lenguaje como medio para el establecimiento de relaciones interpersonales y, en particular, para la comunicación humana. Estas teorías del significado como uso, parten de la idea de Wittgenstein (1988) en la que la pluralidad de usos posibles del lenguaje hace que la función representativa pierda

su posición privilegiada (Corredor, 1999). Para este filósofo el medio del lenguaje no sirve, de modo preeminente, para describir o fijar hechos; junto a estos usos del lenguaje se sitúan otros muchos, como el dar órdenes o seguirlas, hacer promesas, narrar cuentos, saludar o apostar. Austin (1982) intenta sistematizar y ordenar los posibles usos del lenguaje dentro de unos pocos modos de empleo básicos, valiéndose para ello de los llamados verbos realizativos explícitos. Ello le permitió analizar el doble rendimiento de los actos de habla. Por medio de ellos, el hablante, al tiempo que dice algo, hace algo con palabras.

Esta manera de abordar el significado se apoya en la idea de que el que usa de la lengua está sometido a convenciones sociales. La existencia de pautas establecidas está muy clara en aquellos casos en los que hay una institución que respalda el contenido asociado a la expresión lingüística, como en el caso de las declaraciones de Bautizo respaldadas por la Iglesia o las de matrimonio respaldada también por la iglesia y/o el Estado. Son estas instituciones las que determinan la validez de las expresiones. La situación es menos clara, en cambio, cuando no hay convenciones tan firmes que asienten el uso. De esta manera, una expresión como 11 puede emplearse para realizar una enorme variedad de actividades lingüísticas: desde informar a advertir, desde amenazar hasta consolar; los usos son potencialmente muy amplios.

11) “es tarde”.

En este sentido, dado que para cada expresión la situación de uso generaría un significado distinto, estaríamos frente a una polisemia múltiple y los significados estarían afectos a las convenciones propias de la situación. Esto quiere decir que el significado está internamente vinculado con las pretensiones de validez que el hablante presenta ante su interlocutor en la realización del acto de habla (afirmación, declaración, promesa, compromiso, etc.) y que permiten a este último adoptar una posición racionalmente crítica, pues el vehículo entre significado y pretensiones de validez se establece mediante un potencial de justificación que opera mediante razones. En otras palabras, el significado es validado o invalidado por el interlocutor en atención a las razones presentadas por el hablante, las cuales pueden ser esbozadas, por ejemplo, desde la experiencia personal, desde la autoridad en algún campo del conocimiento, desde la tradición, desde una comunidad discursiva particular, desde una ideología, etc. (Escandell, 2004).

Estas perspectivas teóricas se plantean como teorías intersubjetivistas del significado, esto es, el significado pragmático (o fuerza ilocutiva) de habla se explica en términos de un conjunto de dimensiones de validez, que han de verse como pretensiones de reconocimiento intersubjetivo (Corredor, 1999; Escandell, 2004).

El estudio de estos significados pragmáticos han tomado fuerza en los últimos años con las distintas vertientes del análisis crítico del discurso, en las cuales las expresiones contenidas en los textos a analizar significan según las variables ideológicas, sociales y/o políticas en estudio (Fairclough, 1992, 1995; van Dijk, 1997, 1999, 2003; Wodak & Meyer, 2003).

2.4 Tipos de significado

Como hemos visto, existen diversos enfoques teóricos respecto del estudio del significado. Esta misma diversidad teórica ha redundado en la constitución polisémica del concepto de “significado”, es por ello que habrá tantos significados de “significado” como teorías y disciplinas que se ocupen del lenguaje (Osgood, Suci & Tannenbaum, 1976). Diversos autores se han preocupado de deslindar algunos de estos significados, sin embargo, muchos reconocen la pertinencia de la clasificación de Leech (1977). Este distingue al menos ocho tipos de significados, que revisaremos brevemente.

2.4.1 Significado conceptual: El también llamado significado “denotativo” o “cognoscitivo” es, para Leech (1977), el factor fundamental de la comunicación lingüística. Esto porque el significado conceptual posee una organización sutil y compleja, que le permite relacionarlo con los niveles sintáctico y fonológico del lenguaje, en tanto que todos comparten los principios de contrastividad y el de estructura constituyente (ejes paradigmáticos y sintagmáticos). Los rasgos contrastantes sustentan la clasificación de los conceptos en la semántica estructural, de modo que cualquiera que sea la etiqueta que asignemos a uno de ellos, los rasgos los definen positivamente en virtud de los rasgos que posee y, por implicación, negativamente, en virtud de los rasgos que no posee; así el significado de la palabra “mujer” se podría especificar por los rasgos: [+humano, -masculino, +adulto]. En tanto que “muchacho” podría definirse como: [+humano, +masculino, -adulto]. El principio de la estructura constituyente es aquel que sostiene que

las unidades lingüísticas mayores están compuestas de otras más pequeñas. De esta manera, es posible, por ejemplo, descomponer una oración yendo desde sus constituyentes inmediatos a sus constituyentes últimos, pasando por una serie de estadios jerarquizados y pudiendo graficarse en forma arbórea. Este principio da sustento a la correspondencia entre la estructura semántica y la estructura sintáctica de las lenguas naturales. Con esto se plantea la existencia de una semántica conceptual fundada en los sistemas de la lógica simbólica. Un tercer principio corresponde al de la organización lingüística, según el cual, cualquier pieza del lenguaje está estructurada en dos o más “niveles” simultáneamente, el de la fonología, el de la sintaxis y el de la semántica. Estos niveles son necesarios para rendir plena cuenta de la competencia lingüística, mediante la cual podemos producir o entender diversas locuciones. Esto último, redundaría en que para analizar cualquier oración es preciso elaborar una “representación fonológica”, una “representación sintáctica” y una “representación semántica” y explicitar las derivaciones de un nivel de representación a otro.

El objetivo de la semántica conceptual es, entonces, proporcionar una determinada configuración de símbolos abstractos para cualquier interpretación determinada de una oración. De manera que esa configuración sea la “representación semántica” de la oración en cuestión mostrando con exactitud lo que se precisa saber para diferenciar un significado determinado de todos los demás y, presentando, además, concordancia entre el significado y las formulaciones sintácticas y fonológicas correspondientes.

2.4.2 Significado connotativo: Es entendido como el valor comunicativo que tiene una expresión, atendiendo sólo a lo que ella se refiere, es decir, dejando de lado su contenido puramente conceptual. Aquí la noción de referencia coincide en un alto grado con la de significado conceptual: en el ejemplo de la palabra “mujer”, las propiedades de [+humano], [+adulto], [-masculino] deben suministrar criterios para el uso correcto de esa palabra. Ahora estos rasgos contrastantes, a su vez en el “mundo real”, se convierten en atributos del referente. Pero, sabemos que cualquier referente de mujer posee muchos atributos adicionales, que comprenden atributos físicos ([+bípedo], [+matriz]), atributos psicológicos y sociales ([+gregario], [+instinto maternal]). Además, el significado connotativo puede englobar “propiedades supuestas” del referente, o sea, las que se deban al punto de vista

que adopte un solo individuo, un grupo o una sociedad entera; así, antiguamente la mujer portaba algunos atributos que la sociedad, eminentemente machista, le había asignado ([+débil], [+propensa al llanto], [+cobarde], [+sentimental], etc.). Evidentemente, las connotaciones son susceptibles de variar de una época a otra y de una sociedad a otra, incluso de un individuo a otro dentro de una misma comunidad lingüística. La connotación es de por sí accidental al lenguaje, ya que el significado connotativo no es específico del lenguaje, también lo poseen otros sistemas de comunicación como la música y las artes visuales. Además, las connotaciones son relativamente inestables, varían con la cultura, el período histórico y las experiencias de los individuos. Por último, el significado connotativo es algo indeterminado y sin límites precisos, pues se asienta en nuestros conocimientos y nuestras creencias acerca del universo, contrario a lo que sucede con el significado conceptual, que tiende a ser más estable.

2.4.3 Significado estilístico y significado afectivo: Estos significados están relacionados con la situación en que tiene lugar una expresión. El significado estilístico es lo que un elemento de la lengua expresa acerca de las circunstancias sociales de su empleo; podemos “decodificar” el significado estilístico de un texto solo después de que hayamos reconocido la existencia de distintas dimensiones y niveles de uso dentro del mismo idioma. Así, reconocemos que algunas palabras o expresiones son dialectales, que otras están dadas por las clases sociales de los interlocutores (dimensión diastrática), y otras, por las variaciones en el uso según las situaciones sociales, dimensión diafásica, etc. El significado afectivo, por su parte, puede reflejar las opiniones y las creencias personales del hablante, incluyendo su actitud para con el oyente o su postura ante lo que está hablando. Este significado se transmite a menudo explícitamente a través del contenido conceptual o connotativo de las palabras empleadas. El significado afectivo es, en gran medida, una categoría parasitaria, en el sentido de que para expresar nuestras emociones tenemos que contar con la ayuda de otras categorías del significado (conceptual, connotativo o estilístico). Sólo en el caso de las interjecciones, el significado emocional se reconoce sin la intervención de las otras categorías.

2.4.4 Significado reflejo y significado conlocativo: Estos significados suponen una interconexión en el nivel léxico del lenguaje. El significado reflejo es aquel que se da en los

casos de significado conceptual múltiple, es decir, cuando un sentido de una palabra forma parte de nuestra respuesta a otro sentido. Un ejemplo de lo anterior sería la asignación múltiple de sentido a la palabra cópula, la cual se puede usar tanto en el dominio de la gramática (cópula verbal) como en el de la fisiología del sexo (cópula humana). El significado conlocativo consiste en las asociaciones que una palabra adquiere al tener en cuenta los significados de las palabras que suelen aparecer en su entorno; de esta manera las palabras temblar y estremecerse, si bien son muy similares en su significado conceptual, en su significado conlocativo se diferencian bastante, ya que temblar se asocia a “de frío”, “de miedo”, en cambio, estremecerse se asocia más a “de emoción”, “de pasión”, “de asco”. Sobre este tipo de significado nos extenderemos en 2.4.6.

2.4.5 Significado temático: Este es lo que se comunica gracias a la forma en que el que habla o escribe organiza el mensaje, atendiendo a la ordenación, al foco y al énfasis. Esto se materializa ante todo en la elección que hace el hablante entre construcciones gramaticales alternativas, por ejemplo:

12. a) “Se detuvieron al final del pasillo”

b) “Al final del pasillo, se detuvieron”

Estas oraciones tienen, en un sentido, el “mismo significado”; pero, a pesar de ello, se reconoce que sus efectos comunicativos pueden ser distintos, ya que el énfasis en la acción o en el lugar determinará el uso de una expresión o la otra.

2.4.6 Significado asociativo: Significado reflejo y significado conlocativo, significado afectivo y estilístico tienen más en común con el significado connotativo que con el conceptual: todos tienen el mismo carácter indeterminado y poco preciso en la fijación de sus límites y, además, se prestan mejor al análisis hecho a base de escalas o grados que al que se basa en la elección de una opción que, por fuerza, excluya a las demás; todos ellos pueden agruparse bajo el rótulo de significado asociativo. Para explicar la comunicación a estos niveles nos valemos de la teoría asociativa elemental de las conexiones mentales basadas en la contigüidad de las percepciones empíricas. El significado asociativo contiene tantos factores imponderables que sólo se lo puede estudiar sistemáticamente mediante técnicas estadísticas aproximativas. Osgood, Suci y Tannenbaum (1976) propusieron un

método de análisis parcial del significado asociativo, basado en lo que denominaron “Diferencial Semántico”, que es una técnica estadística para organizar el significado a base de un espacio semántico multidimensional, utilizando como datos los juicios de los hablantes, que se registraban en escalas divididas en siete grados cada una. Estas escalas estaban rotuladas mediante pares de adjetivos contrapuestos tales como alegre-triste, duro-blando, lento-rápido. De tal manera, la persona podía, por ejemplo, registrar sus impresiones sobre una palabra asignándole un valor en la escala propuesta. El resultado más interesante de este análisis fue descubrir que lo realmente esencial parece residir en tres dimensiones principales, a saber: la evaluación (bueno-malo), la potencia (duro-blando) y la actividad (activo-pasivo). Según Leech (1977:37):

“Esta perspectiva de análisis proporciona una explicación parcial y aproximada del significado asociativo, porque, por una parte, entraña una selección de entre las infinitas escalas posibles, las cuales, en cualquier caso, podrían explicar el significado asociativo sólo en la medida en que éste es aplicable a base de aquéllas; y, por otra, debido al muestreo estadístico, y porque una escala dividida en siete grados constituye la división de un continuo en siete segmentos dentro de los cuales no se hace diferencia alguna”.

A pesar de lo anterior, no se puede desconocer el valor de este método de análisis como método de cuantificación del significado asociativo en los estudios de la personalidad y la psicoterapia, sobre todo en el examen de las diferencias existentes entre las reacciones de los individuos, ya que a diferencia del significado conceptual, el significado asociativo es menos estable y varía de acuerdo con la experiencia de los individuos.

Podemos observar claramente una relación entre las teorías expuestas en un primer momento y esta clasificación de Leech (1977). En ella, el significado conceptual se entiende desde la óptica de las teorías referencialistas y denotacionales, basadas en un enfoque simbólico, en donde el significado se obtiene por identificación de los rasgos componenciales y la oposición de estos rasgos en el sistema de la lengua. Por otra parte, el significado temático se sustenta en su relación con el contexto en el cual se realiza una expresión determinada, lo que nos permite comprenderlo desde una óptica centrada en un enfoque pragmático. Por último, el mismo Leech (1977) plantea que el significado connotativo, así como el significado reflejo y significado conlocativo y el significado

afectivo y estilístico se asemejan en su carácter indeterminado y en lo poco preciso en la fijación de sus límites, lo que le permite agruparlos bajo un solo tipo de significado, el significado asociativo. Esta agrupación se sustenta claramente gracias a un enfoque representacional y mentalista, en donde los principios empiristas de la semejanza y la contigüidad de las formas toman relevancia. Por lo tanto, la clasificación de Leech (1977) es reductible según los tres grandes tipos de teorías del significado y, por lo tanto, a tres paradigmas claramente establecidos. Para el significado conceptual, el racionalista simbólico; para el significado asociativo, el empirista asociacionista; y para el temático, el pragmático.

Una vez revisados los diversos tipos de significados, así como las teorías con las cuales se asocian, nos interesa destacar que, debido a que trabajaremos sobre las relaciones de significado existentes entre unidades lingüísticas al interior de un texto, entendido este como el producto semántico-lingüístico de una representación mental, nos enmarcaremos en las teorías representacionales y mentalistas del significado. Como planteábamos más arriba, estas teorías nos permiten entender el significado como una red de relaciones de unidades lingüísticas que permiten a los individuos representarse mentalmente los hechos del mundo. Por lo mismo, el tipo de significado que nos interesa destacar es el que se funda en una visión asociacionista colocacional o conlocativa (como plantea Leech, 1977).

2.4.6.1 Estudio del significado asociativo colocacional

Como ya mencionábamos más arriba el significado asociativo es factible de ser analizado a partir de la identificación de las relaciones entre las unidades lingüísticas, por lo que se presta muy bien para análisis de tipo cuantitativo como los que hicieron Osgood et al., (1976), utilizando el Diferencial Semántico.

En nuestro caso, el análisis de las relaciones semánticas de tipo asociativo que realizaremos se sustenta en la noción de “similitud semántica”, noción que se remonta a los aportes realizados por Russell (1937) con su “teoría de clases y similitud”. Hoy en día, esta noción se implementa como medida probabilística o como el grado de intercambiabilidad de una palabra por otra en un contexto particular, dado que la similitud semántica se concibe bajo el supuesto de que palabras semánticamente similares se comportarán de manera similar

(Manning & Schütze, 2003; Matsumoto, 2003), utilizando para su aplicación un modelo vectorial denominado LSA³.

Una noción compatible de significado por asociación, aunque realizada desde una tradición antimentalista, es la planteada por los trabajos desarrollados por la escuela funcionalista inglesa en lingüística, a partir de la noción de significado colocacional (Palmer, 1980). Como sabemos, esta noción se basa en la teoría contextual del significado, en la cual una palabra adquiere significado por las palabras que la acompañan (Palmer, 1980; Stubbs, 1996, 2001). Bajo esta perspectiva se desarrolla toda la tradición lingüística orientada hacia el análisis de los textos, promovida por la escuela funcionalista firthiana (Stubbs, 1996). Los principios que guían esta escuela pueden sintetizarse, siguiendo a Stubbs (1996), de la siguiente manera:

- 1) La lingüística es esencialmente una ciencia social y una ciencia aplicada, que tiene implicaciones prácticas, especialmente en educación.
- 2) Los datos para el análisis lingüístico deben ser obtenidos del lenguaje en uso, esto implica que el lenguaje debe ser estudiado en situaciones auténticas de realización (no de manera intuitiva y con oraciones inventadas), en textos completos (no en oraciones aisladas o fragmentos textuales) y estos textos deben ser estudiados comparativamente a través de *corpora* textuales.
- 3) La preocupación esencial de la lingüística debe ser el estudio del significado, entendiendo que forma y significado son inseparables y que los estudios léxicos y gramaticales son interdependientes.
- 4) La naturaleza del comportamiento lingüístico involucra rutina y creación y el lenguaje en uso transmite cultura.

De estos principios nos interesa destacar el segundo y el tercero, ya que son los que más se relacionan con esta investigación. Estos son congruentes con nuestra idea de estudiar el lenguaje tal como es utilizado por los usuarios, en nuestro caso los científicos, y el análisis del significado, partir de las formas lingüísticas detectadas en textos completos, utilizando

³ Los conceptos de *similitud semántica* y LSA serán revisados en los apartados 2.7.2 y 2.7.2, respectivamente.

herramientas computacionales. Estas nociones básicas han dado origen a toda una corriente lingüística neofirthiana impulsada, principalmente por Halliday (1991, 1992) y Sinclair (1991), en la cual el uso de herramientas computacionales permite llevar a cabo estudios empíricos del significado utilizando grandes corpora textuales, la lingüística de corpus computacional. Respecto de esto último, Sinclair (1991: 6-7) plantea:

“[...] in all cases so far examined, each meaning can be associated with a distinct formal patterning[...]. There is ultimately no distinctions between form and meaning[...]. The meaning affects the structure and this is [...] the principal observation of corpus linguistics in the last decade”.

Esta idea, según Stubbs (1996), no ha sido del todo aceptada por la lingüística actual bajo el argumento de que el corpus no puede aportar al estudio del significado, siguiendo en muchos casos la clásica idea de que lo importante es la *langue* y no la *parole* (Saussure, 1983) o la competencia lingüística del hablante y no su actuación (Chomsky, 1957, 1965). Durante fines de la década del cincuenta y continuada hasta mediados de los noventa la hegemonía generativista, no solo desestimó el estudio del lenguaje a través de *corpora* de textos naturales, sino que también evadió un enfoque de dimensiones probabilísticas. Sin embargo, y como plantean Bod (2003) y Jurasky (2003) los enfoques probabilísticos y basados en corpus para la investigación científica del lenguaje humano no son incompatibles con una visión cognitiva del mismo. Por el contrario, los aportes del primero pueden ser de gran utilidad para la segunda y viceversa (Parodi, 2005b).

Esta focalización de la lingüística de corpus, planteada indudablemente como un resurgimiento de los estudios empiricistas, esta vez apoyados en una tecnología más poderosa y en el libre acceso a incontables textos en formato electrónico, permite ahondar en los estudios del lenguaje, y a través de este, acercarse a la comprensión de la mente humana, esto es, indagar en la naturaleza del lenguaje como manifestación de la mente (Chafe, 1994).

Con relación al estudio del significado, Stubbs (2001) plantea la posibilidad de estudiar las relaciones léxico-semánticas (ver 2.5), utilizando el enfoque de la lingüística de corpus en el cual a partir de estudios de colocaciones y frecuencias de palabras establece, basándose en ejemplos concretos, que los datos obtenidos de los corpus proveen evidencia respecto

del significado denotativo y connotativo. Así, apoyado en el principio de que el significado está en el uso, promueve los métodos observacionales de una semántica de corpus. En relación con estos métodos propone que la herramienta fundamental de la semántica de corpus es la concordancia, la cual permite estudiar en detalle una palabra objetivo (nodo) y las características de sus colocados. Esto último, unido a estudios de frecuencias en grandes cantidades de textos permite la construcción de patrones semánticos del uso de tales palabras (o grupos de palabras) (ver McEnery & Wilson, 1996; Stubbs, 2001; Biber, 2005).

Cabe hacer notar, que esta perspectiva de estudio del significado permite reconocer el significado desde los textos, lo cual se condice con la idea planteada en el Capítulo 1 respecto de la diferencia entre texto y discurso, entendiendo al texto como un nivel de representación lingüística externa y material de los procesos semántico-lingüísticos de los individuos, construidos sociocognitivamente en el seno de una comunidad comunicativa.

Comúnmente se han planteado los análisis de las frecuencias y correlaciones entre las palabras como un primer nivel de descripción de los textos (Manning & Schütze, 2003; Biber, 1988, 1998; McEnery & Wilson, 1996; Oakes, 1998; Manning y Schütze, 2003) en el cual lo importante es la descripción del comportamiento verbal en cuanto a la ocurrencia de las palabras en el texto. De este modo, los textos, con base en la cantidad de palabras totales, presentarán distintos tipo de palabras según los distintos tipos de frecuencias. En particular, las palabras de tipo funcional, sin mayor contenido semántico, son comúnmente las más frecuentes y las de contenido semántico aparecen como las menos frecuentes (Manning & Schütze, 2003). Esto es claramente explicado por la ley del menor esfuerzo lingüístico descrita por Zipf, en la cual pocas palabras se utilizan mucho, y que estas pocas palabras también tienen varias posibilidades significativas (Manning & Schütze, 2003, Sigurd, Eeg-Olofsson & van de Weijer, 2004).

El interés por el estudio de estas frecuencias comúnmente ha estado en reconocer el uso característico de las palabras, según los tipos de textos. Estos estudios de frecuencias se han preocupado por estudiar las comunalidades y especificidades de palabras en distintos tipos de textos, puesto que esto permite conocer el léxico compartido entre dos o más textos y, por otra parte, permite identificar palabras de uso terminológico, en particular en textos de las ciencias (Ventura, Hu, Graesser, Louwerse & Olney, 2004; Cademártori, Parodi &

Venegas, en prensa). Además, es posible, utilizando técnicas multifactoriales, agrupar variables (palabras o textos) características de un corpus. De este modo es posible observar cuáles son las variables de mayor importancia o características de los textos respecto de las cuales se agrupan y sus aplicaciones en estudios tanto textuales como psicolingüísticos (Baayen & Lieber, 1997; Lebart, Salem & Bécue, 2000; Hinto, Lupker & Pexman, 2002; van Overschelde, 2002; Sigurd et al., 2004; Rodríguez, 2004).

Sin embargo, respecto de estos métodos se ha criticado que la mayoría de los estudios de frecuencias en lingüística de corpus se han concentrado en descubrir cuáles son las frecuencias más altas de una unidad en un nivel lingüístico determinado (Oakes, 1998; Pérez, 2002; Manning & Schütze, 2003). Así, cuando este tipo de estudios se limita al recuento aislado de las unidades más frecuentes, oculta diversos aspectos interesantes que dicen relación con unidades de frecuencia nula, mínima o media. Al respecto, McEnery y Wilson (1996) reconocen que el análisis cuantitativo tiende a dejar de lado las ocurrencias muy escasas, lo que hace de este tipo de análisis muy apto para generalizaciones; pero, por lo mismo, tiende a perder riqueza en los datos.

Para evitar lo más posible esta pérdida de información aportada por las frecuencias más bajas o nulas, se han tomado en cuenta en los análisis las frecuencias inversas de los textos y los *hapax logomema* (frecuencia de palabra igual a 1). Existe una constante que surge al analizar las frecuencias de las formas de un corpus: sin importar cuán grande sea el corpus ni el tipo de documentos que contiene, la mitad de las formas de un corpus son *hapax logomema* (Rojo, 2002). Por otra parte, el cálculo de las frecuencias medias junto con el cálculo de las frecuencias más altas permite obtener un índice del vocabulario básico (*core vocabulary*) de una lengua o sublengua específica (McEnery & Wilson, 1996; Rojo, 2002). Este cálculo es de gran importancia al momento de diseñar métodos adecuados para la enseñanza de segundas lenguas (McEnery & Wilson, 1996; Rojo, 2002; Pérez, 2002).

Ahora bien, teniendo en cuenta los aspectos anteriores, podemos plantear que, para estudiar cuantitativamente relaciones léxico-semánticas en corpus, no podemos centrarnos solo en las más altas frecuencias sino en todo el rango de frecuencias de ocurrencias, e incluso más, para un estudio completo se requiere considerar, además, las co-ocurrencias

entre palabras y otras unidades textuales, así como las relaciones que aparecen debilitadas en por el tamaño de los corpus.

Dado lo anterior optamos por un método basado en un análisis semántico latente, que permite reconocer mediante medidas de vectores (o dimensionales), no solo las relaciones entre las palabras, sino que incluso las similitudes semánticas existentes entre las unidades lingüísticas o textuales a partir de las correlaciones entre, no solo, palabras, sino que entre palabras y documentos (entendidos como párrafos), generando un segundo nivel de descripción matemático (más abstracto), a través de procesos algorítmicos de reducción de dimensiones (sobre este método nos extenderemos en 2.7.3).

Para cerrar este apartado, volveremos a la idea inicial planteada en el apartado 2.3 en el cual se manifestó que el significado es la asociación entre la expresión lingüística y otro elemento. Esta asociación o relación semántica, como hemos visto, ha sido el objeto de estudio de buena parte de la actividad científica interesada en la comunicación lingüística, puesto que las relaciones semánticas contribuyen a definir la estructura semántica, que a su vez, sirve para configurar el sistema lingüístico.

Por otra parte, la actividad de los sujetos que consiste en manejar un sistema lingüístico implica también el manejo de las relaciones semánticas, pues estas organizan en buena medida el lexicón mental, intervienen como variables significativas en el procesamiento interno, facilitan la interdependencia con el sistema cognitivo y se usan permanentemente para producir y comprender locuciones o textos (Mayor, 2000). Como podemos percatarnos, a la base de esta preocupación por las relaciones semánticas no sólo está la comunicación lingüística, sino que también el interés científico por dilucidar la manera en que los seres humanos representamos conocimiento, proceso que se basa, sin duda alguna, en el descubrimiento y constatación de relaciones entre unidades del lenguaje.

2.5 Relaciones semánticas

Las relaciones semánticas son aquellas relaciones que, por una parte, implican la presencia del significado y que, por otra, se dan entre unidades lingüísticas en función de su significado (Mayor, 2000).

Para las primeras cabe interpretar las relaciones semánticas de dos formas diferentes. Una en la que el significado es el campo de la relación y otra en que las relaciones se dan entre elementos o componentes del significado. Un ejemplo para el primer tipo de relaciones es el análisis del significado que hacen Ogden y Richards (1984) en función de las relaciones entre el pensamiento (el concepto), el símbolo (la palabra) y el referente (la cosa). Para el segundo tipo, el ejemplo podría estar dado por las distintas teorías que asumen que el significado debe ser representado en términos de sus unidades mínimas. En esta línea destaca el análisis componencial del significado, que tiene por objeto especificar los diferentes rasgos o componentes semánticos que lo constituyen, así como las relaciones entre ellos. Este tipo de análisis ha sido llevado a cabo tanto en Europa (por ejemplo, Greimas, 1966; Hjelmslev, 1974; Pottier, 1983; Coseriu, 1987) con el nombre de análisis sémico, como en Norteamérica, destacándose aquí los trabajos de los semantistas generativos (por ejemplo, Lakoff, 1970; Katz & Fodor, 1976; Nida, 1975). Desde un punto de vista más actual, Escandell (2004) retoma este tipo de análisis siguiendo el enfoque formalista y utilizando constructos teóricos de la teoría de conjuntos. Otra propuesta muy desarrollada en esta área es la que tiene relación con los primitivos semánticos (Wierzbicka, 1992, 1997). La meta de esta investigadora es establecer un grupo completo y estable de primitivos semánticos a través de investigaciones de universales léxicos. La principal crítica al enfoque composicional del léxico es la de no precisar el modo de establecer cuáles son en definitiva los primitivos semánticos y determinar cómo operacionalizarlos y ponerlos en práctica (Mora, 2001).

Otra forma de interpretar las relaciones semánticas nos lleva a considerar la relación entre el significado y otro elemento, constituyendo entre ambos un campo más amplio que el del significado, pero que lo incluye. Tal es el caso de la concepción saussureana del signo lingüístico que se constituye como tal a partir de la relación existente entre significante y significado (Saussure, 1983). Un ejemplo de este tipo de análisis es el realizado por Pottier (1974) quien estudia las relaciones semánticas entre: un significante y varios significados (relación *homonímica*, *polisémica* y *metafórica*); entre un significado y varios significantes (relación de *polimorfismo* u *homosemia*, es decir, *sinonimia*, *metonimia* y *motivación del significante*) y relaciones por afinidad entre significados (*parasinonimia*) y entre significantes (*paronimia*).

Para la realización de los análisis de similitud semántica que nos proponemos en este trabajo, las relaciones semánticas que nos interesa son aquellas que se dan entre las palabras y los párrafos y entre párrafos, presentes en los textos, utilizando un método de cuantificación automático, que se basa solo en información estadística obtenida del texto. Esto nos obliga a centrarnos en las relaciones entre unidades portadoras de significado y no, exclusivamente, en las relaciones que se dan entre los componentes del signo lingüístico.

A continuación, revisaremos los diferentes tipos de relaciones semánticas en función de los principales tipos de unidades lingüísticas. Las primeras relaciones que analizaremos serán las que se dan entre unidades léxicas, las que en conjunto conforman el lexicón mental y/o los diccionarios.

Luego analizaremos las relaciones semánticas entre unidades léxicas, que asociadas entre sí de acuerdo con reglas sintácticas, constituyen enunciados o proposiciones.

Por último, consideraremos las relaciones semánticas entre enunciados o proposiciones, que combinados entre sí de acuerdo a reglas textuales, constituyen textos organizados, según una intención comunicativa enmarcada según las pautas impuestas por la comunidad discursiva en la cual se produce el texto.

2.5.1 Relaciones semánticas en el léxico

Según Mayor (2000), las primeras aproximaciones a una clasificación de las relaciones semánticas entre palabras procede de las investigaciones sobre la asociación de ideas, que se solapan y son continuadas por las llevadas a cabo sobre las asociaciones verbales.

Aristóteles (1966) ya había sugerido tres leyes de asociación de nuestros pensamientos y recuerdos: la *semejanza*, el *contraste* y la *contigüidad*, y una primera distinción entre ellas: las dos primeras se basarían en conexiones lógico-semánticas, y la tercera, en conexiones experienciales (Puente, 1998; Valiña & Martin, 2000; Mayor, 2000; Pozo, 2001).

Los asociacionistas ingleses, en particular Locke (1956), recogen estos principios de asociación y sugieren que estos se basan en las relaciones que existen entre cosas, entre

ideas y entre palabras. Las relaciones más básicas son las de *causa – efecto*, *las temporales*, *las espaciales*, *las de identidad* y *las de diversidad*.

La tradición asociacionista siguió por este camino, en especial desde la psicología. En los albores de esta disciplina, Wundt (1893) clasificó las asociaciones en *internas* y *externas*. Las internas (conexiones que surgen del interior del sujeto) se dividen en tres tipos. El primer tipo incluye la *supraordenación*, *subordinación* y *la coordinación*, el segundo considera *la identidad* y *el contraste*, y el tercero incluye *la causalidad* y *la finalidad*. Las externas (que vienen dadas desde el mundo exterior) se dividen en *simultáneas* (de la parte al todo, del todo a la parte y de coexistencia) y *sucesivas* (auditivas y visuales). Estas relaciones, si bien en su momento no son consideradas semánticas, constituyen una clasificación de relaciones semánticas (Mayor, 2000).

Por otra parte, en los inicios de la lingüística moderna una clasificación importante propuesta por Saussure (1983) y desarrollada posteriormente por los estructuralistas (por ejemplo, Hjelmslev, 1974) y los funcionalistas (por ejemplo, Jakobson & Halle, 1956; Martinet, 1965, 1971) es la que distingue entre asociaciones (o relaciones) *in absentia* e *in praesentia*, es decir, relaciones paradigmáticas y sintagmáticas.

Otra aproximación al estudio de las relaciones semánticas y a su clasificación es la que se lleva a cabo desde las investigaciones sobre la organización de la memoria y el procesamiento del léxico. Los estudios de memoria, en especial de la memoria semántica y las representaciones, no han estado exentos de controversias. Como plantea Pozo (2001), el origen de esta controversia está en la adopción en las ciencias cognitivas de la teoría de la información, ya que esta alberga la posibilidad de considerar que la memoria (y por ello la mente) se caracterizaría como un sistema de cómputo de representaciones lógicas, abstractas, basada en proposiciones, o como el resultado parcial o total de la interconexión de una serie de nodos o unidades de procesamiento a través de diferentes tipos de relaciones semánticas en la cual las representaciones son difusas y están distribuidas (Rumelhart, McClelland & el grupo PDP, 1992).

Otros autores utilizan los casos propuestos por Fillmore (1971) como criterio para clasificar las relaciones semánticas que estructuran la memoria; así, por ejemplo, Lindsay y Norman

(1972) ofrecen una lista completa de los casos utilizados para describir acontecimientos: *acción, agente, condicional, instrumento, ubicación, objeto, propósito, cualidad, receptor, tiempo, verdad*. Otros autores se centran solo en algunas relaciones semánticas a las que atribuyen un valor decisivo para organizar la memoria. Collins y Quillian (1969) incluyen una red de relaciones semánticas y unos mecanismos que operan sobre esa red; proponen relaciones como la *subordinación*, la *supraordenación*, la *conjunción*, la *disyunción*, la *determinación*, etc., aunque en los experimentos sólo utilizan redes categoriales y relaciones jerárquicas.

Una alternativa a la teoría de redes jerárquicas es la propuesta por Collins y Loftus (1975). La idea básica es que la estructura de la memoria semántica está estrechamente interconectada con relaciones que van en múltiples direcciones, lo que permitiría explicar cómo ocurren las inferencias. Esta estructura se fundamenta en el nivel de probabilidad de las conexiones, según estas hayan sido previamente activadas o no, y en la distancia semántica entre los conceptos. Se retoma el principio fundamental de la propagación de la activación. Esto se explica porque el origen de la activación ocurre cada vez que algún estímulo externo o interno excita un concepto. Este concepto activado irradia su acción multidireccionalmente hacia otros conceptos próximos, y estos, a su vez, hacia los más distantes. Esta propagación se caracteriza por decrecer con el paso del tiempo y en función de la distancia entre los conceptos, por ser secuencial (se activa un concepto a la vez) y por verse facilitada en virtud de factores como la semejanza de los conceptos, la frecuencia de uso y el número de conexiones que posea cada concepto (Puente, 1998).

En este contexto, las relaciones semánticas juegan un rol muy importante en el proceso de comprensión de los enunciados, es por ello que los estudios sobre acceso al léxico y reconocimiento de palabras, así como sobre los efectos de facilitación e interferencia del contexto, ponen de relieve la importancia de las relaciones semánticas como variables del procesamiento cognitivo en la comprensión. En este contexto de investigación se han utilizado relaciones de *semejanza* y *antonimia* y relaciones *taxonómicas* y *partonímicas* (Mayor, 2000).

En los estudios sobre la estructura del significado y la semántica léxica el tema de las relaciones semánticas ha sido ampliamente trabajado. Destaca aquí el trabajo de Cruse

(1986), quien utiliza la intersección de conjuntos para clasificar las relaciones semánticas entre unidades léxicas: *la identidad* (dando lugar a la *sinonimia*), *la inclusión* (dando lugar a la *hiponimia*), *el solapamiento* (dando lugar a la *compatibilidad*), y *la disyunción* (dando lugar a la *incompatibilidad*). A su vez, analiza las configuraciones léxicas en cuyo seno se dan las relaciones semánticas, que también se clasifican en cuatro grupos: *jerarquías* (que son ramificadas, como las *taxonomías* y las *meronomías*), las *series proporcionales* (que dan lugar a las *endonimias*, las *analogías*, el *parentesco lexical* y las *paronimias*), los *dobletes* (pares opuestos que pueden ser *antónimos*, *complementarios* y *direccionales*) y los *agrupamientos o clusters* (a los que pertenecen los *sinónimos* y los *plesiónimos*).

Otra taxonomía de las relaciones semánticas es la propuesta por Winston, Chaffin y Hermann (1987). Esta es construida a partir del resultado de indagaciones críticas y empíricas acerca de la identificación de los diferentes tipos de relaciones semánticas y de indagaciones sobre la percepción de semejanzas entre esas relaciones. Las agrupaciones encontradas fueron reunidas en cinco grandes tipos de relaciones a los que denominan “familias de relación”: 1) de *contrastes* (que incluyen relaciones *contrarias*, *contradictorias*, *inversas*, *direccionales*, *incompatibles*, *contrarias asimétricas*, *pseudoantónimas*, y de *atributos semejantes*); 2) de *semejanza* (que incluye relaciones de *sinonimia*, de *dimensión semejante*, de *atributo necesario*, de *atributo invitado* y de *acción subordinada*); 3) de *inclusión de clases* (que abarca las relaciones *subordinadas perceptuales*, *subordinadas funcionales*, *subordinadas de estado*, *subordinadas de actividad*, *subordinadas geográficas*, de *lugares*); 4) de *caso* (entre las que destacan las relaciones *agente-acción*, *agente-instrumento*, *agente-objeto*, *acción-receptor* y *acción-instrumento*); y 5) de *parte-todo* (que incluye relaciones de *objeto funcional*, de *colección*, de *grupo*, de *ingrediente*, de *localización funcional*, de *organización*, y de *medida*).

Un acercamiento más actualizado de estas relaciones semánticas lo propone Mayor (2000) quien plantea dos grandes categorías: *las relaciones estructurales* que se basan en la estructura de los dos significados y las *relaciones funcionales*, en las que uno de los dos significados está en función del otro. Las relaciones estructurales son: *taxonómicas* (hiponímicas, hiperonímicas y cohiponímicas), *partonómicas* (metonímicas, holonímicas, comeronímicas), de *orden* (espacio-temporales, experienciales, abstractas), de *semejanza*

(sinonimia, analogía) y de *oposición* (antonímicas, inversas, recíprocas, direccionales, contradictorias). Las relaciones funcionales son: *causales* (de causalidad eficiente, antecedente-consecuente, finales), *predicativas* (agente-acción, agente-objeto, agente-beneficiario, agente-instrumento, agente-circunstancia, acción-agente, acción-objeto, etc.), *atributivas*, de *pertenencia*, de *implicación*, de *contigüidad* (espacial, temporal, experiencial) y *sintagmáticas*.

2.5.2 Relaciones semánticas en el enunciado.

En la gramática de casos de Fillmore (1971) se toman en cuenta las relaciones semánticas que existen entre complejos nominales y verbos. El propósito de Fillmore (1971), como el de los semánticos generativos, era corregir algunas de las limitaciones de la gramática de Chomsky (1957, 1965), específicamente en lo que al rol del significado respecta. Recordemos que en los modelos clásicos el componente semántico, que relaciona la estructura generada por el componente sintáctico a ciertas representaciones semánticas, era considerado como puramente interpretativo. Este componente utiliza la información proveniente del componente sintáctico en lo concerniente a los formativos, sus propiedades inherentes, y sus interrelaciones en una oración dada. Consecuentemente, en cada oración, el componente sintáctico debe especificar, para cada oración, una estructura profunda que determina su interpretación semántica y una estructura de superficie que determina su interpretación fonética.

En este enfoque más clásico de la lingüística generativa, por lo tanto, el rol de la semántica es puramente interpretativo y no generativo (Tognini-Bonelli, 2001) o como lo refiere Lyons (1980), diciendo que en tanto la buena formación de la oración sea alcanzada, la semántica es residual.

Fillmore (1971) se inspira en la lógica de predicados para afirmar que una proposición consta de un predicado y de uno o más argumentos, siendo críticas las relaciones existentes entre cada argumento y el predicado. Según su hipótesis, el número de papeles que desempeñan los argumentos respecto del predicado es muy limitado, y estos roles equivalen a conceptos mentales universales de carácter innato. Los casos propuestos son: *agente* (argumento que especifica la causante del suceso descrito por el predicado),

experimentador (argumento cuyo referente [vivo] experimenta lo que expresa el verbo), *instrumento* (argumento que expresa la causa inmediata o el estímulo del suceso descrito en el predicado), *objeto* (argumento cuyo referente experimenta el cambio expresado en el predicado), *fuentes* (argumento que caracteriza el punto de partida del suceso expresado en el predicado), *fin* (argumento que caracteriza el punto final del suceso expresado en el predicado), *locación* (argumento que caracteriza el lugar del suceso expresado por el predicado) y *tiempo* (argumento que caracteriza el momento en el que ocurre el suceso expresado en el predicado).

Desde la perspectiva de la semántica generativa, se entiende que el significado gramatical puede describirse en términos de las relaciones semánticas entre los constituyentes de la proposición, a los que Nida (1975) denomina “relaciones semánticas primarias”. Estas se pueden agrupar en seis diferentes configuraciones a) *acción*: Esta configuración incluye eventos y elementos que los acompañan, sin que se pueda establecer con precisión cuántas relaciones deben ser consideradas como descriptoras de eventos y cuántas reconocidas como tipo de relación de caso -relaciones entre elementos y eventos-; entre las más representativas están las relaciones *agente-acción*, *evento-recipiente*, *evento-beneficiario*, *instrumento-evento*, *sustancia-evento*), b) *descripción*: Esta consta de entidades, eventos o abstracciones a los que se atribuyen constituyentes cualificadores o modificadores, c) *asociación*: Esta incluye entidades que pueden estar en relación *serial*, *alternativa* o *posesiva*, d) *clasificación*: Esta reúne a entidades que se clasifican como pertenecientes a una clase de entidades más amplia, e) *identificación*: Esta consta de dos diferentes designaciones para el mismo referente, f) *configuración axial*: Esta incluye las relaciones entre entidades o eventos y preposiciones y conjunciones; si implica entidades expresa *posición*, *movimiento* y *agencia*; si implica eventos expresa *tiempo*, *razón*, *propósito*, *concesión* y *medios*.

Un último ejemplo de relaciones semánticas a nivel del enunciado es el que sugiere Crombie (1985). Este autor propone clasificar las diferentes relaciones que se dan entre argumentos y predicados, también llamados roles o casos, guiado por la lógica de la coherencia intraproposicional. La clasificación de estas relaciones se inicia con la

clasificación de los tipos de predicados semánticos. Estos pueden ser verbales o nominales y adjetivales.

Los predicados verbales son de tres tipos: a) *dinámicos*, es decir, implican actividad física o mental y pueden referirse a una actividad general, a una actividad momentánea, a un evento transicional, a una actividad mental o a un factitivo, b) *de proceso*, esto es, implican procesos en los que no hay actividad consciente, y c) *de estado*, aquellos que implican percepción y cognición o relaciones. Los predicados nominales o adjetivales son los que complementan una cópula y se dividen en: *estados* materiales y experienciales y *procesos* materiales y experienciales.

La forma en que los argumentos funcionan dentro de las proposiciones depende, en gran medida, de los tipos de predicados semánticos a los que se relacionan. Estos roles o casos semánticos se pueden clasificar en cinco tipos principales: a) *roles causales*, que implica una acción que causa un efecto; las relaciones de los argumentos con estas acciones pueden ser de varias clases: de *agente* (un ser vivo que lleva a cabo una acción), de *instrumento* (una entidad por medio de la cual se lleva a cabo la acción) y de *fuerza* (algo no vivo que causa algo sin intervención de un agente); b) *roles de participación*, que incluyen cualquier intervención o participación no causal de una entidad o abstracción en una actividad, proceso o estado, pudiéndose clasificar en *roles de participación en actividades* -con algunos tipos de predicados dinámicos- (paciente y asignado), *roles factitivos* -con predicados factitivos- (material y resultado), *roles de participación en procesos* -con predicados de proceso- (mutante) y *roles de participación en estados* -con predicados de estado- (lo que se mantiene, el que experimenta, lo experimentado); c) *roles de transición*, que conciernen a movimientos o localización y que pueden clasificarse en *objeto* (entidad en movimiento o localizada), *fuerza* (localización de una entidad al comienzo del movimiento), *fin* (localización de una entidad al final del movimiento) y *campo* (localización de una entidad y camino o área atravesados por una entidad móvil); d) *roles relacionales*, que unen argumentos con predicados relacionales y que pueden ser el *referido* (entidad viva a que está ligado el predicado relacional), el *referente* (entidad viva o no viva a la que el referido está ligado a través del predicado relacional) y la *cantidad* (que expresa la medida en la que una entidad está ligada a un predicado relacional); y e) *roles abaxiales*,

que expresan la entidad viva afectada por una acción o proceso, pero que no está directamente implicada en esa acción o proceso.

Se debe destacar que estos tipos de relaciones semánticas son utilizados para la representación simbólica de conocimiento. En estos, el significado de las relaciones está estrechamente unido a las configuraciones sintácticas. Más adelante, veremos que son estos tipos de relaciones semánticas los que permiten una representación computacional del significado basado en lenguajes representacionales.

2.5.3 Relaciones semánticas en el texto

El estudio de las relaciones semánticas entre las diversas unidades textuales, en particular entre las proposiciones que las constituyen, ha sido abordado desde distintos enfoques, puesto que es esencial en la determinación de estructuras textuales (Mayor, 2000).

Desde la perspectiva de la poética o retórica de corte estructuralista, representada por Todorov (1975), ya se suponía que todo texto era descomponible en unidades mínimas y que el tipo de relaciones entre estas unidades servirían para distinguir entre varias estructuras textuales. De esta manera, las relaciones semánticas entre las unidades pueden clasificarse en tres órdenes: *lógico* (causalidad, disyunción, conjunción, exclusión e inclusión), *temporal* (digresión, inversión y visión estereoscópica) y *espacial* (simetría, gradaciones, paralelismo y antítesis).

Por otra parte, Rumelhart (1975) proponía junto a una estructura sintáctica del relato una estructura semántica basada en una serie de relaciones semánticas, entre las que destacan la *conjunción* (*Y*), un simple predicado conjuntivo de cualquier número de argumentos; la *autorización* (permitir), que relaciona un evento que hace posible un segundo evento, pero no lo causa directamente; la *iniciación* (iniciar), que relaciona un evento externo con la reacción voluntaria de un ser consciente a ese evento; la *motivación* (motivar), que relaciona una respuesta interna a una acción resultante de ella; la *causa* (causar), que relaciona dos eventos, el primero de los cuales causa el segundo; y la *secuenciación* (después), que relaciona una secuencia temporal de eventos.

Más tarde, en el marco de las primeras investigaciones psicolingüísticas respecto de la comprensión, Kintsch y van Dijk (1978) exponen su primer modelo para el procesamiento cognitivo del texto escrito. Este modelo opera en el ámbito de las estructuras semánticas subyacentes en términos de proposiciones y de relaciones entre proposiciones. Plantean que existen dos niveles de representación cognitiva del texto, a saber, el código de superficie y la base textual. El primero de ellos constituye el nivel *verbatim* o de superficie, siendo estrictamente lingüístico. El segundo corresponde a la estructura semántica que construye el lector y que puede ser descrita en dos niveles, en un nivel de microestructura y en uno de macroestructura. La microestructura es el nivel local del discurso, es decir, la estructura de las proposiciones individuales y sus relaciones. La macroestructura es de naturaleza más global y describe el contenido del discurso⁴ como un todo. Estos niveles se relacionan por medio de macrorreglas que son recursivas. Las macrorreglas que distinguen son las de supresión, generalización y construcción, teniendo estas por función relacionar las secuencias de proposiciones de la microestructura con secuencias de proposiciones en un nivel superior, por medio de la supresión o generalización de todas las proposiciones que son irrelevantes o redundantes y por la construcción inferencial de nuevas proposiciones. Se obtiene de este modo un orden jerárquico en el que cada nivel superior corresponde a integraciones de los niveles inferiores. Así, un discurso es coherente para un lector solo si sus oraciones y proposiciones respectivas están relacionadas y organizadas globalmente al nivel de la macroestructura.

Según Kintsch y van Dijk (1978), un texto es comprendido cuando su contenido es representado en la estructura de conocimiento semántico del lector, esto es cuando el lector construye la base textual a partir del establecimiento de coherencia entre las oraciones (coherencia local) y entre el contenido de todos los párrafos del texto (coherencia global).

Más tarde, van Dijk y Kintsch (1983) afirman que el modelo de 1978 puede considerarse como un modelo subespecífico y lo consideran una versión simplificada del modelo de 1983. Este último modelo es considerado como un modelo estratégico proposicional en el cual la comprensión de un texto involucra la construcción de una representación semántica

⁴ Kintsch y Van Dijk (1978) utilizan los conceptos de texto y discurso indistintamente.

de los contenidos del texto, constituida por varios niveles y a través de la acción de varios procesos estratégicos, en los cuales se reduce y organiza la información que se está interpretando. Según estos modelos el significado de las palabras de la oración forma la microestructura de un texto. Estas microestructuras se organizan en conjuntos de representaciones proposicionales interrelacionadas de las frases, cláusulas y oraciones del texto. Las macrorreglas traducen estas secuencias de proposiciones a conjuntos menores de proposiciones más generales, eliminando las proposiciones que son menos importantes para el significado global del texto, generalizando proposiciones en conjuntos mayores y construyendo nuevas unidades textuales que reemplazan el significado del antiguo conjunto. Las macrorreglas operan en forma recursiva de modo que las macroestructuras que se forman mediante las macrorreglas pueden ser sometidas a un nuevo ciclo de macrorreglas, generalizando aún más el meollo del texto. Por tanto, las macroestructuras son descripciones semánticas abstractas del contenido semántico del texto, similar al tema y al significado global, las que proveen coherencia global. Debido a que las macroestructuras son descripciones semánticas abstractas del texto de superficie, que operan a un nivel de abstracción más alto que la microestructura, es probable que no siempre estén señalizadas en el texto. Los casos en los que están expresadas corresponden a los títulos y subtítulos de los textos, y también al resumen del contenido. Otro buen indicador de la marcoestructura es la oración tópica al comienzo o al final de un párrafo (van Dijk & Kintsch, 1983).

Las bases textuales de los dos modelos propuestos por Kintsch y van Dijk (1978) y van Dijk y Kintsch (1983) son representaciones proposicionales. Los autores complementan su modelo con la teoría de los esquemas, por cuanto plantean que las macrorreglas de supresión, generalización y construcción se aplican bajo el control de un esquema que limita su operación de manera que las macroestructuras no lleguen a ser abstracciones o generalizaciones sin significado. A partir del esquema, el lector determina estratégicamente qué microproposiciones o generalizaciones de microproposiciones son pertinentes y, de este modo, qué partes del texto constituirán el meollo o su esencia (van Dijk & Kintsch, 1983). Los esquemas además de suministrar un marco coherente para interpretar las unidades semánticas del texto, también proporcionan una base para los procesos más activos de arriba-hacia-abajo (*top down*). Se puede prescindir de la información que falta si se considera insignificante o, de lo contrario, se la puede buscar activamente en el texto.

Las desviaciones con respecto al esquema pueden ser aceptadas y registradas, o si son muy importantes, pueden ser el punto de partida para hacer un esfuerzo en buscar soluciones que traten de explicarlas.

Van Dijk y Kintsch (1983) identifican problemas sin resolver en relación con los esquemas, en particular en lo que se refiere a la adaptación del esquema a un contexto. Por esta razón, en el modelo de 1983, introducen un tercer nivel de representación denominado *modelo de situación*. Mientras que los esquemas o guiones se refieren a sucesos y episodios estereotípicos o prototípicos (Schank & Abelson, 1977) y, por ende, representan información general y abstracta, los modelos de situación representan sucesos específicos y concretos y, por tanto, enfatizan constantes que caracterizan personas, lugares o tiempos. Ya no se representará el texto en sí sino la situación tratada por el texto. En síntesis, mientras en un nivel de representación, la base textual (el conocimiento del texto) es construida, en otro, el modelo de situación incorpora el conocimiento de mundo aportado por el texto.

Van Dijk y Kintsch (1983) afirman que el uso del conocimiento en la comprensión del discurso significa ser capaz de relacionar el discurso con alguna estructura existente de conocimiento que proporciona el modelo de situación para esto. Si no somos capaces de imaginar una situación en la que ciertos individuos tienen las propiedades o relaciones indicadas en el texto, no podemos comprender el texto. Si no captamos cognitivamente las relaciones entre los hechos locales y los hechos globales a los que se refiere el texto, tampoco lo entendemos. De este modo, al leer no sólo construimos una representación semántica del texto, sino también una representación mental de un suceso o de una situación. Este modelo de situación, como lo entienden van Dijk y Kintsch (1983), van Dijk (1995) y Kintsch (1988, 1998), corresponde a una imagen mental o a un esquema mental de la porción del mundo a la cual se hace referencia (persona, evento, situación, etc.), aunque muchos de los elementos que allí intervengan no aparezcan en el texto, ya sea porque la información está implícita o porque se presupone, los lectores la completan con el conocimiento que poseen en el modelo de situación. Este está formado por toda la información integrada a los conocimientos generales, sociales y culturales del lector.

En síntesis, el modelo propuesto por van Dijk y Kintsch (1983) aborda dos de los tres temas que según Sanders y Spooren (2001) son los de mayor relevancia en los estudios de la representación del texto, a saber: representación multiniveles, subespecificación de las representaciones y representaciones dinámicas.

El primero es la consideración de la idea de múltiples niveles de representaciones del texto, los cuales son el código de superficie (una representación efímera del material lingüístico de las oraciones tal como se presenta en el texto), la base del texto (que contiene las proposiciones expresadas por las oraciones y sus interrelaciones) y el modelo de situación (en el cual el material lingüístico es integrado con el conocimiento previo del lector).

Además, podemos notar que las expresiones lingüísticas constituyentes del texto, en estos modelos cognitivos, se consideran como instrucciones o pistas para actualizar una representación mental del texto que está basada en discursos previos, conocimientos previos e inferencias realizadas por el lector, esto es lo que Sanders y Spooren (2001) denominan la *subespecificación de las representaciones*. Al respecto afirman que:

“Contrary to what is maintained in the standard coding theory of meaning, it is fairly generally accepted in many branches of linguistics and psychology that what an utterance means cannot in any easy, transparent and compositional way be connected to the meaning of the individual elements in the utterance and their interrelations. An utterance explicitly codes only part of the meaning of the utterance into explicit linguistic material, the rest having to be provided by inferencing” (Sanders & Spooren, 2001: 4).

Esta idea de que la suma de las partes y de sus relaciones no conforman completamente el significado del enunciado es un tema que ha sido tratado ampliamente en lingüística (por ejemplo, Grice, 1975; Sperber & Wilson, 1992; Fauconier, 1994; Pomerantz & Fehr, 1997) y en psicolingüística (por ejemplo, Kintsch, 1988, 1998; Gernsbacher, 1990, 1996; Graesser, Singer & Trabasso, 1994; Louwerse, 2004).

Cabe hacer notar que el tema de la subespecificación de las representaciones se funda en el supuesto de que los sujetos construyen el significado del texto gracias a que son capaces construir un todo coherente a partir de lo manifestado materialmente en el texto. Esta búsqueda constante de coherencia, o lo que es lo mismo, el establecimiento de relaciones

semánticas en la mente del lector, es la que permite formar las representaciones mentales de un texto dado.

Recordemos que, como ya mencionáramos en el capítulo primero, en este trabajo los conceptos de texto y discurso no se consideran como equivalentes, ya que el texto es considerado como una instancia particular de la actividad discursiva, esto es, como la representación lingüística externa material de un proceso semántico-lingüístico (*érgon*, en palabras de Coseriu, 1991) en el cual se organiza la información, integrando unidades lingüísticas de los distintos niveles del lenguaje (léxicas, sintácticas y semánticas) que un escritor pretende dar a conocer al lector.

Ahora bien, definir el texto como un producto lingüístico, que representa procesos de semánticos implica usar las nociones de cohesión y coherencia. De este modo, para algunos autores el rasgo semántico característico de los textos es la coherencia; la secuencia de oraciones correspondería a la actualización de una secuencia de proposiciones que, de ser coherentes entre sí, con el todo y el contexto, constituiría la unidad texto (van Dijk & Kintsch, 1983; Peronard, 1998; Sanders & Spooren, 2001). Cabe señalar aquí que esta coherencia puede ser local o global (van Dijk & Kintsch, 1983), externa o interna (Peronard, 1998), referencial o relacional (Sanders & Spooren, 2001).

En cuanto a la coherencia textual local y global, van Dijk y Kintsch (1983: 15), señalan para la primera que:

“The main abstract condition on local coherence is that the complex propositions, expressed by the respective clauses or sentences, denote *facts* of some possible world that are related, conditionally or by inclusion”.

Sin embargo, este nivel de coherencia no es suficiente, según van Dijk y Kintsch (1983), puesto que para que una secuencia de proposiciones sea un texto es necesario que también sea coherente a nivel global, es decir, que el conjunto de los significados abstractos que engloban los significados de secuencias parciales de proposiciones, cumplan las mismas condiciones señaladas para la coherencia local. El conjunto de estas macroproposiciones constituye una macroestructura que corresponde al asunto, tema o meollo de un texto (En el apartado 2.5.3 se desarrolla más extensamente este tema).

En cuanto a la coherencia externa e interna, la primera está orientada a explicar aquellos textos muy breves cuyo significado sólo se percibe en relación con el contexto comunicativo en el cual se emiten. En este caso, el sentido del texto resulta de la conjunción de una expresión con el marco social en el cual se da. En cuanto a la coherencia interna, según Peronard (1998), existe consenso acerca de que una secuencia de oraciones es un texto si las proposiciones que ella expresa están relacionadas directa o indirectamente, formando un todo y/o si los hechos y entes del mundo al cual hacen referencia se consideran relacionados entre sí; es decir, se considera coherente a una secuencia de oraciones si tanto sus significados intensionales (proposiciones) como sus significados extensionales (referentes) forman totalidades conexas (Peronard, 1998).

Por último, en cuanto a la coherencia referencial y relacional, Sanders y Spooren (2001: 7) plantean que:

“Generally speaking, there are two respects in which a texts can cohere: a) Referential coherence: units are connected by repeated reference to the same object; b) Relational coherence: text segments are connected by establishing coherence relations like cause-consequence between them”.

Para el primer tipo de coherencia, los autores destacan la relevancia de indicadores lingüísticos como los pronombres y otros dispositivos para la referencia anafórica. Particularmente, argumentan respecto de la importancia que tiene la gramática, entendida desde un punto de vista funcional y cognitivo, en las operaciones mentales para la conexión de la información nueva con las representaciones mentales existentes. En cuanto a la coherencia relacional, los autores la definen como la relación de significado por la cual se conectan dos segmentos de texto. Ejemplos de relaciones de este tipo serían: causa-consecuencia, listados y problema-solución. Agregan que estas relaciones son conceptuales y que ellas pueden estar, aunque no necesariamente, realizadas explícitamente por marcadores lingüísticos.

Halliday y Hasan (1976), usan solo el término cohesión para referirse a las relaciones de significado que existe al interior del texto. Para estos últimos autores un texto se define por su “textura”, es decir, la textura es la que permite distinguir entre un texto y un “no texto”. Ahora bien esta textura se construye a partir de relaciones cohesivas entre las oraciones. A

cada una de estas relaciones se les denomina *ties* (ligazones) y se clasifican en diferentes tipos. Así, encontramos relaciones cohesivas por referencia, substitución, elipsis, conjunción y léxicas. A la unión de estas relaciones entre las distintas oraciones del texto y de los párrafos se le denomina cohesión, es decir, a las relaciones de significado que existen al interior del texto y que lo conforman como tal. Cabe hacer notar que Halliday y Hasan (1976) distinguen entre cohesión gramatical y cohesión léxica. Finalmente, otros describen la coherencia como coherencia semántica y cohesión como la manifestación gramatical de la coherencia subyacente (van Dijk & Kintsch, 1983).

De Beaugrande y Dressler (1997) plantean un uso de los conceptos de cohesión y coherencia bastante distintos. Para ellos, la cohesión consiste en la interrelación entre secuencias oracionales que componen la superficie textual, a través de relaciones gramaticales, como la repetición, las formas pronominales, la correferencia, la elisión o la conexión. Junto a esto, De Beaugrande y Dressler (1997) demuestran, mediante un sistema de análisis dinámico, basado en redes cognitivas, que la cohesión, en contra de lo que planteaban Halliday y Hasan (1976), no consiste únicamente en un conjunto de relaciones superficiales que interconecta gramaticalmente los diversos componentes textuales, sino que cada elemento lingüístico dirige y mediatiza la operación de acceso a otros elementos lingüísticos con los que se interrelaciona. En este sentido para De Beaugrande y Dressler (1997) la cohesión funciona asegurando que se mantenga activada en la memoria la información relevante, de manera que cuando aparezca, por ejemplo, un elemento pronominal, el receptor no tenga que trasladar su atención hacia atrás en el texto, sino que bastará con que recupere en su memoria activa tal información. En cuanto a la coherencia, plantean que un texto posee coherencia cuando los conceptos que componen su universo del discurso están interconectados a través de relaciones de diversa naturaleza, por ejemplo, la causalidad. En suma, los autores proponen que la cohesión y la coherencia son, por un lado, restricciones inscritas en el texto por el productor, cuya función es orientar los procesos cognitivos interpretativos que han de ponerse en funcionamiento en los lectores; y, por otro, propiedades constituyentes de la actividad interpretativa. Esto permitiría, según De Beaugrande y Dressler (1997), explicar el hecho de que un texto con imperfecciones formales, que presente deterioro más o menos grave en su cohesión o en su coherencia, pueda ser interpretado por los receptores textuales.

Estas nociones de cohesión y coherencia son complementarias con perspectivas psicolingüísticas más actuales y más acotadas a los ámbitos de análisis de textos desde perspectivas más computacionales.

Louwerse (2004), así como ya también lo esbozaba Bernárdez (1995), usa el término coherencia para las relaciones representacionales y el de cohesión para las indicaciones textuales a partir de las que debería construirse una representación coherente. De este modo, *“la cohesión es la consistencia de los elementos en el texto y la coherencia es la consistencia de los elementos en la representación en la mente del comprendedor”* (Louwerse, 2004: 42). Para este autor, la cohesión no es ni suficiente ni necesaria para la coherencia, esto quiere decir que la cohesión no siempre garantiza la coherencia mientras que la ausencia de la cohesión no impide la coherencia.

Indudablemente, la coherencia y la cohesión están interrelacionadas. En este sentido, el mismo Louwerse (2004: 42) plantea:

“La cohesión, sin duda, solo puede ser interpretada (como coherencia), pero de igual manera, la coherencia emerge debido a los lazos cohesivos en el texto. Sin embargo, la distinción entre ambas es importante para numerosos enfoques de lingüística computacional, lingüística de corpus y enfoques psicolingüísticos”.

Según Louwerse (2004), muchos estudios psicolingüísticos han mostrado que, en general, la cohesión facilita la coherencia. Sin embargo, la literatura también ha encontrado una interacción entre cohesión y coherencia, al mostrar que la cohesión no siempre lleva a beneficios para la comprensión. En los estudios realizados por McNamara (2004) se ha puesto de manifiesto la fuerte interacción entre cohesión y coherencia tanto en lectores con bajo conocimiento previo como en los de alto conocimiento previo. En una serie de experimentos realizados, concluyeron que los lectores con bajo conocimiento se beneficiaban más de textos muy cohesivos, tanto en tareas de comprensión como en las de recuerdo, pero los lectores que tienen gran conocimiento previo obtienen mejores logros en los textos poco cohesivos. Esto se puede explicar suponiendo que los lectores de gran conocimiento se ven inducidos a leer el texto en forma más activa, eliminando los vacíos cohesivos. Los lectores que poseen un alto conocimiento del tema son capaces de hacerlo,

porque poseen la información previa necesaria. Los lectores que tienen poca información, por otra parte, necesitan las pistas cohesivas para construir una representación mental coherente.

En función de lo anterior es posible concluir que un texto como entidad material está conformado por unidades lingüísticas que representan las relaciones semánticas por medio de las relaciones de cohesión que sustentan la conectividad del texto: la *temporalidad* (antes, después, durante, y, entonces), la *causalidad* (porque, así), la *intencionalidad* (para, con el propósito de), la *oposición* (pero, mas, sin embargo), la *implicación lógica* (por consiguiente, por lo tanto, entonces) y que lo constituyen en una unidad semántica (Graesser, Wiemer-Hasting & Wiemer-Hasting, 2001). Ahora, en cuanto representación cognitiva, el texto exige del lector construir relaciones de coherencia que son guiadas tanto por las marcas de relaciones cohesivas, a nivel del léxico y de la gramática (si las hubiere) como por los conocimientos del lector que le permiten reconocer esas relaciones o agregar relaciones propias, es decir, producir inferencias que son activadas durante y después del procesamiento cognitivo del texto, construyendo finalmente una representación mental coherente a partir de lo leído.

Por último, si las representaciones mentales de un texto se construyen como lo hemos planteado más arriba, deben tenerse en cuenta las relaciones de cohesión (temporal, causal, de intencionalidad, de oposición, de implicación lógica, referencial, de sustitución, de elipsis, de conjunción y léxicas), así como los distintos tipos de inferencias que pueden permitir la construcción de coherencia de un texto (por ejemplo, inferencias “puente”, referenciales, causales, elaborativas, temáticas, etc.) (ver Graesser, Singer & Trabasso, 1994; León, 2003; Parodi, 2005c).

Como podemos observar, tanto la noción de cohesión como la de coherencia, así como el estatus del texto como representación externa de los procesos semántico-lingüísticos que se dan en la mente de los individuos, son temas que aún no se zanján completamente desde las diversas disciplinas en las cuales el texto y el discurso son objetos de estudio. Sin embargo, es posible percatarse que actualmente son las aproximaciones psicolingüísticas la que tienden a proponer mayores avances en estas temáticas.

Un tercer tema comúnmente tratado en los estudios de representación del texto se refiere a las representaciones dinámicas. Para Sanders y Spooren (2001:5) el que un texto sea representado dinámicamente significa que:

“The effect of a language element on a representation is dependent on the current state of that representation, which is updated incrementally.”

Según Sanders y Spooren (2001), este tema está fundado en un acercamiento cognitivo conexionista respecto del procesamiento del texto. En este contexto, el conexionismo propone capturar la flexibilidad necesaria para modelar el curso temporal de la comprensión y producción de discurso. La idea de que el proceso cognitivo de producción e interpretación pueda ser modelado como un proceso dinámico en el cual las activaciones fluctúan, y que este proceso es influenciado o determinado por las características del texto, involucra uno de los más importantes cambios en el campo de los estudios lingüísticos y psicolingüísticos del discurso. Este aspecto de la representación del discurso propone que una visión dinámica en el proceso de la comprensión del discurso guía a las expectativas, que se realizan mientras el lector recorre el texto, la activación de conceptos, hechos y eventos como parte de la fluctuación constante de la representación del discurso. Varios modelos de comprensión se basan en esta idea y en resultados empíricos llevados a cabo bajo este enfoque. Algunos son el *modelo de construcción de estructuras* (Gernsbacher, 1990, 1996), el *modelo de paisaje de lectura* (van den Broek, Rinden, Fletcher & Thurlow, 1996; van den Broek, Young, Tzeng & Linderholm, 1999) y, en parte, el *modelo de construcción-integración* (Kintsch, 1988, 1998). Este último es un modelo híbrido, ya que presenta una integración de los modelos simbólicos, basados en proposiciones, y los conexionistas, basados en fluctuaciones de activación de conceptos.

En suma, la presentación de las relaciones semánticas en el ámbito del léxico, en el del enunciado y en el del texto nos permite tener una amplia panorámica de las variables que constituye la actividad tanto lingüística como mental en torno a la construcción del significado. En este sentido, podemos observar que en la formación del significado interactúan estrechamente no solo las unidades lingüísticas y sus relaciones, sino que también las “restricciones” psicosociales que afectan a la actividad discursiva. Además, es interesante notar, como lo hace Mayor (2000: 98), que:

“[...] las relaciones semánticas, tanto desde una perspectiva descriptiva (estructural) como funcional y procesual, son sumamente complejas y, a la vez, están organizadas en niveles de generalidad y especificidad variable”.

A la vez, plantea que considerando esta dificultad es posible agrupar estas relaciones semánticas en seis grandes categorías: *oposición, gradación, causalidad, atribución, agencia y contigüidad*.

Es interesante destacar que en los tres ámbitos observados (léxico, enunciado y texto) conviven en el estudio de las relaciones semánticas enfoques cognitivos de orden simbólico y de orden conexionista. Dado que nuestra investigación centra su atención en las relaciones semánticas léxicas, que pueden en alguna medida dar cuenta de las relaciones textuales, se hace necesario especificar cómo se llevan a cabo las representaciones del significado desde estos dos enfoques cognitivos, fundamentalmente, porque cada uno de estos enfoques darán origen a distintos tipos de representaciones computacionales del significado.

2.6 Representaciones simbólicas y conexionistas del significado

2.6.1 El enfoque cognitivista

Como plantean diversos autores (De Vega, 1984; Gardner, 1988, Martínez, 1997; Puente, 1998; Valiña & Martín, 2000; Pozo, 2001) la cognición humana, para el cognitismo, es la manipulación de símbolos a la manera de los computadores digitales. Esto es lo que se conoce como la “metáfora fuerte” del cognitismo (De Vega, 1984). Dicho de otro modo, la cognición es una representación mental y la mente manipula símbolos que representan rasgos del mundo o el mundo como “siendo” de alguna manera.

Martínez (1997) plantea que es un hecho incontrovertible para el cognitivista que la conducta inteligente (y por lo tanto intencionada) presupone algún tipo de habilidad para representar el mundo. Sin embargo, resulta complejo explicar cómo es que esas representaciones son físicamente realizadas en forma de un código simbólico en el cerebro o en una máquina (ver Fodor, 1984; Pinker, 1997).

El cognitivismo, al proponer un modelo computacional de la cognición, se sustenta en lo que Martínez (1997) denomina “un profundo prejuicio antibiológico”, ya que:

“[...] al cognitivista le resulta irrelevante saber si esos procesos son realizados en una substancia inmaterial, en un tejido humano o en un circuito de ordenador. El modelo cognitivista hace, por tanto, abstracción de las realizaciones biológicas de las estructuras cognitivas.” (Martínez, 1997: 208).

A pesar de lo anterior, este modelo es considerado por teóricos de las ciencias cognitivas como el único mecanismo conocido que puede compararse con la cognición. Esto, porque su funcionamiento formal le otorga un carácter de universalidad que permite a la máquina ser programada para realizar cualquier función formalmente especificada, al expresarse mediante sistemas simbólico (Martínez, 1997). Respecto de lo anterior Pylyshyn (1989: 50-51) plantea que:

“At the most abstract level, the class of mechanisms called computers are the only known mechanisms that are sufficiently plastic in their behavior to match the plasticity of human cognition. They are also the only known mechanism capable of producing behavior that can be described as “knowledge dependent”. Because of such properties computing remains the primary candidate for meeting the dual needs of (a) explaining cognition in mechanistic terms and (b) accounting for certain other wise problematic aspects of cognition – in particular the fact that behavior can be systematically influenced by – inducing differences in beliefs or goals.”

En suma, se puede determinar, siguiendo estos razonamientos, que como la cognición humana, la computación es representacional y, por consiguiente, tiene un valor semántico. De esta manera, es posible hablar de relaciones semánticas existentes entre las expresiones simbólicas al interior del sistema computacional, ya que como plantea Martínez (1997: 209) “No hay computación sin representación”. Ahora bien, el nivel semántico de las expresiones simbólicas está determinado en un computador por la sintaxis, dado que ésta es la que refleja la proyección semántica. En este sentido, si consideramos que los computadores construyen significado guiados por la sintaxis, cabría pensar que esto es de la misma manera en los seres humanos. Sin embargo, esta conclusión es debatible, pues no está totalmente claro que todas las diferencias semánticas relevantes para una explicación

de la conducta estén reflejadas sintácticamente (Martínez, 1997; Pozo, 2001; Landauer, 2002).

No obstante, el enfoque cognitivista sigue viendo en el paralelismo entre sintaxis y semántica la posibilidad de que tanto la inteligencia como la intencionalidad sean traducibles mecánicamente, pudiéndose construir un modelo mecánico del pensamiento.

Ahora bien, el reducir la semántica a la sintaxis hace que un ordenador manipule símbolos sin considerar lo que se representa en esos símbolos, es decir, construye un lenguaje en términos de símbolos no interpretados y, por ello, no puede dar cuenta del carácter contextual del significado, lo anterior le hace caer en un solipsismo metodológico, según el cual se estudian los estados mentales sin prestar ninguna atención al mundo externo que representan y sin asumir que existe (Martínez, 1997). Por otra parte, al distinguir en los computadores distintos niveles de organización (el nivel de conocimiento, el nivel simbólico y el nivel físico o biológico), como hace Pylyshyn (1989), y asimilarlos al comportamiento de la mente humana, se presentan los mecanismos y se crean las condiciones necesarias para que se dé en ellos una conducta inteligente e integrada. Sin embargo, estos supuestos son cuestionables, puesto que la sintaxis no es requerimiento indispensable de la semántica (Landauer, 2002) y, porque es más fácil comprobar que humanos y computadores no compartimos la misma arquitectura cognitiva, que lo contrario.

2.6.2 El enfoque conexionista

Desde mediados de los años ochenta, comenzaron a surgir alternativas en el marco de la psicología cognitiva que pusieron en duda los cimientos de la arquitectura simbólica basada en el procesamiento de información y, en especial, de la reducción de la semántica a la sintaxis. De entre esas alternativas, la que más radicalmente removió esos cimientos ha sido la propuesta de entender la mente y la memoria humana como un sistema de Procesamiento Distribuido en Paralelo (PDP), que:

“[...] asume que el procesamiento de información tiene lugar por medio de la interacción de un gran número de elementos de procesamiento simple,

llamados ‘unidades’, cada uno de los cuales envía señales excitatorias e inhibitorias a otras unidades” (McClelland, Rumelhart & Hinton, 1992: 46).

Pozo (2001) nos recuerda que, en la búsqueda del significado y la eficiencia representacional, el procesamiento de la información había ido buscando unidades de análisis cada vez más amplias o molares, como los esquemas, de los que el propio Rumelhart (1975) fue uno de los principales teóricos. Lo que hace el enfoque PDP es considerar que el nivel de análisis representacional-simbólico, que caracteriza al procesamiento de información, es demasiado molar y propone volver a un análisis subatómico para explicar esos niveles molares. El conexionismo, por tanto, se sitúa en la microestructura cognitiva. En este sentido, este enfoque cognitivo:

“[...] no niega que existe una macroestructura, igual que el estudio de las partículas subatómicas no niega la existencia de interacciones entre átomos. Lo que los modelos PDP hacen es describir la estructura interna de las unidades más grandes, igual que la física subatómica describe la estructura interna de los átomos que forman los componentes de unidades más grandes de la estructura química” (McClelland, Rumelhart & Hinton, 1992: 48).

Está claro que esta postura supone un cambio radical frente a la postura simbólica, tal como lo manifiesta Norman (citado en Pozo, 2001: 63) al dar cuenta de las diferencias entre ambos enfoques:

“En la vieja concepción, el procesamiento lo hacían componentes separados que se comunicaban pasándose mensajes –símbolos- entre sí. El sistema perceptivo se comunicaba con el sistema de memoria. La memoria se comunicaba con la percepción y solución de problemas. Los sistemas de lenguaje se comunicaban con los demás. Las comunicaciones consistían en mensajes simbólicos interpretados a través de un sistema representacional que implicaba una correspondencia entre símbolo y referente y, en consecuencia, un mecanismo de interpretación[...]. Se suponía el énfasis en la representación simbólica y por tanto se ponía énfasis en las reglas de pensamiento, solución de problemas y memoria. [...] la nueva concepción es muy diferente. [...] el procesamiento lo hacen redes PDP que se configuran a sí mismas para ajustarse a los datos que llegan con la mínima discrepancia o conflicto[...]. La información pasa de unas unidades a otras no mediante mensajes sino mediante valores de activación, mediante escalares, no mediante símbolos. La interpretación del procesamiento no se hace en términos de los mensajes que se envían sino más bien de los estados que están activos. Así, lo importante son las unidades que están activas, no qué mensajes se envían”.

El PDP supone, por tanto, un nuevo intento reduccionista, pero que toma como modelo computacional, ya no el computador sino el propio cerebro. Las unidades de información serían las neuronas, cuyas pautas de activación conformarían redes neuronales. Todo lo que sucede por encima de ese nivel subsimbólico puede ser reducido al funcionamiento de una red neuronal (McClelland, Rumelhart & Hinton, 1992).

Este reduccionismo permite establecer que para el conexionismo, como para cualquier corriente asociacionista, las estructuras de la memoria (o todo lo que esté por encima del nivel subsimbólico), pueden ser reducidas a sus elementos componentes. Incluso, el conexionismo niega la existencia de tales estructuras, ya que lo único que hay en la memoria son pautas de activación (Pozo, 2001).

En función de lo anterior, cabe preguntarse cómo se organizan las representaciones si no hay estructuras en la memoria que almacenen información. Según Pozo (2001), lo que hace un sistema de memoria conexionista es reflejar o copiar, en forma de ciertas pautas de activación, la estructura del ambiente, las regularidades que se producen en el entorno, de acuerdo a un principio de correspondencia entre el mundo interno (en este caso la red neuronal) y el mundo externo (el entorno al que se enfrenta esa red neuronal), característico de los enfoques asociacionistas radicales. Como plantea Pozo (2001: 66): *“El nuevo asociacionismo del PDP ha sustituido la caja negra por un caja vacía, que se limita a reflejar, mediante cálculos muy elaborados, la estructura del ambiente”*.

Ahora bien, la recuperación de un determinado contenido representacional previamente activado en la red neuronal, supone necesariamente una reconstrucción o reactivación, dado que el proceso computacional total está descrito en términos de las interacciones entre miles de procesadores restringidos. De esta forma, la representación de conocimiento es distribuida a través de estas redes formadas por unidades y conexiones (Iza & Ezquerro, 1999).

Aunque el enfoque conexionista y los tratamientos basados en el uso de redes neuronales han alcanzado un desarrollo notable en los últimos años (García, 1992; Moreno, 1998; Pozo, 2001), los modelos conexionistas adolecen de algunas limitaciones que es preciso conocer para valorar adecuadamente las posibilidades de este enfoque.

En primer lugar, existen fenómenos relacionados con el aprendizaje que no resultan fáciles de implementar mediante el uso de redes conexionistas. Por ejemplo, la generación de nuevas construcciones cognitivas que aportan significados más ricos y extensos a los conceptos ya conocidos. Otra situación similar se produce cuando se intenta simular procesos de carácter implícito o procedimientos complejos, siendo estos fundamentales en la enseñanza de las ciencias (Campanario, 2004). Por otra parte, muchos de los logros de los modelos conexionistas actuales tienen que ver con el reconocimiento de patrones, pero estos procesos son mucho más limitados que los procesos cognitivos implicados en la comprensión y aprendizaje complejo de conceptos, principios y teorías. Además, un concepto clave de los modelos conexionistas, el de activación, no tiene una traducción inmediata y unívoca en términos relacionados con procesos cognitivos.

Los mecanismos y aplicaciones matemáticas que utilizan los modelos conexionistas están tan alejados de los enfoques clásicos en psicología que Fodor (1984) y Pylyshin (1989), dos críticos bien conocidos, señalan con cierta ironía que *"todas las razones que hay para pensar que los modelos conexionistas podrían ser verdaderos, son razones para pensar que pudieran no ser Psicología"* (citado en García-Madruga, 1992: 25).

El lector interesado en profundizar en las limitaciones de los modelos conexionistas puede encontrar análisis más detallados en Pozo (2001) o De Vega (1984)

También puede ser interesante analizar la introducción de la edición española de la obra de Rumelhart, McClelland y el grupo PDP (1992), a cargo de García-Madruga (1992).

2.7 Representaciones computacionales del significado

2.7.1 Lenguajes de representación semántica

Desde el punto de vista de la lingüística computacional, las aproximaciones a la semántica han sido entendidas bajo la noción simbólica de que el significado de las expresiones lingüísticas puede ser capturado en estructuras formales, las que son denominadas representación del significado. Correspondientemente, el marco usado para especificar la sintaxis y semántica de esas representaciones es conocido como "lenguajes de

representación del significado” (Jurafsky & Martin, 2000). Sin embargo, cabe advertir que, en términos, rigurosos la “representación del significado” llevado a cabo por los computadores, no es más que un procedimiento de cálculo, en el cual el computador transforma ciertas instrucciones programadas en un lenguaje binario que le permite procesar una información de entrada y convertirla en una información de salida, según las condiciones programadas. En este sentido, solo los seres humanos pueden representarse el significado de un texto, en tanto que los computadores solo permiten modelar tal representación. En lo que sigue el concepto de representación será utilizado entendido como el modelamiento computacional de las representaciones de significado.

La necesidad de modelar la representación del significado surge, como lo plantean Jurafsky y Martin (2000), cuando ni los datos lingüísticos puros ni los modelamientos fonológicos, morfológicos y sintácticos logran dar cuenta del tipo de procesamiento semántico requerido para el análisis de tareas específicas en donde se requiere una representación conceptual de los datos lingüísticos.

Algunas tareas para las cuales se requiere el modelamiento de representaciones conceptuales son: responder a preguntas de ensayo en un examen, decidir qué ordenar por medio de la lectura de un menú, aprender a utilizar un programa de computadora leyendo un manual, seguir las instrucciones de una receta, etc. Estas tareas tienen en común el requerimiento de acceder a representaciones que permitan vincular los elementos lingüísticos involucrados con el conocimiento de mundo no lingüístico necesario para llevarlas a cabo con éxito. Por ejemplo, responder y evaluar un ensayo requiere conocimiento previo sobre el tema de las preguntas, conocimiento respecto del nivel del estudiante y el nivel que debiera lograr, y conocimiento de cómo estas preguntas son normalmente respondidas.

En este enfoque representacional se toman los datos lingüísticos y se construyen representaciones del significado que son creadas del mismo tipo de material que es usado para representar el conocimiento del mundo cotidiano. A este procedimiento, en el cual tales representaciones son creadas y asignadas a los datos lingüísticos, se le denomina *análisis semántico* (Jurafsky & Martin, 2000). De esta forma, una oración será representada a través de los lenguajes de representación de significado. Los más conocidos son: el

Cálculo de Predicado de Primer Orden (FOPI, por su sigla en inglés), la red semántica, el diagrama de dependencia contextual y la representación basada en marcos (Jurafsky & Martin, 2000; Lappin, 2003). Si bien, estos lenguajes se diferencian bastante, a un nivel de abstracción, ellos comparten como fundamento común la noción de que la representación del significado consiste en estructuras compuestas a partir de un grupo de símbolos. Cuando se ordenan apropiadamente, estas estructuras de símbolos se corresponden con objetos y relaciones entre los objetos en un mundo representado.

Veamos algunos ejemplos de lenguajes representacionales, presentados por Jurafsky y Martin (2000: 503), para la oración del idioma inglés: “*I have a car*”.

FOPC: $\exists x,y \text{Having}(x) \wedge \text{Haver}(\text{Speaker}, x) \wedge \text{HadThing}(y,x) \wedge \text{Car}(y)$

Red Semántica:

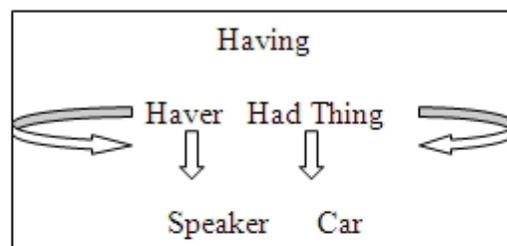
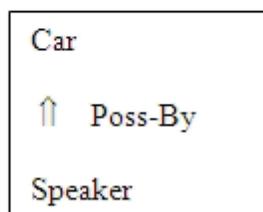
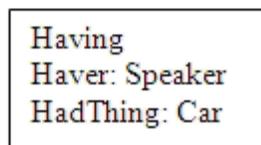


Diagrama de Dependencia Contextual:



Representación basada en Marcos:



Cabe hacer notar que estas representaciones pueden ser vistas de, al menos, dos modos: como representación del significado de un dato lingüístico particular y como representación del estado de la situación en un mundo particular. Es esta perspectiva dual la que permite que las representaciones sean usadas para vincular los datos lingüísticos con el mundo y nuestro conocimiento respecto de él.

La mayor importancia de estos modelos de representación está en la posibilidad de una representación computacional del significado, por ello deben, como lenguajes de representación del significado, responder a las necesidades prácticas del procesamiento semántico. Entre ellas, está la necesidad de determinar la verdad de las proposiciones, resolver los problemas de representaciones lingüísticas ambiguas (en términos lógico-formales), representar las variables y las relaciones entre ellas para permitir la activación de inferencias lógicas y, por último, ser lenguajes suficientemente expresivos para representar semánticamente una amplia gama de posibilidades (idealmente un lenguaje de representación del significado que pueda representar el significado de cualquier expresión de lenguaje natural) (Jurafsky & Martin, 2000; Lappin, 2003).

Cabe destacar que entre los lenguajes de representación, el FOPC es el lenguaje que ha sido más importante en la investigación de los problemas de semántica computacional, ya que satisface adecuadamente los requerimientos computacionales mencionados más arriba (Jurafsky & Martin, 2000; Manning & Schütze, 2003; Lappin, 2003).

Dejaremos hasta aquí esta breve revisión de los modelos de representación puesto que, si bien permiten representar el significado de las oraciones (sobre todo las que tienen significado literal), responden a un enfoque fundamentalmente basado en proposiciones lógico-simbólicas. Esto los hace, por una parte, poco útiles en la representación del significado de grandes cantidades de información textual de lenguaje naturales y, por otra, requieren de un amplio conocimiento previo y de reglas programadas que respondan a todas las representaciones del significado (Allen, 1995; Jurafsky & Martin, 2000; Lappin, 2003).

Como hemos observado sucintamente, la representación computacional del significado, en los términos expuestos anteriormente, se centra en la representación del enunciado. Estas representaciones se analizan composicionalmente a partir de una semántica guiada por la sintaxis, en donde cada componente se identifica con un constituyente sintáctico.

2.7.2. Similitud Semántica

Existen otras formas de representar el significado, fundamentalmente léxico, que no se basan en la lógica simbólica y que responden más bien a parámetros empiricistas e inductivos. Diversos investigadores del ámbito de la psicolingüística computacional han trabajado utilizando esta perspectiva, en particular enfocándose en la adquisición de las propiedades semánticas de las palabras a partir de *corpora* textuales. Estos autores han desarrollado modelos que permiten dar cuenta, de forma más o menos acuciosa, del procesamiento del lenguaje natural y, en particular, de la modelación de los distintos tipos de memoria de largo plazo (léxica, sintáctica y semántica). Algunos de estos modelos son: *Latent Semantic Analysis* (Landauer & Dumais, 1996, 1997; Landauer, Foltz & Laham, 1998; Kintsch, 1998, 2001; Landauer, 2002), *The Syntagmatic Paradigmatic Model* (Dennis & Harrington, 2001; Dennis, 2003), *Semi Discrete Matrix Decomposition* (Kolda & O'Leary, 1998), *Pooled Adjacent Context Model* (Redington, Chater, & Finch, 1998), *Probabilistic Latent Semantic Indexing* (Hofmann, 2001), *Latent Dirichlet Allocation* (Blei, Ng, & Jordan, 2002), *The Topics Model* (Griffiths & Steyvers, 2002), *Word Association* (Steyvers, Shiffrin & Nelson, 2004), *Non-negative matrix factorization* (Lee & Sung, 1999; Ge & Iwata, 2002), *Local Linear Embedding* (Roweis & Saul, 2000), *UNICON* (Lin & Pantel, 2001).

Todos los modelos mencionados tienen en común el hecho de que plantean algoritmos, a través de los cuales se pretende explicar cómo ocurre el procesamiento semántico del lenguaje natural (fundamentalmente léxico) a partir de textos naturales y que utilizan la noción de similitud semántica. Esta noción supone que determinar automáticamente cuán similar es una palabra nueva en relación a una conocida es más fácil que determinar el significado propiamente tal de ella (Manning & Schütze, 2003). A pesar de todas las limitaciones que puedan existir, en especial aquellas que tiene que ver con las adecuaciones a los valores de verdad del significado (Quine, 1968), la similitud semántica es la medida empírica más útil que existe y funciona bajo el supuesto de que palabras semánticamente similares se comportarán de manera similar (Manning & Schütze, 2003; Matsumoto, 2003). Se observa aquí una estrecha relación con el estudio colocacional del significado realizado por la escuela neofirthiana esbozado brevemente en 2.4.6.1.

Esta técnica de estudio del significado ha sido aplicada a diversas tareas de procesamiento natural del lenguaje como: la desambiguación de significados, la extracción de restricciones de selección, el análisis sintáctico, la asignación de grupos preposicionales, la estructuración de los compuestos nominales, la resolución de la referencia, la clasificación o la agrupación. También se ha utilizado en los sistemas de recuperación de la información, las consultas en internet, la extracción de información y el resumen automático o la traducción automática (Matsumoto, 2003; Rodríguez, 2004).

Manning y Schütze (2003) advierten que la similitud semántica no es una noción intuitiva y clara como podría pensarse. Esto, debido a que la noción de similitud semántica ha sido utilizada en investigaciones de variadas disciplinas, desde la psicología cognitiva, pasando por la neurofisiología, la filosofía y la lingüística, hasta la inteligencia artificial, por nombrar solo algunas.

Según Latorres (2004), un iniciador de los modelos de similitud semántica fue Russell (1937) con su “teoría de clases y similitud”, en la cual se planteaba un modelo que consideraba todos los aspectos del mundo real. Sin embargo, esta teoría contenía varias paradojas que la invalidaron. Una de estas paradojas fue el problema de “la clase de clases”, para el cual, y en cooperación con Frege (1972), se propuso la solución que planteaba que ninguna clase puede ser una extensión de una super-clase que cae debajo de la primera, resolviendo el problema, pero haciendo que la solución sea mucho más restrictiva, y al final, menos capaz de representar la realidad. De todas formas, como asevera Latorres (2004), estos modelos afectaron la forma de percibir los problemas de la realidad entre los futuros filósofos, los lógicos, y desde ellos hasta los estudiosos del lenguaje desde una perspectiva probabilística y computacional.

Entre los que tomaron una visión psicolingüística, Osgood (1952) y Osgood, et al. (1976) analizaron el significado de las palabras experimentando con la interpretación humana. Este tipo de evaluación, basada en el análisis de humanos realizando evaluación de sinónimos, fue validada luego por Miller y Charles (1991). Los resultados de estos experimentos son utilizados en varios estudios recientes para probar y validar estadísticamente varios algoritmos computacionales a efecto de determinar la similitud semántica entre conceptos expresados por las palabras. Estos estudios se fundamentan en la idea de que el juicio

humano se considera correcto por definición, tal como es afirmado por Budanitsky y Hirst (2001). Estos últimos autores aseveran que la similitud es lo que hacen los humanos, por lo que comparan y evalúan el desempeño de algoritmos computacionales con el desempeño medido en una muestra de humanos realizando la tarea de determinar la similitud semántica entre conceptos, principio que también es utilizado por los modelos mencionados al inicio de este apartado.

Así, mientras para algunos la similitud semántica se entiende como una extensión de sinonimia, otros la entienden en el sentido en el que dos palabras comparten un mismo dominio semántico o tópico. En esta lógica, las palabras son similares si ellas refieren a entidades en el mundo que tienen muchas posibilidades de co-ocurrir como en “*doctor*”, “*enfermera*”, “*fiebre*” e “*intravenosa*” refiriendo estas palabras a diferentes entidades e incluso, pudiendo pertenecer a categorías sintácticas diferentes. Esto ocurre puesto que los modelos algorítmicos de extracción de información están basados en una interpretación “extrema” de los principios de la semántica composicional (Jurafsky & Martin, 2000). En estos sistemas, el significado de los documentos reside solamente en las palabras que están contenidas en ellos. Es decir, el orden y la constitución de las palabras que forman las oraciones, que a su vez constituyen el texto, no tienen importancia en la determinación de su significado. Debido a que ignoran la información sintáctica, este tipo de aproximación es comúnmente denominado como métodos de “depósito de palabras” (*bag of words*) (Jurafsky & Martin, 2000).

Manning y Schütze (2003) destacan que la idea más relevante en relación con la similitud semántica es que los resultados pueden ser explicados por el grado de intercambiabilidad contextual o el grado en el cual una palabra puede ser substituida por otra en un contexto dado. Desde una perspectiva algorítmica, la medición de la similitud semántica es conceptualizada por los modelos de tipo vectorial como una medida de similitud de vectores para determinar la similitud de dos palabras que son representadas como vectores en un espacio multidimensional. Para llevar a cabo esto, se construye una matriz en la cual se representa numéricamente la co-ocurrencia de las palabras por una unidad mayor, denominada “*documento*” (normalmente oraciones o párrafos). Luego se construyen otras

dos matrices en las cuales se considera la correlación de palabras por palabras y la correlación de documentos por documentos.

2.7.2.1 El modelo vectorial

A continuación presentaremos de manera detallada la forma en que es utilizado el modelo vectorial para la recuperación de información. Nos interesa destacar este modelo, porque es la manera más sencilla de explicar cómo se llevan a cabo las operaciones matemáticas que permiten determinar la similitud entre documentos a partir de las palabras contenidas en ellos.

El modelo vectorial fue definido por Salton (1968) y es ampliamente usado en operaciones de recuperación de información, así como también en operaciones de categorización automática, filtrado de información, etc. (Zazo, Figuerola, Alonso & Gómez, 2002).

Según Zazo, et al. (2002) en el modelo vectorial se intenta recoger la relación de cada documento D_i , de una colección de N documentos, con el conjunto de las m características de la colección. Formalmente un documento puede considerarse como un vector que expresa la relación del documento con cada una de esas características.

$$1. D_t \rightarrow \vec{d} = (c_{t1}, c_{t2}, \dots, c_{tm})$$

Es decir, ese vector identifica en qué grado el documento D_i satisface cada una de las m características. En ese vector, c_{ik} es un valor numérico que expresa en qué grado el documento D_i posee la característica k . La noción de “característica” suele concretarse en la ocurrencia de determinadas palabras o términos en el documento, aunque nada impide tomar en consideración otros aspectos. Respecto de esto último, cabe señalar que este tipo de procedimientos se han utilizado en el reconocimiento de objetos, donde las características son de carácter viso-perceptual (color, forma, etc) (Landauer, 2002).

Si se consideran los términos como características definitorias del documento, el proceso que debe seguir el sistema pasa primero por seleccionar aquellos términos útiles que permitan discriminar unos documentos de otros. En este punto, debemos señalar que no

todas las palabras contribuyen con la misma importancia en la caracterización del documento. Desde el punto de vista de la recuperación de información, existen palabras casi vacías de contenido semántico (*stop words*), como los artículos, preposiciones o conjunciones, que son poco útiles en el proceso. Pero también son poco importantes aquellas palabras que por su frecuencia de aparición en toda la colección de documentos pierden su poder de discriminación. En la recuperación de información todas ellas forman parte del conjunto de palabras vacías, que se eliminan en el proceso de indexación.

Además de la eliminación de palabras vacías, en el proceso se pueden incluir aplicaciones léxicas como lematización o extracción de raíces, etiquetado de términos, detección de unidades multipalabra, etc.

Una vez seleccionado el conjunto de términos caracterizadores de la colección de documentos, es necesario obtener el valor de cada elemento del vector del documento. El caso más simple es utilizar una aproximación binaria, de forma que si en el documento D_i aparece el término k , el valor c_{ik} sería 1, y en caso contrario sería 0.

No obstante, una palabra puede aparecer más de una vez en el mismo documento y, además, unas palabras pueden considerarse con más peso, esto es, más significativas que otras, de forma que el valor numérico de cada uno de los componentes del vector obedece normalmente a cálculos más sofisticados que la simple asignación binaria. Por otro lado, también es importante normalizar los vectores para no privilegiar documentos

$$2. \quad \bar{d}_i = \frac{1}{\sqrt{\sum_{j=1}^m w_{ij}^2}} (w_{i1}, w_{i2}, \dots, w_{im})$$

Se han propuesto diversos métodos para calcular el peso de cada término en el vector documento (Salton & McGill, 1983; Salton & Buckley, 1988; Harman, 1992), pero en general, para estimarlos se parte de dos ideas en cierto sentido contrapuestas: si un término aparece mucho en un documento, es importante para caracterizar ese documento. Pero si aparece en muchos documentos de la colección, no es beneficioso para distinguir un

documento de los demás, dado su escaso poder discriminatorio, resultando poco útil para la recuperación.

Para determinar la capacidad de representación de un término para un documento dado se computa el número de veces que aparece en dicho documento, obteniéndose la frecuencia del término en el documento, tf (*term frequency*).

Por otra parte, si la frecuencia de un término en toda la colección de documentos es extremadamente alta, se opta por eliminarlo del conjunto de términos de la colección (pertenece al conjunto de palabras vacías). Podría decirse que la capacidad de recuperación de un término es inversamente proporcional a su frecuencia en la colección de documentos. Esto es lo que se conoce como idf (*inverse document frequency*).

Así, para calcular el peso de cada elemento del vector que representa al documento se tiene en cuenta la frecuencia inversa del término en la colección, combinándola de alguna forma con la frecuencia del término dentro de cada documento. Normalmente se utiliza para ello el producto simple (*dot product*) (Harman, 1992).

$$3. \quad w_{ij} = tf_i \cdot idf_j$$

Salton y Buckley (1988) experimentaron con más de 200 sistemas de cálculo de pesos, pero uno de los más utilizados viene dado por la ecuación 4, que expresa el peso del término j en el documento i .

$$4. \quad W_{ij} = tf_i \cdot \log \frac{N}{df_j}$$

Donde df_j es el número de documentos en que aparece el término j , y N el número de documentos de la colección.

Ahora bien, el proceso realizado para los documentos también puede aplicarse a las consultas. Efectivamente, una consulta Q , realizada en lenguaje natural está formada por términos y , por tanto, puede verse como un documento más, seguramente bastante breve, aunque no siempre. Así pues, el mecanismo de obtención de pesos también se aplica a las

consultas, para de esta manera poder disponer de representaciones homogéneas de consultas y documentos, que posibiliten obtener el grado de similitud entre ambas representaciones.

El vector representante de la consulta está formado por un vector de igual número de elementos que los vectores de los documentos. Cada elemento de ese vector expresa el grado en que cada uno de los términos de la colección representa las necesidades informativas de la persona que hace la consulta.

$$5. \quad Q \rightarrow \bar{q} = \frac{1}{\sqrt{\sum_{j=1}^m P_j^2}} (p_1, p_2, \dots, p_n)$$

La resolución de la consulta consiste en un proceso para establecer el grado de semejanza entre el vector consulta y el vector de cada uno de los documentos. Para una consulta determinada, cada documento arrojará un grado de similitud determinado; aquéllos cuyo grado de similitud sea más elevado se ajustarán mejor a las necesidades expresadas en la consulta, desde el punto de vista del sistema de recuperación de información. No obstante, es el usuario el que debe decidir la relevancia de los documentos recuperados, siendo ésta una característica totalmente subjetiva del mismo.

El modo más simple de calcular la similitud entre una consulta y un documento, utilizando el modelo vectorial, es realizar el producto escalar de los vectores que los representan (ecuación 6). En esta ecuación se incluye la normalización de los vectores, a fin de obviar distorsiones producidas por los diferentes tamaños de los documentos. El índice de similitud más utilizado es el coseno del ángulo formado por ambos vectores. Para una consulta Q, el índice de similitud con un documento D_i es:

$$6. \quad \text{simil}(Q, D_i) = \frac{\sum_{j=1}^m p_j d_{ij}}{\sqrt{\sum_{j=1}^m p_j^2 \cdot \sum_{j=1}^m p_{ij}^2}}$$

Existen otros métodos propuestos para calcular la similitud. Algunos ejemplos son: el coeficiente de emparejamiento (*matching coefficient*), el coeficiente de Dice, el coeficiente de Jaccard (o Tanimoto) y el coeficiente de solapamiento (*overlap coefficient*). Síntesis con la descripción de estas medidas de similitud pueden encontrarse en Jurafsky y Martin (2000), Manning y Schütze (2003), Tzoukermann, Klavans y Strzalkowski (2003).

Los resultados obtenidos a partir del cálculo del índice de similitud entre la consulta y todos los documentos permite ordenar los resultados en orden decreciente (Zazo, et al, 2002). De esta manera se le ofrecen al usuario primero los documentos que el sistema de recuperación considera más similares con la consulta, y que pueden coincidir, o no, con lo esperado por el usuario. La relevancia es la medida subjetiva que el usuario tiene para determinar si los resultados son adecuados, y en qué grado, a sus necesidades informativas.

A continuación presentamos un ejemplo, aparecido en Jackson y Moulinier (2002), que puede ayudar a comprender cómo se asocian semánticamente los documentos por medio de la medida de su similitud semántica en un espacio vectorial.

Consideremos un caso simple tridimensional con un corpus de tres documentos. Las dimensiones son: “sí”, “no”, “quizás” y los documentos son:

D₁: ‘sí sí sí’

D₂: ‘no no no’

D₃: ‘sí quizás sí’

Las dimensiones de este espacio son concebidas como características que permiten distinguir un documento de otro. Los componentes de los vectores de los documentos pueden ser vistos como “pesos” que codifican la importancia de los rasgos correspondientes para un documento. Podemos representar cada documento en nuestro corpus con un vector tridimensional. Asumiendo que los componentes del vector son frecuencias brutas asociadas con las dimensiones y apareamiento en el orden ‘sí’, ‘quizás’, ‘no’, entonces los vectores que se generan son:

d₁: (3,0,0)

$d_2: (0,0,3)$

$d_3: (2,1,0)$

Estos vectores pueden usarse para representar los documentos D_1 , D_2 y D_3 , respectivamente.

Se puede observar a simple vista que d_1 es más cercano a d_3 que a d_2 en el espacio vectorial definido por los términos. Los vectores d_1 y d_2 están en ángulo recto en el espacio vectorial, mientras que d_1 y d_3 se encuentran en ángulo agudo. Esto está acorde a la simple intuición de que D_1 es más similar al D_3 que al D_2 .

La similitud entre documentos puede ser medida por el producto simple de sus correspondientes vectores. Como se ha visto, la medida de coseno es el producto simple de los vectores normalizado por el largo del vector. En el ejemplo omitimos la normalización, por motivos didácticos, así:

$$7. \text{simil}(D_1, D_2) = d_1 \cdot d_2 = (3,0,0) \cdot (0,0,3) = 0$$

Mientras que:

$$8. \text{simil}(D_1, D_3) = d_1 \cdot d_3 = (3,0,0) \cdot (2,1,0) = 6$$

El documento D_1 y D_2 no tienen palabras en común y son, por lo tanto, completamente disímiles. Esto fundado por el hecho geométrico de que sus vectores son ortogonales y por el hecho algebraico de que el producto simple es cero. En tanto D_1 y D_3 comparten la primera dimensión ("sí") por lo que sus vectores se correlacionan y su producto simple no es cero.

En términos más generales, dado dos documentos cualesquiera, D_1 y D_2 , con vectores

$$9. d_1 = (d_{1,1}, \dots, d_{1,t}) \quad \text{y} \quad d_2 = (d_{2,1}, \dots, d_{2,t})$$

la similitud entre los dos documentos puede ser calculada mediante la siguiente fórmula:

$$10. \text{simil}(D_1, D_2) = \sum_{i=1}^{i=t} d_{1,i} \cdot d_{2,i}$$

Insistimos en que este es un ejemplo muy simple, ya que en un cálculo con textos reales se deben considerar múltiples dimensiones (cada palabra será una dimensión), se debe normalizar por el largo de cada vector (por el valor asignado a cada palabra en el documento) y se debe utilizar una función más sofisticada que el cálculo de frecuencia para estimar el peso de cada vector (función que normalmente marca la diferencia en los modelos de recuperación léxico-semántica basados en métodos vectoriales de similitud semántica). Sin embargo, la esencia del procedimiento es la que se ha usado en este ejemplo.

En el caso de nuestra investigación nos interesa, dentro de los modelos basados en el modelo vectorial, el que corresponde al modelo de indexación semántica latente o, como es conocido en psicolingüística computacional, *Análisis Semántico Latente* (LSA). Esto, fundamentalmente, porque es un modelo que ya tiene varios años de utilización tanto en inglés como lenguas de origen románico (francés, portugués y, en muy pocos casos en español), además, porque trabajaremos con una colección de textos estáticos, esto implica que no se están actualizando continuamente, y porque nos interesa realizar búsquedas que nos permitan representar cuantitativamente las relaciones semánticas existentes entre una palabra clave y un párrafo (por ejemplo, la similitud existente entre las palabras clave y el resumen), así como la relación entre párrafos (por ejemplo, similitud entre resumen y párrafos del artículo científico). Otro motivo es que con este modelo no se requiere un marcaje estructural de los textos de ningún tipo, es decir, los textos no deben ser etiquetados morfosintácticamente como lo requieren los sistemas de representación simbólicos del significado, guiados por la sintaxis. Se debe reconocer que, a pesar de la existencia de modelos más modernos y, en algunos, casos más efectivos para desarrollar esta tarea, el LSA es un modelo clásico, probado en múltiples lenguas y con un desarrollo computacional (no sólo teórico y algorítmico) accesible para el desarrollo de este estudio.

En el siguiente apartado, explicitaremos, con más detalle, el funcionamiento y las aplicaciones de este modelo que se plantea hoy en día, según sus seguidores, como un

modelo robusto en la determinación, no sólo, de las relaciones semánticas entre palabras, sino que también de oraciones y párrafos (Landauer, 2002).

2.7.3. Análisis semántico latente

Como hemos visto en los apartados anteriores, una forma de determinar el significado de las palabras es por medio del establecimiento de sus relaciones en los contextos de palabras en los cuales aparecen. Hemos establecido también que estas relaciones pueden ser representadas por medio de sistemas computacionales, los cuales responden fundamentalmente a dos enfoques subyacentes. El primero es el enfoque simbólico-deductivo en el que el significado se determina mediante la construcción de reglas de proyección semántica de los rasgos sintácticos de cada unidad lingüística (ver Jurfasky & Martin, 2000; Manning & Schütze, 2003; Lappin, 2003). El segundo enfoque es empírico-inductivo, es decir, no se sustenta en reglas previas que establezcan el orden en el que las unidades lingüísticas serán relacionadas para obtener la representación, sino que utilizan la información de manera directa desde el corpus de textos, esto es, sin realizar análisis previos de estas unidades. Estos sistemas utilizan técnicas matemáticas y estadísticas para determinar tanto los valores de las unidades lingüísticas como los valores de las fuerzas de relación entre estas unidades. Esto último, permite configurar una representación cuantitativa de las unidades lingüísticas utilizadas para el análisis.

Un método empírico ampliamente utilizado por los sistemas de recuperación de información es el modelo vectorial, como veíamos en el apartado anterior, el cual sobre la base del cálculo de la similitud de las unidades lingüísticas pueden establecer por medio de medidas de coseno, entre otras, la cercanía o la lejanía de dos palabras u otras unidades lingüísticas mayores en un espacio vectorial, lo cual a su vez da cuenta de su similitud de significado. Entre estos modelos de espacios vectoriales destaca el análisis que se hace utilizando la “indexación semántica latente” (LSI, por su uso en inglés) (Tzoukermann, Klavans & Strzalkowski, 2003). A través de este método se pretende solucionar los problemas de indexación y recuperación de documentos que se producen al utilizar métodos basados en palabras clave. Para Deerwester, Dumais, Furnas, Landauer y Harshman (1990: 392) el problema se define de la siguiente manera:

“The problem is that users want to retrieve on the basis of conceptual content, and individual words provide unreliable evidence about the conceptual topic or meaning of a document. There are usually many ways to express a given concept, so the literal terms in a user’s query may not match those of a relevant document. In addition, most words have multiple meanings, so terms in a user’s query will literally match terms in documents that are not of interest to the user”.

Es debido a este problema que ellos suponen que:

“[...] there is some underlying latent semantic structure in the data that is partially obscured by the randomness of word choice with respect to retrieval” (Deerwester et al., 1990: 392).

Es importante destacar que para los autores el término “semántica” es utilizado aquí para referirse al hecho de que los términos que componen un documento pueden ser tomados como referentes del documento mismo o de su tema (en términos más lingüísticos nos referimos a una semántica de tipo léxico). En tanto que con “estructura semántica” quieren significar la estructura correlacional del modo en que palabras individuales co-ocurren con los documentos en los que aparecen (Deerwester et al., 1990). Esta idea de co-ocurrencia se explica por el hecho de que dos o más términos ocurren en los mismos documentos más a menudo que el sólo azar (Manning & Schütze, 2003). Por último, el concepto de “latente” implica que estas asociaciones semánticas están inherentes en la estructura de los documentos (Tzoukermann, Klavans & Strzalkowski, 2003).

En este modelo, las consultas y los documentos a recuperar son proyectados en un espacio con dimensiones semánticas latentes. La co-ocurrencia de términos son proyectadas en las mismas dimensiones, en tanto los términos no co-ocurrentes son proyectados en diferentes dimensiones. En el espacio semántico latente una consulta y un documento pueden tener una alta similitud, medida por su coseno, aun si es que ellos no comparten ningún término, siempre que sus términos sean semánticamente similares de acuerdo al análisis de co-ocurrencia (Deerwester et al., 1990; Manning & Schütze, 2003).

El espacio semántico puede ser explicado de una manera muy simplificada y didáctica. Los párrafos con contenido semántico similar están localizados juntos en el espacio y las

palabras que han ocurrido en un párrafo similar están localizadas también cerca unas de otras en el espacio (Quesada, Kintsch & Gómez, 2002).

Esta mirada da cuenta de uno de los beneficios de la reducción de dimensiones: las palabras que no co-ocurren, pero ocurren en contextos similares (por ejemplo: doctor y médico) serán agrupadas juntas. Dado que las palabras y párrafos en el espacio están ordenados de acuerdo a su contenido semántico, es que se le ha denominado “espacio semántico” (Quesada, 2003).

Un aspecto muy importante de este modelo es que el espacio semántico proyectado tiene menos dimensiones que el espacio semántico original, el cual tiene tantas dimensiones como términos. Esto quiere decir que se genera una representación matemática del texto en la que se produce eliminación de aquellos términos, entendidos como variables matemáticas, que se consideran menos relevantes para el análisis. Es por ello que el LSI es conocido como un método de reducción de dimensiones.

Como técnica de reducción de la dimensionalidad, el LSI utiliza una técnica de descomposición de valores singulares (SVD sigla tomada del inglés), la cual es similar al análisis factorial, en particular al análisis de componentes principales (PCA sigla tomada del inglés), pero se diferencia de él en que el PCA se utiliza para matrices cuadradas, es decir, matrices en las cuales se considera una variable (términos o documentos), en tanto el SVD puede ser aplicado a cualquier matriz no cuadrada. En el caso de la representación de textos, se considera que tanto los términos (palabras o agrupación de palabras, por ejemplo, sintagmas fijos) como documentos (oraciones o párrafos) son variables que son tomadas conjuntamente en la matriz. Otra diferencia es que la representación generada por el SVD toma en consideración el realizar los menores cambios posibles de la matriz original para la construcción del espacio semántico latente.

En síntesis, el análisis de co-ocurrencia y la reducción de la dimensionalidad son las dos características más relevantes del LSI. Estas son las que permiten la recuperación de documentos desde una base de datos aunque los términos ingresados en la consulta no aparezcan explicitados en los documentos. Este procedimiento, en comparación a los

sistemas de recuperación de información basados en palabras claves, constituye entre un 20 a un 30% de mejoramiento en la recuperación de documentos (Dumais, 1997).

Luego de múltiples experiencias con el uso del método de indexación semántica latente, un grupo de investigadores interesados en los aspectos cognitivos subsimbólicos han encontrado en este modelo de indexación un recurso que les ha sido muy útil y que les ha servido de base para explicar, según ellos, el proceso de adquisición y representación del conocimiento (Deweester et al, 1990; Foltz, 1990; Landauer & Dumais, 1996, 1997; Landauer, Foltz & Laham, 1998; Kintsch, 1998, 2001; Landauer, 2002; Quesada, Kintsch & Gómez, 2002; Quesada, 2003). Así, surge a partir del LSI un modelo teórico computacional denominado Análisis Semántico Latente que pretende dar cuenta de los procesos de adquisición y representación del conocimiento, sustentado en un enfoque probabilístico-inductivo relacionado con los modelos de redes neuronales, pero basado en un valor de descomposición singular, el cual permite construir un espacio semántico latente que puede representar el contenido de grandes *corpora* textuales. A su vez estos investigadores plantean que la representación construida matemáticamente desde estos textos sería análoga a la representación cognitiva que hacen los seres humanos a partir de las experiencias lingüísticas a las que ellos están afectos durante su vida. De esta manera, plantean que el espacio semántico ideal se encuentra en la mente del individuo y que este espacio puede ser representado utilizando los textos que han permitido la formación de ese espacio semántico (Deerwester et al., 1990; Landauer & Dumais, 1997).

En lo que sigue a continuación, daremos a conocer los fundamentos propuestos por los investigadores para considerar al LSA como un modelo teórico, para más adelante especificar su funcionamiento y las aplicaciones en las cuales se ha utilizado. Cerraremos este apartado con una discusión crítica sobre los aspectos más relevantes con relación al LSA.

2.7.3.1 Perspectiva teórica del LSA

Landauer y Dumais (1997: 1) basándose en los resultados obtenidos a partir del análisis con LSI de grandes corpus textuales, plantean la posibilidad de pensar en lo que ellos denominan una “*Theory of acquisition, induction and representation of knowledge*”. Estos

autores dan inicio a los estudios psicolingüísticos rebautizando al LSI como Análisis Semántico Latente. En este sentido, Landauer, Foltz y Laham (1998: 260) argumentan que:

“Word and passage meaning representations derived by LSA have been found capable of simulating a variety of human cognitive phenomena, ranging from developmental acquisition of recognition vocabulary to word-categorization, sentence-word semantic priming, discourse comprensión, and judgments of essay quality”.

La teoría que postulan estos autores descansa en la noción de que algunos dominios de conocimiento contienen inmensos números de interrelaciones débiles o latentes, que si son aprovechadas se pueden amplificar produciendo aprendizaje a través de procesos de inferencia. El método de inducción propuesto depende de la reconstrucción de un sistema de relaciones de similitud múltiples en un espacio multidimensional. Se supone que la co-ocurrencia de eventos, en particular de palabras en contextos locales se generan y se reflejan por su similitud en algún lugar de este espacio multidimensional. Utilizando los métodos estadísticos referidos en Deerwester et al. (1990), se concluye que el LSA puede usarse para predecir fenómenos tales como la coherencia textual, comprensión, desambiguación contextual de homógrafos y generación del significado central inferido de un párrafo. Definen al LSA como un modelo teórico y un método para extraer y representar el significado contextual del uso de palabras a través de computación estadística aplicada a un gran corpus textual (Landauer & Dumais, 1997).

Lo interesante de este enfoque teórico es que permite pensar en una representación del conocimiento fundada en los textos mismos, es decir, se defiende una postura empírico-inductiva de la adquisición y representación del conocimiento. Al respecto, Landauer y Dumais (1997) mantienen que la propiedad inductiva del aprendizaje por el que las personas adquieren mucho más conocimiento del que parece estar disponible en la experiencia, es un verdadero misterio. Este es el conocido problema de "la pobreza del estímulo" o "problema de Platón".

La adquisición del lenguaje ha sido un problema ampliamente debatido, sin embargo, la explicación que ha predominado en la ciencia cognitiva es la que sugiere, siguiendo a Chomsky (1991), que la exposición por parte de los niños al lenguaje de los padres no

proporciona evidencia adecuada para explicar el hecho de que sean capaces de producir y entender construcciones sintácticas y léxicas nunca antes oídas. Por ello, Chomsky (1991) propone que los seres humanos adquieren las reglas sintácticas de manera independiente de la lengua en la cual están inmersos, gracias a un dispositivo de adquisición del lenguaje asentado en la genética humana. Pinker (1994), aunando evidencia obtenida desde las distintas perspectivas teóricas que estudian el origen del lenguaje (estudios evolutivos, historia, anatomía, fisiología y desarrollo del lenguaje), sintetiza este supuesto planteando que:

“[...] language learning must be based on a very strong and specific innate foundation, a set of general rules and predilections which need parameter-setting and filling in, but not acquisition as such, from experience.”

Al respecto, Deerwester et al. (1990: 395) plantean que:

“While this "language instinct" position is debatable as stated, it rests on an idea that is surely correct, that some powerful mechanism exists in the minds of children that can use the finite information they receive to turn them into competent users of human language. What we want to know, of course, is what this mechanism is, what it does, how it works. Unfortunately the rest of the instinctivist answers are as yet of limited help. The fact that the mechanism is given by biology or that it exists as an autonomous mental or physical "module" (if it does), tells us next to nothing about how the mind solves the basic inductive problem.”

Es, en este sentido, que los autores que defienden la postura inductiva del aprendizaje plantean que la manera en que la mente resuelve este problema inductivo básico es explicable a partir del mecanismo del LSA, que consiste en la acomodación simultánea de un número muy grande de relaciones de co-ocurrencia locales en un espacio multidimensional de una dimensionalidad adecuada. Hipotéticamente, el espacio multidimensional óptimo para la reconstrucción debe tener la misma dimensionalidad que la fuente que produce el discurso, esto es, el espacio semántico del hablante o escritor humano (Landauer, Foltz & Laham, 1998).

El supuesto subyacente del LSA es que las similitudes y diferencias en el significado de las palabras pueden, en gran medida, ser inducido desde las similitudes y diferencias que ocurren en el contexto del discurso. A su vez, similitudes en el significado de una unidad

lingüística mayor a la palabra pueden ser inducidas en gran medida desde la combinación (en sentido matemático) de las palabras que contiene. Este supuesto implica que el determinante del significado verbal que generalmente domina es la elección de palabras y la combinación de las palabras en la expresión, por lo tanto, para muchos propósitos el orden de las palabras en los pasajes puede ser ignorado en la estimación del significado con una mínima pérdida de exactitud. La pregunta respecto de si estos supuestos son correctos ha sido indagada y respondida afirmativamente, por medio de la evaluación de la habilidad de los modelos computacionales basados en estos supuestos para simular un amplio rango de fenómenos verbales humanos (Landauer & Dumais, 1996, 1997; Landauer, Foltz & Laham, 1998; Kintsch, 1998, 2000, 2001, 2002; Rehder, Schreiner, Wolfe, Laham, Landauer & Kintsch, 1998; Wolfe, Schreiner, Rehder, Laham, Foltz, Kintsch & Landauer, 1998; Landauer & Pstorka, 2000; Landauer, 2002; Quesada, Kintsch & Gómez, 2002; Quesada, 2003).

Cabe hacer notar aquí que, si bien en los estudios se demuestra que los resultados obtenidos con el LSA se correlacionan significativamente con los resultados mostrados por los seres humanos ante tareas similares de determinación de relaciones de significado, nada pueden plantear respecto de la adquisición de reglas sintácticas y menos respecto de los significados referenciales, dado que no incluyen variables correspondientes al conocimiento previo o, más aún, a variables innatas en la asignación del significado. Por otra parte, lo que se está comparando es el producto final de dos tipos de procesamiento que no son comparables directamente; esto es, un procesamiento matemático-estadístico realizado por una máquina computacional y el procesamiento cognitivo realizado por un ser humano, en el cual existe una multitud de variables que no son controladas o al menos emuladas en el procesamiento computacional.

Este modelo teórico no ha estado exento de crítica, en particular respecto de su plausibilidad psicológica para explicar el procesamiento cognitivo de los textos por parte de los seres humanos.

Entre quienes más han criticado el LSA como modelo teórico está Perfetti (1998), quien niega la posibilidad de que el análisis de co-ocurrencias pueda servir para dar cuenta de las relaciones semánticas que los seres humanos realizan. Este autor plantea que algo tan

general, que es indiferente a una amplia variedad de propiedades estructurales, definitorias para la cognición, es un pobre candidato para ser modelo de procesos de la cognición dentro de un dominio específico. Perfetti (1998) especifica que las principales fallas del LSA son la arbitrariedad del tamaño del corpus; esto debido a la imposibilidad de determinar con certeza una muestra ideal de textos en calidad y cantidad que sea similar a la que ha sido experimentada por los seres humanos. Por otra parte, critica la no inclusión de la sintaxis en el análisis semántico y las fallas denominadas "de equivocación", como por ejemplo, que para el inglés la palabra "*physician*" aparezca más relacionada la palabra "*nurse*" que "*doctor*", este tipo de errores los adjudica al procesamiento por co-ocurrencia. Sin embargo, y a pesar de lo anterior, considera que el LSA es una útil herramienta de investigación en psicolingüística.

Otros autores, como De Vega (2002), critican el LSA diciendo que es una mala teoría psicológica, ya que no tiene en cuenta la estructura sintáctica, ni la morfología, ni el conocimiento del mundo, ni las características funcionales de la memoria. Además, incorpora el problema de la "circularidad simbólica", según el cual los símbolos sólo se refieren a otros símbolos, esto implica un problema de "fundación" (grounding) del lenguaje con los objetos o eventos que son percibidos sin el lenguaje. Por último, destaca algunos errores de asignación de puntaje del LSA, de la misma manera como lo plantea Perfetti (1998), debido a que no toma en cuenta el orden de palabras.

En suma, el LSA como una teoría de adquisición y la representación del significado parece no considerar aspectos que han sido considerados relevantes para explicar este tipo de procesamiento cognitivo. A este respecto, el mismo Landauer (2002: 31) plantea a modo de aclaración que:

"If any of my presentations of LSA have given cause to believe that LSA is a to be considered a complete theory of language and knowledge, or even lexical semantics, I regret it profoundly. LSA is a theory of (about) those things, but not of everything about them".

Además, agrega que:

"LSA is, of course, incomplete as a theory of language, or even as a theory of verbal semantics. It includes no model of language production, or of the

dynamic processes of comprehension. Nor does it deal with discourse and conversation conventions, or with pragmatic factors in semantics. That no current theory is more complete, and none as able to model full-scale vocabulary learning, is no excuse. We need more. However, LSA provides a good base camp for further exploration. I give an example of an effective computation for some important aspects of the problem and opens up paths that were previously closed by incorrect assumptions” (Landauer, 2002: 28).

En este trabajo, la perspectiva en cuanto al LSA como modelo teórico concuerda en parte con lo expuesto por Landauer (2002). Consideramos que no es un modelo teórico capaz de explicar el procesamiento cognitivo que los seres humanos realizan. Esto, fundamentalmente, debido a que no consideramos plausible explicar ningún procesamiento cognitivo a partir de un modelo matemático-estadístico, aunque los resultados obtenidos utilizándolo se correlacionen con los resultados humanos. En definitiva, la similitud en el producto no implica similitud en el proceso. Por otra parte, las representaciones matemáticas realizadas por la reducción de dimensionalidad a partir del material lingüístico explícito de los textos (palabras) no dan cuenta de la profunda interrelación existente entre el texto, el contexto sociocultural y la mente del lector.

Sin embargo, los resultados obtenidos a partir de esta técnica nos hacen pensar en que puede ser muy útil en la descripción de las relaciones semánticas de tipo asociativo, en particular de las relaciones de similitud semántica basadas en la cuantificación de las co-ocurrencias colocacionales de las palabras en los textos. Específicamente, su valor reside en la capacidad que tiene, como implementación computacional, para ser utilizada con grandes cantidades de texto. En suma, nuestro interés por el LSA se sustenta en la funcionalidad práctica que tiene como herramienta (implementada computacionalmente) para la identificación de relaciones semánticas no guiadas por la sintaxis, asumiendo que con ello no se puede explicar el funcionamiento cognitivo de los individuos, pero que sí permite obtener resultados similares a los obtenidos por los seres humanos ante tareas de similitud léxico-semántica.

En lo que sigue, daremos a conocer el funcionamiento del LSA como técnica matemático-estadística para identificar cuantitativamente la similitud semántica entre palabras, entre palabras y párrafos y entre párrafos.

2.7.3.2 Funcionamiento del LSA

Landauer, Foltz y Laham (1998) plantean que el LSA es una técnica matemático-estadística totalmente automática para extraer e inferir relaciones del uso contextual esperado de palabras en pasajes de texto. No es un procesamiento de lenguaje natural tradicional o programa de inteligencia artificial (de tipo simbólico), no usa ningún diccionario construido humanamente, bases de conocimiento, redes semánticas, gramáticas, segmentadores sintácticos, o morfologías y toma como datos de entrada sólo la segmentación del texto en palabras, pasajes, frases o párrafos.

El LSA no usa ninguna información lingüística previa o conocimiento perceptual, esto significa que está solamente basado en un método de aprendizaje matemático general que logra efectos inductivos poderosos, extrayendo un adecuado número de dimensiones para representar objetos y contextos.

En definitiva, el LSA extrae sus representaciones de significado a partir de palabras y párrafos, es decir, exclusivamente del análisis matemático-estadístico del texto. Nada de su conocimiento viene desde la información perceptual sobre el mundo físico, del instinto, o de la experiencia generada por funciones corporales, sentimientos y/o intenciones. Así, su representación del significado es parcial y limitada, puesto que no hace uso del orden de palabra, así como tampoco de relaciones sintácticas o lógicas, o de la morfología. Sobre este punto, Landauer (2002) explica que el 80% de la información potencial en el lenguaje está en la elección de palabras sin tener en cuenta el orden en el que ellas aparecen. Esto es lo que Jurafsky y Martin (2000) denominan métodos de “depósito de palabras” (*bag of words*).

Junto con esta idea de representación, sin sintaxis, aparece la idea de que en estas grandes cantidades de *corpora* existen interrelaciones semánticas débiles entre palabras que son potenciadas por el método de reducción de dimensiones SVD.

Un buen ejemplo de cómo se potencian y reconocen estas relaciones semánticas latentes en los textos es considerar lo siguiente (Landauer, 2002):

ecks + wye+ aye = foo

ecks + wye + bie = foo

Donde ecks y wye co-ocurren en el mismo contexto, en tanto aye y bie no lo hacen. En conjunto ambas ecuaciones implican que aye y bie debieran tener el mismo significado, pero no implican nada de la relación entre ecks y wye. Así, podemos plantear que el hecho de que las palabras aparezcan juntas en un mismo contexto no determina que tengan un mismo significado.

Ahora si adicionamos dos ecuaciones más:

ecks + wye + aye = foo

ecks + wye + bie = foo

ecks + wye + cee = bar

ecks + wye + dee = bar

Sabemos que cee y dee son sinónimos. Finalmente, consideremos:

aye + cee = oof

bie + dee = rab

Siendo consistentes con las ecuaciones anteriores, en las cuales aye=bie y cee=dee, podemos determinar que estos dos pasajes tienen que tener el mismo significado (oof=rab) aun cuando no exista ninguna palabra en común. De esta misma manera es como se computan los datos de los textos para determinar la latencia semántica y sus relaciones en los corpus de textos.

En suma, la metáfora que subyace al término “latente” es que por medio de la reducción de dimensionalidad que se realiza usando el SVD, se obtiene una buena representación de las relaciones existentes entre las palabras en un corpus textual, en el cual estas relaciones son muy débiles debido al gran número de palabras. De este modo, el LSA recupera en gran medida las relaciones semánticas originales (en términos de dimensiones) que representan

mejor los agrupamientos (los que se podrían interpretar como tópicos) de palabras y documentos.

El procedimiento que se lleva cabo al utilizar el LSA para la representación de los textos en espacios semánticos multidimensionales es el siguiente:

El primer paso es construir, a partir de los textos, una matriz de co-ocurrencias en la cual cada fila representa a un documento o cotexto (generalmente párrafos) (d) y cada columna representa una palabra del texto (t). Cada celda contiene la frecuencia con la cual la palabra de la fila aparece en el pasaje de texto denotado por su columna (t, d). La entrada en la celda está sujeta a una transformación doble por la cual cada celda es “pesada” por una función que expresa tanto la importancia de un pasaje particular de texto como el grado por el cual el tipo de palabra transporta información en el dominio del discurso en general. Esto quiere decir, que la frecuencia de la palabra en cada celda es convertida en su logaritmo (log). Luego, se calcula para cada palabra una medición de la teoría de información, entropía, esto es $-p(x) \log p(x)$ para todas entradas en la fila, después cada entrada en la celda de la matriz es dividida por el valor entrópico de cada fila. Esto implica, en palabras de Landauer, Foltz y Laham (1998: 276), que:

“The effect of this transformation is to weight each word-type occurrence directly by an estimate of its importance in the passage and inversely with the degree to which knowing that a word occurs provides information about which passage it appeared in”.

La importancia de esta doble transformación es que permite representar correctamente un pasaje de texto como una combinación de las palabras que contiene, porque se enfatizan las palabras portadoras de significado específico, reduciendo la influencia de términos que ocurren en una gran cantidad de documentos y enfatizando aquellas que están más asociadas con un tipo particular de documentos.

El segundo paso, luego de esta doble transformación, es aplicar a la matriz resultante una procedimiento denominado descomposición en valor singular (SVD). El SVD descompone la matriz rectangular (aquella que considera diferentes entidades en las filas y en las

columnas, por ejemplo, términos y documentos) en el producto de otras tres matrices, esto es en una matriz de menores dimensiones, pero que representa a la matriz original.

$$X = \{W\}\{S\}\{P\} = X'$$

Donde: X = matriz original de términos por documentos ($t \times d$)

$\{W\}$ = matriz de los valores singulares de los términos (filas)

$\{S\}$ = matriz diagonal de valores singulares (esta matriz es la que da el peso de cada una de las dimensiones)

$\{P\}$ = matriz de valores singulares de los documentos (columnas)

X' = matriz resultante, que contiene menos dimensiones y que representa a la original.

Como plantea Landauer, Foltz y Laham (1998: 266):

“LSA applies singular value decomposition (SVD) to the matrix. This is a form of factor analysis, or more properly the mathematical generalization of which factor analysis is a special case. In SVD, a rectangular matrix is decomposed into the product of three other matrices. One component matrix describes the original row entities as vectors of derived orthogonal factor values, another describes the original column entities in the same way, and the third is a diagonal matrix containing scaling values such that when the three components are matrix-multiplied, the original matrix is reconstructed. There is a mathematical proof that any matrix can be decomposed perfectly, using no more factor than the smallest dimension of the original matrix. When fewer than the necessary number of factors are used, the reconstructed matrix is a least-squares best fit. One can reduce the dimensionality of the solution simply by deleting coefficients in the diagonal matrix, ordinarily starting with the smallest. (In practice, for computational reasons, for very large corpora only a limited number of dimensions—currently a few thousand— can be constructed.)”

Para propósitos explicativos es útil interpretar la SVD en términos geométricos. Esto significa que los valores de las filas y columnas de la matriz reducida son tomados como coordenadas de puntos que representan los documentos y los términos en un espacio multidimensional de k -dimensiones (donde k significa que las dimensiones son menores y representan a las dimensiones originales de la matriz de co-ocurrencias entre palabras y

documentos, esto es n -dimensiones). La cantidad de dimensiones (k) es la dimensionalidad a usar en la construcción del espacio multidimensional creado por la SVD. Normalmente, las dimensiones a utilizar están entre 50 y 400 dimensiones (Landauer, Foltz y Laham, 1998). De esta forma, cada vector de un texto o documento (columna en la matriz) es visto como un vector espacial en el espacio de k -dimensiones. La similitud entre vectores es calculada usando medidas de coseno, cuyos valores van de 1 para vectores con la misma dirección (esto significa que lo medido es igual) a 0 para aquellos vectores ortogonales (perpendiculares en el espacio multidimensional, es decir, que lo medido es completamente distinto). Teóricamente, los valores de coseno podrían alcanzar hasta -1 , pero porque la información original de co-ocurrencias va de 0 al número de ocurrencias de un término en un documento, los valores de coseno nunca bajan mucho de 0 (Wiemer-Hasting, 2004). Los valores deben ser normalizados, de modo de hacer más efectiva la comparación entre ellos, ya que sino se hace, vectores más largos (correspondiente a documentos más largos) podrían tener una ventaja injusta respecto de los vectores más cortos. Además, la normalización de los valores de coseno permite que estos sean calculados como un producto simple (multiplicación de los vectores).

Otra medida, que a menudo es útil, es "el largo del vector". El largo del vector nos dice cuánta información tiene el LSA sobre tal o cual vector. Así, el largo del vector de la frase es generalmente mayor que el de la palabra, y el largo de vector del párrafo es aún mayor. Las palabras sobre las que el LSA sabe mucho (porque aparecen frecuentemente en el corpus de entrenamiento, en muchos contextos diferentes) tienen longitudes de vector mayores que de aquellas que el LSA no conoce bien. En síntesis, el largo del vector es un indicador de la probabilidad por la que una palabra, o un conjunto de palabras (párrafos) se relacionan con otras en el análisis semántico latente (Kintsch, 2001). Sin embargo, para tener un valor de comparación de similitudes adecuado las cantidades de palabras y/o textos debieran ser homogéneas, lo cual no necesariamente ocurre cuando se trabaja con textos naturales (tal como se usan en la realidad). Es por ello que normalmente se recomienda usar el valor de coseno normalizado. Al respecto, Manning y Schütze (2003: 300) plantean:

“The *cosine* [...] penalizes less in cases where the number of non-zeros entries is very different.[...]. This property of the cosine is important in Statistical NLP since we often compare words or objects that we have

different amounts of data for, but we don't want to say they are dissimilar just because of that".

En cuanto a la importancia de la dimensionalidad del espacio, Wiemer-Hasting (2004) plantea que la dimensionalidad escogida para representar la matriz original está correlacionada con la ocurrencia de los términos en la matriz original, y que partir de ella se puede segmentar el espacio semántico en un buen número de categorías subsimbólicas que pueden ser combinadas significativamente.

Para calcular el número de dimensiones en términos prácticos, Wiemer-Hasting (2004) propone elegir entre 200 a 400 dimensiones; sin embargo, plantea que mientras mayor cantidad de dimensiones se elija mayor cantidad de documentos en el corpus de entrenamiento debe haber. Otra propuesta del mismo autor es elegir entre 1/30 y 1/20 del número de documentos que se tienen en el corpus.

Ahora bien, para calcular los valores de similitud a partir de los vectores en este espacio de, por ejemplo, 300 dimensiones, se utiliza el cálculo del producto entre los valores singulares de la matriz reducida. De esta manera, se pueden efectuar comparaciones entre dos términos, entre dos documentos y entre un término y un documento (Deerwester et al. 1990).

Para la comparación entre dos términos se calcula el producto entre dos vectores correspondientes a dos filas de la matriz X' , las cuales reflejan la extensión (cada celda de término por documento) por la cual dos términos tienen un patrón similar de ocurrencias a través de los documentos. El análisis para la comparación entre dos documentos es similar, excepto que en este caso es el producto entre dos vectores correspondientes a dos columnas de la matriz X' . Esto nos permite distinguir qué documentos tienen perfiles de términos similares (es decir, qué términos constituyen los documentos). En cuanto a la comparación entre un término y un documento, la comparación es diferente a las anteriores; en este caso, el valor de comparación está dado por el valor individual de cada celda en la matriz X' .

A continuación se presentan algunos ejemplos de comparaciones obtenidas utilizando uno de los espacios semánticos construidos en esta investigación. El espacio sobre el cual se harán las comparaciones es el espacio semántico de artículos de investigación científica

(ES-ARTICO). Este está construido sobre la base de 675 artículos de investigación científica de tres áreas científicas: **Ciencias exactas:** Ciencias químicas, Física e Ingeniería química. **Ciencias biológicas:** Ciencias veterinarias, Zoología y Oceanología, Ciencias morfológicas (Anatomía), Ecología y subdisciplinas de la ecología, Infectología pediátrica y/o de adultos y Microbiología. **Ciencias sociales:** Ciencias de la comunicación, Ciencias de la información documental, Antropología social y cultural, Arqueología, Economía, Sociología, Políticas culturales y sociales.

En términos de palabras, el total de palabras es de 3.631.124. El espacio semántico (matriz reducida) resultante tiene un total de 60.229 palabras únicas (types) y se consideraron 294 dimensiones en su construcción.

a) Comparación entre términos:

Tabla 3. Ejemplo de comparación entre palabras según ES-ARTICO

Término 1	Término 2	Valor	Similitud
Química	Física	0,784856	Alta
Cabeza	Cuerpo	0,7460052	Alta
Química	Documentación	-0,0116183	Baja
Cuerpo	Agua	0,0083383	Baja

En el ejemplo presentado en la Tabla 3 se han seleccionado dos pares de palabras que al ser comparadas en el espacio semántico ES-ARTICO demuestra tener alta similitud y baja similitud respectivamente. Estos resultados también se pueden corroborar desde un punto de vista intuitivo. En el caso del primer par de palabras estas corresponden al nombre de dos disciplinas muy próximas del área de las ciencias exactas. En el segundo par de palabras se observa una relación metonímica (de parte-todo). En el tercer par de palabras, si bien ambas palabras pueden designar disciplinas científicas, no comparten el área ni su posible intercambio en algún contexto lingüístico. En el último par de palabras, la relación es entre

dos elementos que no parecen tener mayor relación (alguien podría especular que hay una relación en términos de que el agua es un elemento importante para el cuerpo, pero no es una relación que se haya establecido en los textos, o al menos reconocido con el método).

b) Comparación entre documentos:

Tabla 4. Ejemplo de alta similitud semántica entre 2 textos

Documento 1	Documento 2	Valor	Similitud
A principios del siglo XXI, el paradigma de integración económica dominante induce la agrupación de países afectados por profundas asimetrías. Este esquema obliga a prestar atención a las medidas de tratamiento especial y diferenciado adoptadas para enfrentar las desigualdades entre naciones en distintos acuerdos de integración económica. Este artículo presenta los rasgos principales del paradigma y las razones que lo originaron; compara el tratamiento especial en distintos acuerdos y las críticas suscitadas; y analiza las propuestas planteadas en las negociaciones del ALCA y discute sus posibilidades dentro del cuadro de fuerzas regionales. Implicaciones para el ALCA.	En este artículo se analizan las razones que explican el surgimiento de la tendencia a asociar en acuerdos de integración a países en vías de desarrollo y desarrollados; en segundo lugar, se discuten los instrumentos de tratamiento especial y diferenciado utilizados para mitigar las asimetrías entre naciones dentro de esos acuerdos, y finalmente se evalúan las propuestas hechas en las negociaciones del ALCA.	0.81452703	Alta

Tabla 5. Ejemplo de baja similitud semántica entre dos textos

Documento 1	Documento 2	Valor	Similitud
<p>La vía aérea ha sido propuesta como modelo de diseño óptimo desde una perspectiva física. Su diseño se ha asociado con un adecuado flujo de gases a los alvéolos, una mínima producción de entropía y un mínimo costo en materia y energía. Se ha propuesto un decrecimiento exponencial del diámetro de los bronquios (dG) en función de la generación: $dG = d_0 \cdot 2^{-G/3}$, asociado a una mínima producción de entropía. También se ha propuesto un modelo de renormalización: $dG = A_n \cdot G^{-u}$ donde u es un exponente y A_n una función que introduce desviaciones periódicas en la escala, es decir más de una escala, evitando la propagación distal de errores aleatorios en el calibre de un bronquio. Sin embargo, este último resultado podría ser consecuencia en árboles asimétricos de la relación entre el diámetro y el orden del bronquio y no de la generación. En este trabajo estudiamos la asimetría y el</p>	<p>Las especies animales muestran una gran diversidad en su diseño estructural a consecuencia de su adaptación a diferentes condiciones de vida. Esta observación sugiere preguntas como ¿está su diseño optimizado? o ¿los animales presentan un diseño económico con no más estructura que la necesaria para llevar a cabo su función?.</p>	0.2383657	Baja

<p>decrecimiento del diámetro bronquial en dos especies. Se utiliza el modelo de Zamir como un sistema externo de medida de la optimización. Encontramos una clara asimetría del árbol bronquial. Comprobamos que la relación exponencial diámetro-orden es siempre muy buena ($R^2 \approx 0,8$) y que en cambio la relación exponencial diámetro-generación es menos clara ($R^2 < 0,6$). La supuesta modulación armónica desaparece al considerar el orden y no la generación. Se determina un alto grado de optimización de la estructura en las dos especies.</p>			
---	--	--	--

En los ejemplos entregados de comparación de documentos la comparación presentada en la Tabla 4 muestra una alta similitud, en tanto que en la comparación presentada en la Tabla 5 la similitud es baja. Este resultado está influenciado tanto porque existe en la primera comparación un mayor número de palabras iguales, tanto como porque hay mayor relación temática entre ambos textos, esto es, el resumen y el último párrafo de la introducción de un artículo del área de ciencias sociales (NS7-2003 186: 1). Recordemos que, como plantea Swales (1990), en el último párrafo de la introducción se adelantan los temas que serán tratados en el artículo. En cambio, en la comparación 2 los textos presentan menor cantidad de palabras comunes o relacionadas y porque la relación temática general es muy baja, en este caso, resumen y 1er párrafo de un artículo del área de biología (RCHN1-2002 72(2): 1). Siguiendo a Swales (1990), el primer párrafo tiene la función de establecer el territorio

de la investigación, lo cual no necesariamente está en directa relación temática con el contenido del artículo.

C) Comparación entre palabra y documento

Palabras	Documento	Valor
Triafulveno	El triafulveno H ₂ C ₃ -CH ₂ es el compuesto más simple que presenta conjugación cruzada no alternante. En el presente trabajo se expone un estudio teórico Ab Initio a nivel Hartree-Fock (HF) de la polarizabilidad molecular de triafulveno y análogos del tipo C ₃ H ₂ -X, donde X = O y S. Adicionalmente se evaluaron los efectos de correlación electrónica empleando Teoría de Perturbaciones de segundo orden Moller-Plesset (MP2) y Funcional de la densidad con el funcional híbrido BLYP. Los cálculos fueron realizados con estructuras optimizadas siguiendo restricciones de simetría C _{2v} y con los conjuntos base STO/6-31+G(d,p), STO/6-31+G(3d,3p) y el conjunto base de Sadlej. Los resultados de propiedades estructurales y electrónicas obtenidos para el triafulveno (metilenciclopropeno) reproducen los valores reportados en la literatura. La inclusión de correlación electrónica a nivel MP2 tiene importantes efectos sobre la polarizabilidad promedio dependiendo del conjunto base con el cual se evalúe. La polarizabilidad obtenida mediante la Teoría del Funcional de la Densidad mostró la misma tendencia indicando la importancia del conjunto base en el cálculo de esta propiedad.	0.6252556
Polarizabilidad		0.6369575
Momento bipolar		0.28118393
Ab initio		0.30547398
DFT		0.41627967

En el ejemplo se han tomado las palabras clave de un artículo y se han comparado con el resumen del mismo. Se observa que las dos primeras palabras tienen mayor similitud que las tres últimas, siendo ‘momento bipolar’ la que menos similitud presenta con el resumen. Básicamente, las dos primeras palabras presentan un valor mayor porque se repiten con mayor frecuencia en el texto. Cabe hacer notar que DFT tiene mayor valor que Ab initio, a pesar de que Ab initio aparece una vez en el texto y DFT ninguna, sin embargo, aparece la expresión “Teoría del Funcional de la densidad” y “Funcional de la densidad” las cuales

están en directa relación por abreviatura con DFT, pudiendo ser opciones léxicas en un contexto léxico particular.

2.7.3.3 Aplicaciones del LSA

En este apartado daremos a conocer algunas de las aplicaciones que se le han dado al análisis semántico latente, particularmente desde la perspectiva educacional ligada al estudio de la interacción texto-lector.

Por una parte, Foltz, Kinstch y Landauer (1998) han utilizado el LSA para evaluar los escritos de los alumnos, midiendo el grado de coherencia interna de los textos. Ellos estiman que el LSA se puede utilizar como una técnica para medir la coherencia de los textos, comparando los vectores para dos segmentos inmediatos de texto en un espacio semántico multidimensional. El método proporciona una caracterización del grado de relación semántica entre los segmentos. Los resultados indican que el método puede predecir el efecto de coherencia del texto en la comprensión, de esta manera, el LSA puede aplicarse como un método que produce predicciones de coherencia, similar al modelamiento proposicional. Ellos describen estudios adicionales que investigan la aplicación del LSA en el análisis de la estructura del discurso y examinan el potencial del LSA como modelo psicológico de efectos de coherencia en la comprensión de textos.

Cabe señalar que la noción de coherencia está usada aquí en relación con las relaciones de significado internas del texto y no como hemos planteado en 3.5.3, como relaciones de significado que ocurren en la mente del lector.

Kintsch (2001, 2002) continúa con la labor de utilizar la potencialidad del LSA en la creación y evaluación de resúmenes y de coherencia para el análisis de Predicados y Macroproposiciones con el fin de que el LSA se convierta en una herramienta adecuada para su teoría de Construcción-Integración (Kintsch, 1998, 2000, 2001, 2002). En este sentido, propone un algoritmo en el cual se considera el potencial del LSA para la identificación automática de proposiciones. Por otra parte, destacan los trabajos de Kintsch (2001) y Kintsch y Bowles (2002), sobre temas como el uso del LSA para el análisis

interpretativo de metáforas, los estudios sobre inferencias causales y la desambiguación homonímica.

Por otra parte, Rehder, Schreiner, Wolfe, Laham, Landauer y Kintsch (1998) confirman las investigaciones de Wolfe, Schreiner, Rehder, Laham, Foltz, Kintsch y Landauer (1998) en donde establecen que el LSA puede usarse para evaluar el conocimiento del estudiante, es decir, cómo los ensayos pueden ser categorizados según su dificultad por el LSA y cómo el LSA puede determinar los textos instruccionales más apropiados para cada estudiante. Ellos compararon ensayos escritos por estudiantes con uno o más de los textos instruccionales del corpus, en términos del coseno entre la representación del vector del ensayo del estudiante y el texto instruccional en cuestión. Este simple método fue usado eficazmente, según ellos, en determinar cuál de los textos del corpus permitiría a los estudiantes aprender más. Establecen, por una parte, que el vocabulario técnico no afecta en la evaluación del conocimiento del estudiante. Por otra, que la longitud del ensayo no es problema en la valoración del conocimiento. En definitiva, el LSA sirve como predictor de aprendizaje de los estudiantes a partir de un texto determinado.

Otra aplicación interesante de destacar es el programa denominado *Asesor Inteligente de Ensayos*, que varias universidades estadounidenses y una escuela elemental de Nuevo México están probando, como software experimental diseñado para evaluar a estudiantes que escriben ensayos (Foltz, Landauer & Laham, 1998). La función del programa es evaluar ensayos tan rápido como las pruebas de selección múltiple con el objetivo de, finalmente, reemplazarlas. Los estudiantes envían su ensayo vía internet y reciben retroalimentación de su trabajo casi inmediatamente, también vía internet. Usando este programa, afirman los autores, los estudiantes aumentaron sus calificaciones en aproximadamente tres repeticiones.

En las evaluaciones del programa, el sistema ha demostrado ser tan fiable como los maestros humanos en evaluar el texto del estudiante. Los investigadores establecen que el software concuerda con humanos tan a menudo como estos están de acuerdo entre sí (entre 80 y 90 por ciento del tiempo). El *Asesor Inteligente de Ensayo* es totalmente automatizado y puede enseñarse a sí mismo sobre un asunto particular, a partir de textos, sin ninguna entrada humana de datos. Los investigadores especulan que el sistema podría, teóricamente,

aprender sobre cualquier asunto, aunque hasta ahora su conocimiento se restringe a las áreas de psicología, medicina e historia, entre otras (Foltz, Landauer & Laham, 1998).

Otro programa innovador que utiliza LSA en el área de la educación es el AutoTutor, desarrollado a fines de los años noventa por Arthur Graesser y un amplio equipo de investigación multidisciplinar de la Universidad de Memphis, Estados Unidos. El supuesto principal que está detrás de este desarrollo computacional es que la forma más poderosa para construir conocimiento es a través de un método tutorial cara a cara. De hecho la experiencia dice que el aprendizaje por tutoría humana es en extremo efectivo en comparación al ambiente tradicional de clase (Graesser, Person, Harter & Tutoring Research Group, 2001). Así, este equipo desarrolló un tutor computacional que simula los patrones de discurso y estrategias pedagógicas de un tutor humano típico. Las tareas fundamentales que desarrolla este AutoTutor son: hacer preguntas y plantear problemas, comprender las respuestas (escritas) del estudiante, retroalimentar las respuestas por medio de verbalizaciones y gestos, plantear sugerencias sobre información específica, agregar información perdida, corregir errores sobre algún tópico, entre otras. Para su funcionamiento, este tutor virtual consta de 7 módulos interrelacionados, a saber, un módulo que permite llevar a cabo la extracción de lenguaje, un clasificador de actos de habla, un módulo de LSA, un plan curricular, una interfaz visual (un rostro con movimientos gestuales), un administrador de diálogo y un módulo selector de problemas.

El uso del LSA en este contexto está en relación, como ya hemos visto, con el análisis de grandes cantidades de textos que, en este caso, se utilizan para representar el conocimiento de mundo que tiene el programa y evaluar las respuestas de los estudiantes mediante las similitudes semánticas existentes entre las respuestas ideales contenidas en el programa y las otorgadas por los estudiantes. En suma, según el equipo de investigadores del proyecto AutoTutor, el LSA puede evaluar la calidad de las respuestas de los estudiantes tan bien como lo haría un tutor humano (Graesser et al., 2001). Sin embargo, plantean que el LSA no tiene la capacidad de comprender textos a un nivel profundo. Es, esencialmente, un "comparador de textos" que utiliza un modelo basado en conocimiento estadístico.

Según Graesser, Wiemer-Hastings, Wiemer-Hastings, Person, Harter y TRG (2000), el LSA hace un trabajo impresionante evaluando las similitudes entre las contribuciones del

estudiante y las respuestas esperadas, asignando a estudiantes con mucha habilidad altos valores de similitud y bajos valores de similitud a respuestas erróneas. Sin embargo, plantean que el LSA no está equipado para considerar el orden de palabras, la sintaxis, las expresiones lógicas, la cuantificación, las negaciones, las relaciones retóricas entre las cláusulas, y otros componentes analíticos de comprensión. Para realizar estos procesos analíticos se requiere de arquitecturas simbólicas más tradicionales en inteligencia artificial y se necesita mayor desarrollo en lingüística computacional (Graesser et al., 2001). La apuesta de estos investigadores está en la combinación híbrida entre el LSA y sistemas simbólicos tradicionales, aunque como explican, el LSA contribuiría mejor que el sistema simbólico a los desafíos que presenta, por ejemplo, el discurso conversacional (Graesser et al., 2001).

Otra aplicación interesante de destacar, llevada a cabo por Graesser, Hu, Olde, Ventura, Olney, Louwerse, Franceschetti y Pearson (2002), es el proyecto HURRA (*Human Use Regulatory Affairs Advisor*) el cual es un programa que enseña a los oficiales militares acerca de los aspectos éticos en las investigaciones con sujetos humanos. Este programa funciona sobre una plataforma web (en internet) con un sistema de consulta que permite acceder documentos, utilizando LSA, a través preguntas realizadas en lenguaje natural.

2.7.3.4 Críticas y Proyecciones

Durante este capítulo, hemos presentado distintos aspectos de los estudios relacionados con el significado. Particularmente, hemos planteado como opción de estudio que el significado de los textos se puede investigar desde una perspectiva léxico-semántica, basándonos en una orientación empírica en la cual el significado de las relaciones semánticas entre palabras es accesible por medio del estudio de las asociaciones colocacionales de las mismas. Para la concretización de este enfoque hemos propuesto dos modelos complementarios, los cuales se basan en las frecuencias y co-ocurrencias de las palabras en los textos.

Como ya sabemos, el estudio de frecuencias, si bien nos puede portar una amplia descripción de la constitución léxica de los textos a estudiar, plantea algunos problemas en cuanto a la posibilidad de estudiar cuantitativamente y en detalle las relaciones entre

palabras. Esto, fundamentalmente, porque no se pueden hacer generalizaciones basándose exclusivamente en las listas de palabras, en tanto que requieren un análisis cualitativo *in profundis* de cada una, lo cual tornaría la investigación en una empresa interminable. Insistimos en que es un primer nivel que aporta mucha información sobre las palabras de los textos, pero, a su vez, escasa, respecto de las relaciones de similitudes entre las palabras y los párrafos y entre los párrafos en los textos.

En cuanto al LSA, podemos establecer que, como modelo teórico, no proporciona una explicación acabada del proceso de adquisición y representación del significado. Esto, debido a que no es posible explicar ninguno de estos procesos cognitivos utilizando solo asociaciones de palabras en grandes corpus de textos. Los procesos cognitivos humanos no solo funcionan gracias a la experiencia con el lenguaje escrito, sino también gracias a los sistemas de memoria y de atención que no son considerados en los análisis matemático-estadísticos realizados con LSA.

Por otra parte, las explicaciones dadas por los teóricos del LSA se fundan en resultados basados en los datos entregados por el sistema computacional y su correlación con los resultados obtenidos por seres humanos, lo cual solo da cuenta de la actuación de esos seres humanos ante aquellas tareas, la extrapolación de esos resultados a una explicación de cómo funciona la mente es más bien especulativo, puesto que no se toman en cuenta todas las variables del procesamiento cognitivo humano ante esas tareas. Como dijimos más arriba, la similitud en el producto no implica similitud en el proceso. Por otra parte, las representaciones matemáticas realizadas por reducción de dimensionalidad a partir del material lingüístico explícito de los textos (palabras) no dan cuenta de la profunda interrelación existente entre el texto, el contexto sociocultural y la mente del lector.

En términos más particulares, el LSA aporta datos importantes para explicar la comprensión a partir de la extracción de sentido de los textos, aunque todavía no proporciona soluciones a los problemas de la sintaxis, las expresiones lógicas, la cuantificación, las negaciones, las relaciones retóricas entre las cláusulas, el reconocimiento de los antónimos y otros componentes analíticos de la comprensión, razón por la cual pierde capacidad explicativa (Landauer, 2002).

Con el objetivo de responder a las críticas, sobre todo las que dicen relación con la sintaxis, hay investigaciones que, por medio del marcaje morfosintáctico de los textos, han incorporado estos aspectos (Wiemer-Hastings, 2000; Wiemer-Hastings & Zipitria, 2001; Kanejiya, Kumar & Prasad, 2003). En función de estos resultados, se ha afirmado que la contribución de la sintaxis al significado de la oración no es desatendible, ya que entrega información adicional significativa para el cálculo de la similitud de las diferentes unidades lingüísticas (Wiemer-Hasting & Zipitria, 2001).

En cuanto herramienta, como hasta ahora la conocemos, está afecta a varios aspectos que pueden ser problemáticos. Por una parte, nos encontramos con que el concepto de similitud semántica sobre el cual se basa el LSA es en realidad un concepto estadístico, en donde lo semántico es entendido en función de la correlación de las colocaciones de las palabras en los textos. Además, esta correlación colocacional de palabras es entendida en ocasiones como sinonimia, lo cual no necesariamente es así, ya que el hecho de que una palabra ocupe el lugar de otra palabra en un contexto verbal similar no indica necesariamente sinonimia, sino que puede indicar cualquier relación léxico-semántica: antonimia, hiperonimia, hiponimia, meronimia, etc. Esto último, nos permite afirmar, además, que si bien la similitud semántica permite calcular un índice de relación entre dos palabras, no establece el tipo de relación existente, la cual debe ser identificada por el analista.

Con relación a este último aspecto, dado que no se considera la sintaxis en el procesamiento, las relaciones semánticas a nivel del enunciado no pueden ser establecidas, al menos en los términos expresados en el apartado 3.5.2. Por lo mismo, al comparar documentos, la determinación de las relaciones debe ser interpretada completamente por el investigador a partir de otros análisis complementarios.

Otra noción que aún no está bien definida es la de dimensión. Prueba de ello es lo que Wiemer-Hasting (2004: 4) plantea respecto de cómo se conceptualiza la idea de dimensión:

“No one really knows. The only thing that we do know is that the highest ranked dimension is closely correlated with the frequency of occurrence of terms. The intuition is that the dimensions somehow segment the semantic space into a good number of sub-symbolic categories that can be meaningfully combined. But that’s all speculation”.

Esto, sin lugar a dudas, es un aspecto sobre el cual se debe trabajar con mayor profundidad desde una perspectiva teórica, puesto que desde un punto de vista práctico esta noción se asocia a la idea de que cada fila o columna en la matriz reducida conforma un vector que juntos a los demás vectores forman un espacio multivectorial (desde la perspectiva geométrica) o multidimensional (desde la perspectiva algebraica). Al respecto, Hair, Anderson, Tatham y Black (1999) plantean que, si bien el análisis multidimensional tiene la ventaja de reducir la influencia del investigador al no exigir la especificación de las variables que se utilizan en la comparación de objetos, también tiene la desventaja de que el investigador no está realmente seguro de qué variables está utilizando el encuestado para realizar las comparaciones, siendo este otro argumento para aseverar que el LSA como teoría es insuficiente para explicar el fenómeno humano de la adquisición y representación del conocimiento.

Otro aspecto interesante de problematizar es la cantidad de texto necesario para construir espacios semánticos útiles para extrapolar resultados confiables. En este sentido, se ha planteado que mientras más cantidad de textos, mejor; sin embargo, dado que ante una investigación con esta herramienta se debe tener en cuenta el ámbito discursivo (o dominio) respecto del cual se quiere llevar a cabo los análisis de relación semántica, la cantidad y el tipo de textos no siempre son fáciles de adquirir. Si bien la creación de corpus electrónicos de textos es un trabajo que ha tenido grandes avances en los últimos tiempos, el acceso, el costo asociado y, particularmente en nuestro caso, el idioma en el cual se requieren (español) son factores que dificultan la creación de espacios semánticos con grandes cantidades de palabras y documentos. Un aspecto asociado a esto es la capacidad de procesamiento del computador en el cual se procese el espacio semántico. En este sentido, existe un costo importante en términos de disponer de un computador capaz de procesar tanta información.

Cabe señalar, que las investigaciones en idiomas distintos al inglés usando el LSA son escasas. Se conocen algunos trabajos realizados usando LSI o LSA en estudios que involucran el uso de corpora en lenguas romances, por ejemplo, para el francés se conocen los trabajos de Dumais, Landauer y Littman (1996), Dessus (2000), Lemaire, Bianco, Silvestre y Noveck (2001); para el portugués (Moreira & Huyck, 2002), para el español

(Serafín, Di Eugenio & Glass, 2003; Pérez, Gliozzo, Strapparava, Alfonseca, Rodríguez y Magnini, 2005). Pero se hace necesario mencionar que los trabajos en español han usado pequeños corpus en español y han sido publicados en inglés. El primer estudio utiliza dos corpus orales: DIAG (607 palabras distintas) y CallHome en español (12.066 palabras diferentes) y el segundo es:

“It is a small corpus composed by 1.929 *Student Answers* collected in an Operating Systems course. They have been automatically translated from Spanish to English by using Altavista Babelfish” (Pérez et al, 2005: en prensa)

En definitiva, dado que en español existen muy pocos estudios que utilicen LSA, el umbral de la cantidad de textos, representados por sus palabras diferentes (*types*) que asegure la confiabilidad de los resultados, aún no se conoce claramente y sigue siendo un aspecto que debe ser solucionado empíricamente.

La utilización del LSA como herramienta de investigación requiere de una disposición multidisciplinar bastante exigente, puesto que, en la creación de los espacios semánticos y los resultados que se obtienen de su uso, exigen conocimientos asociados a la informática, a las matemáticas, a la lingüística, a la psicología, entre otros. Por lo mismo, en el diseño investigativo, se hace necesario contar con un equipo multidisciplinar que permita comprender y aplicar adecuadamente la herramienta.

A pesar de las críticas realizadas, no deja de sorprender la capacidad de esta herramienta computacional para dar cuenta de las similitudes entre palabras y la medición de las relaciones de similitud semántica de párrafos y textos. Así también, destaca el buen resultado que han obtenido los productos derivados de estas investigaciones para la educación en Estados Unidos y, en parte, en Francia. De esta manera, el LSA como herramienta tiene proyecciones muy interesantes, sobre todo en el ámbito de la educación, puesto que es allí donde el lenguaje y sus matices semánticos toman mayor relevancia para nuestros propósitos. De esta manera, la utilización del LSA como herramienta es una opción, a nuestro parecer, muy adecuada para llevar a cabo el estudio de las relaciones de similitudes léxico-semánticas en español, en tanto se ha demostrado su utilidad en aplicaciones muy similares en idioma inglés.

De hecho, utilizar el análisis semántico latente con corpus en español es de por sí una proyección muy interesante dado que no existe mucha investigación en este idioma, a pesar de que, como sabemos, es la cuarta lengua más hablada en el mundo después del mandarín, el hindi y el inglés y la segunda más ocupada como instrumento de comunicación, después del inglés.

En suma, nuestro interés por el LSA se sustenta en la funcionalidad práctica que tiene como herramienta computacional para la identificación de relaciones léxico-semánticas independientes de la sintaxis, asumiendo que con ello no se puede explicar el funcionamiento cognitivo de los individuos, pero sí permite obtener resultados similares a los obtenidos por los seres humanos ante tareas de similitud léxico-semántica entre palabras, entre las partes de un texto o entre textos distintos.

SEGUNDA PARTE
MARCO METODOLÓGICO

CAPÍTULO 3

“Metodología de la investigación”

3.1 Objetivo del capítulo

El objetivo de este capítulo es presentar los criterios metodológicos utilizados en el desarrollo de este trabajo. En primer lugar, se contextualizará y se describirá el problema de investigación que abordaremos. Luego, se recapitularán las decisiones teóricas que hemos adoptado a lo largo de los Capítulos 1 y 2. A partir de estas decisiones, se determinará el tipo de estudio y se detallarán tanto las variables como las hipótesis a investigar. Luego, se describirá el corpus y la conformación del espacio semántico especializado. Por último, se presentará el método de análisis que se utilizará para comprobar las hipótesis planteadas.

3.2 Planteamiento del problema

En general, los países de Latinoamérica presentan serias dificultades en cuanto al manejo del código escrito por parte de sus ciudadanos. Así, según los resultados de la medición del Programa Internacional para la Evaluación de Estudiantes (PISA+, por su sigla en inglés) del año 2001 que incluyó a 11 países no miembros de la Organización para la Cooperación de Desarrollo Económico (OECD, por su sigla en inglés), entre ellos 4 Latinoamericanos, queda claro que los alumnos de 15 años de la región obtienen resultados muy por debajo al de los de países desarrollados. En esta medición, Chile superó al promedio de los países Latinoamericanos que participaron en PISA+, a pesar de presentar en el área de lectura un puntaje promedio de 410 puntos, estadísticamente equivalente a los de Argentina, Brasil y México; y superior al de Perú. Sin embargo, según el informe, el 20% de los adolescentes chilenos está por debajo del nivel mínimo de lectura y un 28% apenas alcanza este rango, lo que significa que en el mejor de los casos, sólo pueden manejar las tareas más básicas de lectura y que carecen de los cimientos de las destrezas necesarias para lograr mayores

aprendizajes. En definitiva, el 48% de los escolares de 15 años tiene problemas con la lectura y sólo un 1% logra el nivel máximo (SIMCE, 2005a).

Ahora bien, pese a que Chile está entre los 5 últimos lugares de la lista al compararse con los otros 4 países medidos a nivel Latinoamericano, se encuentra en mejores condiciones que Brasil y Perú (cuyos alumnos que están en nivel 1 o bajo este llegan a un 56% y 80%, respectivamente), pero debajo de los resultados obtenidos por Argentina y México (44%). Resultados que están muy lejos del apenas 7% de alumnos finlandeses que está bajo el nivel mínimo de lectura (SIMCE, 2005a).

Estos resultados solo confirman los magros resultados obtenidos por los estudiantes chilenos de 3° y 4° básico en el estudio de la UNESCO-OREALC de 1999, en el que se demuestra que Chile ostenta un nivel mínimo en el manejo del lenguaje escrito (UNESCO-OREALC, 1999).

En cuanto a las mediciones nacionales, los resultados SIMCE (Sistema de Medición de la Calidad en Educación) de los últimos años muestran que no hay diferencias significativas en los resultados de las pruebas en el área de lengua castellana, lo que agudiza el problema descrito (SIMCE, 2005b).

Por otra parte, pruebas no gubernamentales realizadas por especialistas demuestran que durante los últimos años el rendimiento de los estudiantes no supera el 30% en pruebas de comprensión y producción textual (Peronard, 1989, 1998; Parodi, 1998, 1999; Marinkovich & Morán, 1995; Parodi & Núñez, 1998; Marinkovich, 1999; Morán, 1999), siendo este un porcentaje constante de rendimiento que concuerda con las investigaciones internacionales.

Todos estos resultados en el nivel escolar dejan en claro que la educación de la lecto-escritura en nuestro país no está siendo lo suficientemente efectiva como para que nuestros estudiantes logren desarrollar las competencias mínimas requeridas para comprender lo que leen y producir textos que les permitan desenvolverse adecuadamente en la sociedad actual. Sociedad que, cada vez más, exige de sus integrantes altas competencias en el procesamiento y comunicación de la información en constante crecimiento.

Los resultados obtenidos en el desempeño de la comprensión en adultos, tampoco son satisfactorios. Así, según el estudio de la OECD realizado en cooperación con el Departamento de Estadísticas de Canadá y centros especializados de Estados Unidos durante el año 2000, se demostró que en Chile, único país Latinoamericano invitado a participar, entre el 50% y 57% de la población adulta está por debajo de lo que en el estudio se estima un nivel mínimo de comprensión para funcionar en el mundo de hoy, es decir, que no entiende lo que lee. Tal investigación pone en evidencia que en Chile existe un gran déficit de competencias y capacidades en el manejo del código escrito de la población mayor de 15 años. Incluso estos resultados se manifiestan, de cierto modo, entre los adultos que han egresado de una carrera universitaria, ya que el 13% de ellos carece de una buena comprensión lectora (OECD & Statistics Canada, 2000).

En cuanto a la producción escrita, los antecedentes son muy similares. En general, se ha comprobado que los escritores tienen escaso o nulo manejo de los conocimientos que se exigen para redactar un texto coherentemente, incorporando factores como para quién se escribe, sobre qué materia o tópico y en qué estilo o registro. Al mismo tiempo, se ha detectado que la mayoría de los sujetos investigados no poseen dominio sobre los recursos de textualización elementales y tienen serias dificultades para mantener el tópico a lo largo de un escrito. Al parecer, se tiende a una reproducción de conocimientos almacenados, ya sea en la memoria de largo plazo, sobre hechos no directamente ligados al de la tarea de escritura. Respecto de los diversos tipos de textos, es evidente que existe una tendencia a favor del discurso narrativo, tal vez, producto del énfasis otorgado por el sistema educacional, en desmedro de otros tipos textuales (Marinkovich, 1999; Parodi & Núñez, 1999; Benítez & Velásquez, 2000; Parodi, 2001). Así, enfrentados a la tarea de escribir textos argumentativos, por ejemplo, se detecta en muchos casos una clara tendencia a la narrativización de hechos, más o menos esquemáticos o previamente conocidos, más que a una auténtica escritura creativa en torno a las categorías de una argumentación propiamente tal (Sánchez, 1993; Núñez, 1999; Parodi, 2000). Incluso, alumnos de educación superior presentan serias dificultades para producir textos argumentativos que presenten una estructura y calidad argumental adecuada en torno a un tema de su especialidad (Venegas, 2003).

En cuanto al ámbito del desarrollo de la escritura científica, los indicadores de desarrollo científico, en nuestro país, tampoco presentan un panorama muy favorable. Por ejemplo, en los últimos años, el presupuesto anual de la Comisión Nacional de Ciencia y Tecnología (CONICYT) se ha duplicado respecto del año 1999, alcanzando un monto cercano a los sesenta millones de dólares. Aun así, la producción científica, en términos de publicaciones, alcanza apenas un 10,5% de las publicaciones científicas de los países Latinoamericanos más productivos, estando bajo Brasil, Argentina y México, respectivamente, y superando en apenas un 5% a Venezuela. Este valor, en términos internacionales, cae a un 0,18% de publicaciones científicas registradas en ISI (International Scientific Index) entre 1981 y 2002 (CONICYT, 2004). Todo lo anterior indica claramente que la producción científica nacional es muy escasa en el marco del desarrollo científico internacional y que, a pesar de la implementación de nuevas políticas de desarrollo científico, el desarrollo de la escritura científica en términos de artículos de investigación publicados en revistas de corriente principal es muy escaso. Por último, otro dato que confirma esta situación poco afortunada en cuanto al desarrollo científico es el hecho de que a partir del año 2000 se ha invertido por parte de las universidades chilenas, en promedio, casi más del doble del dinero que se invirtiera en las décadas de los ochenta y los noventa (110.193 y 62.510 millones de pesos respectivamente), lo cual ha redundado en un aumento promedio de casi dos veces la cantidad de postgraduados (5.947 y 3.169 respectivamente). Sin embargo, el índice de productividad de artículos por postgraduado en promedio no ha variado, siendo incluso algo menor (0,28 y 0,29 respectivamente), lo que, en otros términos, implica que por artículo existen entre tres a cuatro postgraduados (CONICYT, 2004).

Estos datos nos indican que, en general, se publica muy poco para la cantidad de recursos invertidos y para la cantidad de investigadores que cada año obtienen su postgrado. Esto, seguramente, debido, como hemos visto hasta ahora, a la falta de desarrollo de las competencias del lenguaje escrito en todos los niveles educacionales. Ahora bien, en cuanto a la escritura científica, sabemos que escribir un artículo de investigación puede resultar una tarea muy compleja para el investigador, sobre todo para quienes se inician en el ámbito científico. En particular, esta tarea involucra no solo el conocimiento del área temática en la cual se desarrolló la investigación, sino que, fundamentalmente, exige del autor el conocimiento de las estructuras retóricas básicas que exige este tipo de texto

(Swales, 1990, 2004; Moyano, 2003) y de las competencias discursivas necesarias para la textualización, a partir de las representaciones psicosociodiscursivas que afectan el proceso de investigación y que se plasman en cada una de las partes del artículo mencionadas.

En este sentido, la problemática acerca de la escritura científica se nos presenta como uno de los ejes fundamentales al momento de pensar en el desarrollo científico de un país. Esto porque el artículo de investigación científica es un producto de conocimiento que ostenta, al menos, dos funciones identificables. Una, en cuanto medio de comunicación de nuevos conocimientos en un ámbito científico, esto es, como resultado de un proceso de investigación realizado por científicos desde un ámbito profesional. Y otra, en cuanto instrumento de generación de nuevo conocimiento (*dynamis*), a través de su estudio y enseñanza. Ahora bien, dada esta doble funcionalidad, cabe preguntarse por la calidad de la textualización de estos artículos de modo que estos cumplan efectivamente con el fin de dar a conocer conocimiento nuevo (en términos de originalidad y relevancia) y de manera clara para sus pares (otros investigadores que generalmente también son académicos).

De este modo, insertarse en la comunidad científica de un área especializada requiere aprender a escribir artículos de investigación científica. Desde esta perspectiva, la pregunta obvia es ¿quién debe enseñar a los investigadores a escribir artículos de investigación científica? Las opciones más probables son al menos dos. Por una parte, podemos pensar que es la universidad como institución formadora de nuevos profesionales (y, entre ellos, nuevos académicos e investigadores) la cual debe asumir tal tarea. Por otra parte, podemos plantear que es la propia comunidad científica la que debe hacerse cargo de tal aprendizaje, constituyéndose las instancias de postgrado, así como los pares evaluadores en el proceso de publicación, el espacio de este proceso de aprendizaje. Sin duda, existe aquí una problemática que en nuestro país aún no se ha desarrollado lo suficiente y que exige un tratamiento exhaustivo desde la perspectiva de las políticas de educación superior y de desarrollo científico nacional. Cabe aclarar, que tal problemática ha sido ampliamente abordada en el ámbito de la escritura científica en las universidades de Estados Unidos (ver Swales, 1990; Ivanich, 1998; Baykoucheva, 2001; y algunos recursos para escritura científica en internet como www.reed.edu/mgeselbr/Chem212/Writing.html, entre muchos otros), así como de Europa, a través de la iniciativa ADIEU (Discurso Académico en la

Unión Europea), correspondiente a un proyecto Sócrates destinado a la investigación y enseñanza de los géneros textuales del ámbito académico (Clase Magistral, Monografía, Artículo Científico) (Vázquez, 2001). Otro ejemplo europeo en esta área corresponde a los esfuerzos realizados por los académicos de la Universidad Pompeu Fabra de Barcelona, España, liderados por la Dra. Helena Calsamiglia e involucrados en el proyecto de Análisis discursivo de la divulgación científica (Calsamiglia, Bonilla, Cassany, López & Martí, 2001).

En Latinoamérica, una iniciativa en torno al desarrollo de la escritura en la universidad es la realizada en la Universidad de Buenos Aires, liderada por Dra. Elvira Narvaja de Arnoux (ver Narvaja de Arnoux, Di Stefano, Pereira, 2000). En nuestro país, se reconocen los incipientes esfuerzos realizados por el equipo de la Dra. Ana María Harvey en la Pontificia Universidad Católica de Chile, quienes estudian el discurso académico, en particular la representación y actualización del género informe en la comunicación escrita entre docentes y alumnos (Harvey, 2005). Y, también, aquellos realizados por el programa de desarrollo de la escritura académica dirigido por la Dra. Marcela Oyanedel en la Pontificia Universidad Católica de Chile (Oyanedel, 2004).

Por último, a partir de los antecedentes expuestos hasta ahora, cabe cuestionarnos, ya no solo en relación con la cantidad de artículos de investigación científica que se publican, sino que más bien respecto de su calidad. Siguiendo en esta idea, la medición de la productividad en términos cuantitativos es una labor de conteo estadístico sin mayor interés, sin embargo, parece ser mucho más relevante preguntarse por el cómo evaluar la calidad en un artículo de investigación científica. Esta será sin duda una interrogante nuclear en esta investigación.

Junto a lo anterior, en términos más amplios, cabe también preguntarse si ¿es distinto escribir en un ámbito general a escribir en un ámbito académico o a escribir en uno científico? Tal cuestionamiento es transversal a esta investigación y es centro de interés de estudios focalizados en la comparación de la producción escrita que se realiza en diversos géneros discursivos.

Por último, desde una perspectiva aún más amplia, cabe cuestionarse respecto de los responsables de educar en las competencias de la escritura científica: ¿debe ser la escuela como institución de enseñanza general? ¿debe ser la universidad? o ¿debe ser la comunidad científica? Sin duda, este es un cuestionamiento que debe ser tratado multidisciplinarmente, en especial desde una concepción de desarrollo científico unificado y consensuado por todos sus actores; por ello, es una pregunta que rebasa las fronteras de la ciencia y exige la participación de los ámbitos políticos y económicos que se relacionan con las expectativas de desarrollo no sólo científico sino que nacional.

3.3 La investigación

Debido a su carácter integrador e interdisciplinar, en esta investigación hemos recurrido a ámbitos de estudio diversos, siendo el estudio del lenguaje el eje central de todos, a partir de los cuales hemos podido ir tomando opciones teóricas y metodológicas que nos serán de utilidad.

En el primer capítulo, luego de una revisión de los conceptos de texto y discurso desde las dimensiones lingüística, cognitiva y social, hemos optado por entender al texto como una representación lingüística externa material del proceso psicosociodiscursivo. En este sentido, el texto se constituye en nuestra investigación como un objeto de estudio analizable desde su materialidad y, por ello, factible de ser analizado con un método semiautomático como un conjunto de datos lingüísticos constituido por sistemas léxicos, semánticos, gramaticales y retóricos. Lo anterior nos permite concebir al texto como un producto del proceso discursivo en el cual se reconocen marcas léxico-semánticas propias de la actividad discursiva. Asimismo, en el Capítulo 1 establecimos que los textos especializados contienen rasgos lingüísticos propios de la actividad psicosociodiscursiva de la ciencia (comunicación entre científicos), como del sistema sociodiscursivo (lenguaje de especialidad de la ciencia en español) y sociocultural (ciencia positivista en la cultura occidental), que los constituyen en tipos de textos científicos, siendo el artículo de investigación científica el tipo textual prototípico. Finalizamos el Capítulo 1 estableciendo que, en español, los estudios respecto de los artículos de investigación han sido hasta ahora abordados desde perspectivas léxicas y, fundamentalmente, retórico estructurales. En este

sentido, el eje central de los estudios ha estado en el análisis de las partes constitutivas del artículo de investigación, siendo en muchas ocasiones estudios comparativos interlenguas y con muestras ejemplares de textos.

En el Capítulo 2, establecimos que el tipo de significado que nos interesa estudiar es el significado de tipo asociativo colocacional, fundado en una teoría representacional y mentalista del significado. De este modo, nuestro análisis estará centrado en la descripción léxica de las áreas de la ciencia a partir de los textos que las componen y en la medición de un valor de significado según la colocación de una palabra junto a otras y según la co-ocurrencia de una palabra en determinados párrafos del texto. Como también establecimos en el Capítulo 2, asumimos que es posible realizar tal análisis, al menos en parte, bajo el supuesto cognitivo conexionista. Este supuesto plantea que la representación del significado en los individuos se puede explicar a partir de la asociación y fuerzas de relación entre los elementos, los cuales forman patrones de significado; en nuestro caso, la vinculación de elementos léxico-semánticos y, a partir de ellos, retóricos de los textos.

En suma, en esta investigación describiremos las relaciones léxico-semánticas, desde una perspectiva semántica colocacional de índole cuantitativo, entre ciertas partes retórico-estructurales en artículos de investigación de tres áreas de la ciencia. Para llevar a cabo este objetivo se analizarán los índices de similitud léxico-semántica obtenidos a partir de las comparaciones entre los apartados retórico-estructurales a investigar, utilizando para ello el modelo de análisis semántico ya descrito.

3.3.1 Objetivo general de investigación

Conociendo la importancia que tiene el artículo de investigación científica en la difusión del conocimiento y entendiendo que este tipo de texto es una manifestación material del proceso sociodiscursivo de la ciencia, en este caso en español, nos interesa:

Comparar las relaciones de similitud léxico-semánticas entre las palabras clave, el resumen, la introducción y el contenido de los artículos de investigación científica publicados en revistas científicas indexadas de tres áreas de la ciencia (ciencias exactas, ciencias

biológicas y ciencias sociales), utilizando una herramienta computacional de análisis vectorial denominada Análisis Semántico Latente.

3.3.1.1 Objetivos específicos de investigación

a) Describir, con asistencia de herramientas computacionales y análisis cuantitativo, los patrones de similitud léxico-semántica existentes entre las palabras clave y el contenido de los artículos digitalizados de investigación científica correspondientes a un corpus proveniente de tres áreas del conocimiento científico.

b) Describir, con asistencia de herramientas computacionales y análisis cuantitativo, los patrones de similitud léxico-semántica existentes entre las palabras clave y el resumen de los artículos digitalizados de investigación científica correspondientes a un corpus proveniente de tres áreas del conocimiento científico.

c) Describir, con asistencia de herramientas computacionales y análisis cuantitativo, los patrones de similitud léxico-semántica existentes entre las palabras clave y la introducción de los artículos digitalizados de investigación científica correspondientes a un corpus proveniente de tres áreas del conocimiento científico.

d) Describir, con asistencia de herramientas computacionales y análisis cuantitativo, patrones de similitud léxico-semántica existentes entre el resumen y el contenido de los artículos digitalizados de investigación científica correspondientes a un corpus proveniente de tres áreas del conocimiento científico.

e) Describir, con asistencia de herramientas computacionales y análisis cuantitativo, patrones de similitud léxico-semántica existentes entre el resumen y la introducción de los artículos digitalizados de investigación científica correspondientes a un corpus proveniente de tres áreas del conocimiento científico.

3.3.2 Tipo de estudio

La investigación realizada es de carácter no experimental exploratorio-descriptiva, enmarcada en una metodología cuantitativa. De este modo, decimos que es exploratoria

puesto que no existen estudios respecto de la escritura científica realizados en español utilizando herramientas computacionales de cálculo de similitud léxico-semántica como el Análisis Semántico Latente, en grandes muestras de textos. Además, planteamos que es descriptiva, pues se pretende indagar la incidencia y los valores en que se manifiestan las variables a investigar y, de este modo, caracterizar los artículos de investigación científica de las tres áreas en estudio a partir de su contenido léxico. En este sentido, seguimos a Hernández, Fernández y Baptista (1998), quienes plantean que los estudios descriptivos evalúan diversos aspectos o componentes del fenómeno a investigar, especificando sus propiedades más importantes, de manera independiente, siendo optativo relacionar los diversos componentes estudiados. En nuestro caso, nosotros optamos por presentar un conjunto de hipótesis descriptivas relacionadas con las variables que serán descritas más adelante.

3.3.3 Variables

Las variables consideradas en esta investigación serán:

Variable Área Científica: Corresponde a los artículos de investigación científica publicados en revistas de científicas indexadas, agrupados *a priori* en tres dominios: ciencias exactas, ciencias biológicas y ciencias sociales.

Variable Palabras Clave: Consiste en un grupo de palabras o frases nominales (3 a 5 aproximadamente) que, de forma altamente abstracta, compactan el significado principal del artículo de investigación científica con el que están asociadas.

Variable Resumen: Consiste en un texto breve (entre 200 y 250 palabras aproximadamente) que compacta léxico-semánticamente el significado global del artículo de investigación científica del cual procede.

Variable Introducción: Consiste en un apartado del artículo de investigación científica, compuesto por varios párrafos, en los cuales se presentan los antecedentes que se conocen respecto de los tópicos principales o problemas que serán discutidos en el artículo acorde con los objetivos de la investigación. Además, en la introducción normalmente se

cuestionan o desafían las afirmaciones asentadas en el conocimiento disciplinar y se indican los contenidos, estructura y/o objetivos del artículo en relación con la información teórica existente y el tópico principal, identificando explícitamente el foco del artículo.

Variable Contenido: Consiste en un texto compuesto por varios apartados retórico-estructurales en los cuales se textualizan los procedimientos lógicos seguidos por uno o varios investigadores en una investigación científica y los argumentos que sirven para dar cuenta e interpretar los hallazgos de dicha investigación.

Debemos precisar que, si bien la variable introducción constituye un apartado retórico-estructural de la variable contenido, hemos decidido aislarla, debido a que es un apartado común en todos los artículos de investigación científica independiente del área disciplinar, por lo tanto es una variable que puede aportar información relevante al ser comparada con las variables resumen y palabras clave en los artículos de investigación científica de las tres áreas científicas en estudio.

En este tipo de estudio exploratorio-descriptivo no se declaran variables dependientes e independientes, puesto que las hipótesis que se plantean son de tipo descriptivo y no implican causalidad (Hernández, Fernández, & Baptista, 1998).

Cabe señalar que las variables palabras clave, resumen, introducción y contenido son variables de tipo textual, a diferencia de la variable área de la ciencia que corresponde a una variable disciplinar.

3.3.4 Hipótesis de estudio

3.3.4.1 Hipótesis general

H₁: Al analizar los índices de similitudes léxico-semánticas entre las variables textuales (palabras clave-resumen, palabras clave-introducción, palabras clave-contenido, resumen-introducción y resumen-contenido) de los artículos de investigación científica en un espacio semántico especializado se presentan diferencias significativas entre las tres áreas del conocimiento investigadas.

$$H_1: \bar{X}_1 \neq \bar{X}_2 \neq \bar{X}_3 (\alpha < 0,05)$$

Donde: \bar{X}_1 = media de ciencias exactas

\bar{X}_2 = media de ciencias biológicas

\bar{X}_3 = media de ciencias sociales

($\alpha < 0,05$) = error estimado de un 5%

H_{0.1}: Al analizar los índices de similitudes léxico-semánticas entre las variables textuales (palabras clave-resumen, palabras clave-introducción, palabras clave-contenido, resumen-introducción y resumen-contenido) de los artículos de investigación científica en un espacio semántico especializado no se presentan diferencias significativas entre las tres áreas del conocimiento investigadas.

$$H_{0.1}: \bar{X}_1 = \bar{X}_2 = \bar{X}_3 (\alpha < 0,05)$$

3.3.4.2 Hipótesis específicas

A continuación se presentan las hipótesis relacionadas con la comparación de los índices de similitud léxico-semántica obtenidos entre las variables textuales estudiadas según cada una de las áreas de la ciencia. Cabe señalar que cada una de las hipótesis específicas, correspondientes a las áreas científicas de investigación, agrupan a su vez diez subhipótesis estadísticas, cuya interpretación se realizará en función del porcentaje de subhipótesis aceptadas o rechazadas acorde al error estadístico aceptado ($\alpha < 0,05$).

3.3.4.2.1 Ciencias exactas

H₂: Al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, se esperan diferencias significativas entre los índices de las siguientes relaciones:

Tabla 6. Hipótesis específica de investigación en ciencias exactas

CIENCIAS EXACTAS <i>($\alpha < 0,05$)</i>	P+R	P+I	P+C	R+I	R+C
P+R		≠	≠	≠	≠
P+I			≠	≠	≠
P+C				≠	≠
R+I					≠
R+C					

Donde: P = Palabras Clave; R = Resumen; I = Introducción; C = Contenido del artículo + = relación entre variables.

H_{0,2}: Al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, no existirán diferencias significativas entre los índices de las siguientes relaciones:

Tabla 7. Hipótesis específica nula en ciencias exactas

CIENCIAS EXACTAS <i>($\alpha < 0,05$)</i>	P+R	P+I	P+C	R+I	R+C
P+R		=	=	=	=
P+I			=	=	=
P+C				=	=
R+I					=
R+C					

3.3.4.2.2 Ciencias biológicas

H₃: Al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, se esperan diferencias significativas entre los índices de las siguientes relaciones:

Tabla 8. Hipótesis específica de investigación en ciencias biológicas

CIENCIAS BIOLÓGICAS ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		≠	≠	≠	≠
P+I			≠	≠	≠
P+C				≠	≠
R+I					≠
R+C					

H_{0,3}: Al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, no existirán diferencias significativas entre los índices de las siguientes relaciones:

Tabla 9. Hipótesis específica nula en ciencias exactas

CIENCIAS BIOLÓGICAS ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		=	=	=	=
P+I			=	=	=
P+C				=	=
R+I					=
R+C					

3.3.4.2.3 Ciencias sociales

H₄: Al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, se esperan diferencias significativas entre los índices de las siguientes relaciones:

Tabla 10. Hipótesis específica de investigación en ciencias sociales

CIENCIAS SOCIALES ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		≠	≠	≠	≠
P+I			≠	≠	≠
P+C				≠	≠
R+I					≠
R+C					

H_{0,4}: Al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, no existirán diferencias significativas entre los índices de las siguientes relaciones:

Tabla 6. Hipótesis específica nula en ciencias sociales

CIENCIAS SOCIALES ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		=	=	=	=
P+I			=	=	=
P+C				=	=
R+I					=
R+C					

3.4 El corpus

3.4.1 El concepto de corpus

Explicar el concepto de corpus, tal y como se usa hoy en día en el ámbito de la lingüística o lexicografía de corpus, o en la lingüística computacional en general, no es tan sencillo como podría parecer a primera vista. En principio, se puede llamar corpus a cualquier colección que contenga más de un texto (corpus como cuerpo textual). Sin embargo, cuando este término se usa en la lingüística actual posee una serie de implicaciones que van más allá del análisis de cualquier cuerpo textual (por ejemplo, dos novelas de un autor o un artículo de un periódico) (Pérez, 2002).

Estas implicaciones se hacen patentes en las diferentes definiciones de corpus propuestas en los últimos años. Para Leech (1991), por ejemplo, un corpus computacional es un fenómeno nada excitante, ya que es una gran cantidad de textos almacenados en un computador.

Aquí, se refleja que, aunque sea de un modo bastante simplista, podemos considerar que un corpus no es más que una colección de texto en formato magnético. Leech (1991) completa su definición recalando que la habilidad que poseen los ordenadores para buscar, recuperar, ordenar y hacer cálculos sobre cantidades masivas de texto nos ha brindado la oportunidad de comprender y de explicar el contenido de esos *corpora* de formas que no eran imaginables en la era que él denomina "pre-computacional". De hecho, dado que los avances tecnológicos van tan unidos al desarrollo de la lingüística de corpus, tal y como hoy en día la conocemos, Leech (1991) argumenta que debe denominarse Lingüística de Corpus Computacional, ya que el término "lingüística de corpus" se usaba antes del advenimiento de los ordenadores digitales. Sobre esta terminología existe una discusión aún en desarrollo (ver Caravedo, 1999, Tognini-Bonelli, 2001). Esta discusión, que en sí misma es muy relevante y productiva, está fuera del interés particular de esta investigación. En nuestro caso, entenderemos esta perspectiva lingüística en cuanto una metodología que no puede realizarse hoy en día sin la asistencia de computadores, por lo que la caracterización de "computacional" nos parece innecesaria y bastaría con denominarla Lingüística de Corpus.

Existe, también, cierto consenso en el seno de la comunidad científica relativo al hecho de que un corpus no sólo ofrece información sobre sí mismo, es decir, sobre lo que contiene, sino que representa una sección más amplia de la lengua seleccionada de acuerdo a una tipología específica (Pérez, 2002). De este modo, quizá la definición más estandarizada la ofrece el grupo de trabajo dedicado a los *corpora* textuales, conocido como EAGLES (*Expert Advisory Group on Language Engineering Standards*):

“Corpus: A collection of pieces of language that are selected and ordered according to explicit linguistic criteria, in order to be used as a sample of the language” (EAGLESa, 1996).

En esta definición se recogen dos aspectos fundamentales que deben ser considerados al trabajar con *corpora*: a) un corpus debe estar compuesto por textos producidos en

situaciones reales (“*pieces of language*”) y b) la inclusión de los textos que componen el corpus debe estar guiada por una serie de criterios lingüísticos explícitos para asegurar que pueda usarse como muestra representativa de una lengua. Además, en términos más específicos, agregan que, dado que el corpus es trabajado por medios computacionales, se debe entender al corpus computacional como: “*A corpus which is encoded in a standardised and homogeneous way for open-ended retrieval tasks*” (EAGLESa, 1996).

3.4.2 Criterios metodológicos en la recolección de corpus

Los criterios de representatividad y estandarización en la recolección de corpus han sido dos de los temas más debatidos entre la comunidad científica, con opiniones diversas recogidas en varios artículos y propuestas, algunas de ellas hechas en el seno de importantes proyectos europeos (Stubbs, 1996; Caravedo, 1999; Pérez, 2002; Rojo, 2002). En EAGLES (1996b), por ejemplo, se definen criterios mínimos que deben cumplirse para que un conjunto de textos en formato electrónico pueda ser considerado un corpus (cantidad, calidad, simplicidad y documentación), y se clasifican los diferentes tipos de *corpora* que pueden existir. Las recomendaciones proporcionadas por el proyecto EAGLES (1996b) son: a) el corpus debe ser lo más grande posible de acuerdo con las tecnologías disponibles en cada época, b) debe incluir ejemplos de amplia gama de materiales en función de ser lo más representativo posible, c) debe existir una clasificación intermedia en los géneros entre el corpus en total y las muestras individuales, d) las muestras deben ser de tamaños similares y e) el corpus, como un todo, debe tener una procedencia declarada.

Las dos primeras recomendaciones hechas por (EAGLES, 1996b) recogen la polémica suscitada hace unos años a la que muchos se referían como calidad vs. cantidad, es decir, aquellos que daban más importancia al hecho de que el corpus fuera representativo y equilibrado y aquellos que, además, destacaban la importancia de que el corpus fuera lo más cuantioso posible.

Por razones obvias de espacio no podemos profundizar en esta polémica, pero parece claro que, aunque ambas argumentaciones tienen parte de razón, ninguna postura debe ser llevada a extremos. Se ha hecho mucho énfasis en la representatividad del corpus, y a sus expensas se ha de decidir qué textos o partes de textos han de incluirse o excluirse y los

criterios que deben guiar la composición y el diseño del corpus, pero la representatividad sigue siendo, hasta ahora, un concepto bastante vago. Los estudiosos no parecen ponerse de acuerdo en cuáles son los rasgos (o los tipos de textos) que representan una lengua, ni qué proporción o qué variables (número de lectores/ oyentes, amplitud geográfica de distribución, etc.) deben guiar la inclusión o exclusión de textos (Pérez, 2002; Biber, 1988).

Por otra parte, la postura contraria (expresada coloquialmente con la frase “*more data is better data*” o “*there is no text like more text*”) puede ser, siempre que se lleve a cabo basándose en unos criterios delimitados, mucho más realista y ajustada a nuestro grado de conocimiento del uso de la lengua y a la realidad tecnológica y económica en la que vivimos, ya que no siempre es posible encontrar textos en formato electrónico de todos los tipos deseables.

Lo ideal es que un corpus sea grande y representativo. Para Biber et al. (1998: 243), por ejemplo, la representatividad se define como “*the extent to which a sample includes the full range of variability in a population*”, identificando a continuación una serie de criterios externos e internos para la compilación de un corpus.

Los primeros (criterios externos) son esencialmente criterios no lingüísticos que determinan el tipo de género, modalidad, origen y finalidad de los textos que han de incluirse. Éstos pueden identificarse con relativa facilidad antes de la construcción del corpus y son los que se encargan de que el corpus represente una variedad suficiente de contextos situacionales, por lo que se consideran externos. Un segundo grupo de criterios identifica tipos diferentes de textos, de acuerdo con una serie de categorías lingüísticas (distribución de pronombres, proposiciones o tiempos verbales, por ejemplo). Estos criterios son internos a los textos que componen el corpus (Biber et al., 1998).

Por otra parte, Biber (1988) destaca la importancia metodológica de la comparación entre corpus, pues básicamente todos los textos comparten un conjunto similar de rasgos lingüísticos y algunos de estos rasgos se constituyen en indicadores de diferencias entre registros, debido a que existirán diferencias entre sus distribuciones en los registros. De esta manera se especifica que las diferencias sistemáticas en el uso de un conjunto básico de rasgos proveen una base para distinguir entre registros (Biber et al., 1998). Muchos

registros se distinguen entre sí solo por una especial frecuencia o rara ocurrencia de un conjunto de rasgos. Halliday (1993) plantea, en este sentido, que toda variedad de lengua, puede ser definida y reconocida por ciertos síndromes, o patrones de co-ocurrencia entre rasgos de un nivel lingüístico u otros rasgos de expresión en el caso de los dialectos, rasgos del contenido en el caso de una variedad funcional o registro.

Por lo anterior, es necesario comparar el corpus de investigación con otro corpus que sirva de referencia, con el fin de conocer si una determinada frecuencia de ocurrencias de un rasgo es relevante o no.

Ahora bien, con el fin de alcanzar la mayor representatividad posible y considerando que el grado de uniformidad de las estructuras textuales en los textos científicos dependen de la disciplina a la que pertenecen (Gnutzmann & Oldenburg, 1991), se seleccionó un corpus de artículos de investigación científica digitalizados pertenecientes a un limitado rango de ámbitos de la ciencia, a saber: ciencias exactas, ciencias biológicas y ciencias sociales. Por tanto, el tipo de corpus de investigación seleccionado corresponde a uno de tipo especializado al que hemos denominado ARTICO (Artículos de Investigación Científica Originales). En él se intenta conciliar ambos criterios: calidad y cantidad.

Teniendo en cuenta un criterio cualitativo se han seleccionado textos que son representativos por su presencia prototípica en el discurso de la ciencia (artículos de investigación científica). A la vez, su selección responde a los textos digitalizados publicados en español por revistas de corriente principal en cada una de las áreas de la ciencia, esto es, revistas puestas a disposición para los investigadores en los indexadores electrónicos ScIELO (Scientific Electronic Library Online) y Latindex (Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal) que cumplen con los criterios internacionales de indexación (por ejemplo, ver www.scielo.org o www.latindex.unam.mx). En cuanto al criterio de cantidad, en esta investigación se ha trabajado en función de la selección de la mayor cantidad de textos posibles que cumplan con el criterio textual concerniente a la presencia de las variables a investigar (palabras clave, resumen, introducción y el contenido del texto). Según estos criterios, se obtuvieron 675 artículos de investigación científica, correspondiendo a un total de 3.631.124 palabras, lo cual aporta a la representatividad del corpus seleccionado,

considerando el tipo textual, la lengua en los que están escritos y el soporte del cual se obtienen.

3.4.3 Conformación del corpus ARTICO

Como ya mencionábamos más arriba, en esta investigación se recolectaron 675 artículos de investigación científica digitales obtenidos exclusivamente de revistas disponibles en internet e indexadas en ScIELO y Latindex para cada una de las áreas científicas en investigación. Los criterios de recolección de los artículos de investigación científica utilizados fueron los siguientes:

1. Se seleccionaron artículos de investigación científica que están publicados en una revista científica indexada en alguna de las 3 áreas de la ciencia en estudio. Las disciplinas incluidas en cada área son las siguientes: **Ciencias exactas:** Ciencias químicas, Física e Ingeniería química. **Ciencias biológicas:** Ciencias veterinarias, Zoología y Oceanología, Ciencias morfológicas (Anatomía), Ecología y subdisciplinas de la ecología (e.g. Autoecología, Ecología de poblaciones y comunidades, Ecología del paisaje, Ecosistemáticas, etc.), Infectología y Microbiología. **Ciencias sociales:** Ciencias de la comunicación, Ciencias de la información documental, Antropología social y cultural, Arqueología, Economía, Sociología, Políticas culturales y sociales.
2. Se seleccionaron artículos que aparecen publicados en español.
3. Se seleccionaron artículos publicados entre los años 2000 y 2003. Con excepción de algunos artículos de ciencias exactas, que corresponden al periodo entre 1997 y 2003. Esto debido a que fue muy difícil recopilar artículos de esta área en español y que cumplieran con los criterios anteriores durante el periodo original.
4. Se seleccionaron solamente artículos que presentan los criterios textuales: palabras clave, resumen, introducción y contenido del artículo.

En la Tabla 11 se especifican las revistas, el periodo de publicación y la cantidad de artículos seleccionados y palabras, según el área científica en estudio.

Tabla 11. Conformación del corpus especializado ARTICO

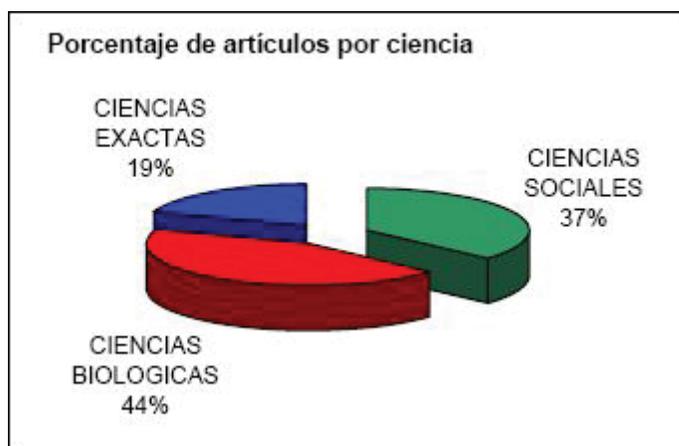
CÓDIGO	Corpus ARTICO	1997	1998	1999	2000	2001	2002
	CIENCIAS EXACTAS						
BSCQ	BOLETÍN DE LA SOCIEDAD CHILENA DE QUIMICA	0	0	15	23	16	24
RCQ	REVISTA COLOMBIANA DE QUIMICA	10	9	10	0	0	0
RPQ	REVISTA PERUANA DE QUÍMICA E INGENIERÍA QUÍMICA	0	0	0	6	0	0
AAL	ACTAS DE LA ACADEMIA LUVENTICUS	0	0	0	0	0	4
ITERC	INTERCIENCIA	0	0	0	0	1	1
ACV	ACTA CIENTÍFICA VENEZOLANA	0	0	5	2	0	0
		10	9	30	31	17	29
	CIENCIAS BIOLÓGICAS						
GC	GAYANA CONCEPCIÓN	0	0	0	6	13	12
RCHA	REVISTA CHILENA DE ANATOMIA	0	0	0	16	20	21
RCI	REVISTA CHILENA DE INFECTOLOGIA	0	0	0	12	7	5
RCHN	REVISTA DE HISTORIA NATURAL	0	0	0	22	36	32
AMV	ARCHIVOS DE MEDICINA VETERINARIA	0	0	0	19	17	16
		0	0	0	75	93	86
	CIENCIAS SOCIALES						
AMB	ÁMBITOS. REVISTA INTERNACIONAL DE COMUNICACIÓN	0	0	0	10	17	20

CHU	CHUNGARÁ.	0	0	0	14	28	10
AD	ANALES DE DOCUMENTACIÓN	0	0	0	9	13	18
NS	NUEVA SOCIEDAD	0	0	0	4	6	5
INE	INVESTIGACIONES ECONÓMICAS	0	0	0	12	10	6
		0	0	0	50	74	59
	TOTAL						

En los Anexos 1 y 2 se entrega mayor información respecto de las revistas (Anexo1: codificación y direcciones electrónicas) y de los artículos por revista de cada una de las áreas (Anexo 2: referencias bibliográficas del corpus ARTICO).

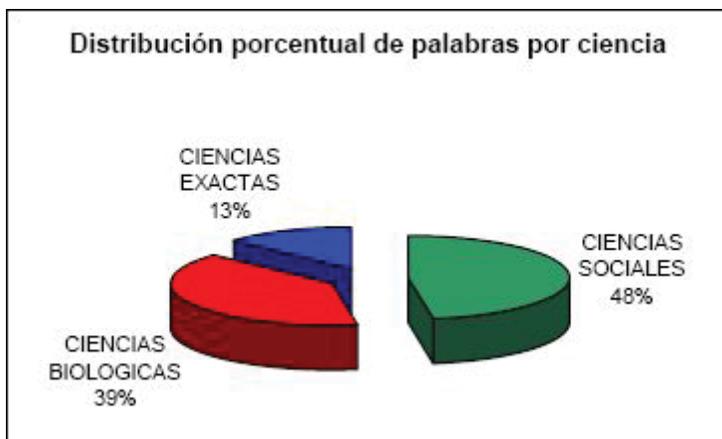
El Gráfico 1 presenta la distribución porcentual de artículos según las diferentes áreas científicas.

Gráfico 1. Distribución de los artículos por área



El Gráfico 2 presenta la distribución porcentual de palabras según las distintas áreas de la ciencia.

Gráfico 2. Distribución de las palabras por áreas



Observamos, en ambos gráficos, que existe una distribución porcentual menor con relación a los artículos y palabras para las ciencias exactas. Esto podría reflejar la preferencia de los científicos de esta área por publicar en inglés más que en español, aunque las revistas de origen pertenezcan a comunidades científicas nacionales de países latinoamericanos, según lo que pudimos comprobar al seleccionar los artículos. Independiente de lo anterior, en el análisis de las palabras y las variables se toman las medidas estadísticas que permiten normalizar los datos de modo que las muestras sean comparables estadísticamente.

Cabe hacer notar también que, en el Gráfico 1 ciencias, biológicas aparece con una mayor cantidad de artículos de investigación comparado con ciencias sociales (44% y 37%, respectivamente); sin embargo, cuando se considera el total palabras los porcentajes se invierten (39% y 48%, respectivamente). Esto nos muestra que los artículos del área de ciencias sociales presentan una tendencia a la utilización de mayor cantidad de palabras promedio por artículo, comparado con las ciencias biológicas. En la Tabla 3 se presentan los promedios de palabras por artículo en cada una de las ciencias, lo que confirma los gráficos anteriores y da cuenta de la mucho menor proporción de palabras por artículo en las ciencias exactas.

Tabla 3. Comparación del promedio de palabras por artículo.

Área de la ciencia	Artículos	Palabras	Promedio de palabras por artículos
Ciencias sociales	245	1.715.632	7.003

Ciencias biológicas	298	1.436.584	4.821
Ciencias exactas	132	478.908	3.628

3.5 Conformación general del espacio semántico

Con el fin de llevar a cabo los análisis de similitud léxico-semántica entre las variables a investigar, utilizando el LSA tal como ha sido descrito en el apartado 2.7.3.2, se debe construir un espacio semántico respecto del cual se puedan obtener los valores de similitud léxico-semántica entre las variables a investigar. Así, en esta investigación hemos decidido construir un espacio semántico o multivectorial a partir del corpus presentado en el apartado anterior. De este modo, construimos un espacio semántico especializado denominado ES-ARTICO.

3.5.1 Conformación del espacio semántico ES-ARTICO

El espacio semántico ES-ARTICO es, según hemos visto en el Capítulo 2, una representación vectorial construida utilizando el procedimiento computacional de reducción dimensional denominado *Singular Value Decomposition* (SVD) incluido en el método LSA, utilizado por el equipo de investigadores del IIS (*Institute of Intelligent Systems*) de la Universidad de Memphis, Estados Unidos.

El procedimiento desarrollado para construir el espacio semántico especializado ES-ARTICO es el que se detalla a continuación:

El primer paso consistió en eliminar toda la información no verbal contenida en los artículos (imágenes, fórmulas, símbolos gráficos, etc). El segundo paso fue separar, automáticamente, cada texto según sus párrafos, considerando para ello los puntos aparte. El tercer paso fue construir, computacionalmente, una gran matriz de frecuencias a partir de las 3.631.124 de palabras totales del corpus. Esta matriz se compuso utilizando el total de palabras únicas (*types*) del corpus, esto es 138.534 *types*, dispuestos en la primera columna, en tanto que en la primera fila se ordenó la identificación de cada uno de los 31.400

párrafos contenidos en los 675 artículos y, finalmente, en cada celda se consignó la frecuencia de aparición de cada palabra por cada párrafo. El cuarto paso fue aplicar la fórmula que permite asignar a cada celda de esta matriz rectangular de ocurrencias un peso estadístico, es decir, la frecuencia de la palabra en cada celda es convertida en su logaritmo (log), calculándose luego su entropía, esto es, $-p \log p$ para todas entradas en la fila. Luego, cada entrada en las celdas de la matriz es dividida por el valor entrópico de cada fila, este procedimiento permite normalizar las frecuencias y darle mayor peso a las palabras que co-ocurren menos frecuentemente en los documentos. El quinto paso fue aplicar el algoritmo SVD, que permite reducir la matriz original, obteniéndose con ello una matriz reducida de 60.229 palabras únicas y 294 dimensiones, convirtiéndose esta (a través de la multiplicación de los valores de las tres matrices que resultan del SVD) en el espacio semántico especializado respecto del cual realizaremos las comparaciones de similitud semántica.

3.6 Método de análisis

El método de análisis utilizado, como ya hemos planteado, se basa en la co-ocurrencia de palabras por párrafos, lo cual permite determinar los valores de similitud léxico-semántica de las relaciones de variables descritas en el apartado 3.3.4. Para llevar a cabo este análisis, se utiliza una interfaz computacional desarrollada para tal efecto, la cual se basa en la co-ocurrencia de las palabras por párrafo de los textos a comparar y luego cuantifica las distancias o similitudes semánticas acorde a los valores entregados por la matriz reducida correspondiente al espacio semántico especializado ES-ARTICO.

3.6.1 Análisis de similitudes léxico-semánticas con LSA

Este método está en concordancia con lo que hemos descrito en los apartados anteriores y especificado en detalle en el apartado 2.7.3.2 del Capítulo 2.

Para realizar la cuantificación de las similitudes léxico-semánticas entre las variables textuales: palabras clave-resumen, palabras clave-contenido, palabras clave-introducción, resumen-introducción y resumen-contenido, se determinó estadísticamente una muestra estratificada de artículos de investigación científica por cada una de las áreas. De este

modo, se seleccionaron por azar simple 5 artículos de ciencias exactas, 12 de ciencias biológicas y 10 de ciencias sociales. Esta muestra estratificada, según el dato estadístico, es representativa del universo de artículos del corpus, puesto que tiene un porcentaje de confianza del 97% y un porcentaje de error de apenas un 3%.

En la Tabla 12 se presentan los artículos seleccionados para llevar a cabo la cuantificación de las similitudes semánticas. Para mayores detalles bibliográficos ver Anexo 3 (Codificación y referencias bibliográficas de la muestra de investigación)

Muestra	Ciencias exactas	Palabras	Ciencias biológicas	Palabras	Ciencias Sociales	Palabras
1	BSCQ1 2000 45(3)	3570	AMV3-2002 34(1)	4908	AMB1 2001 6	5550
2	BSCQ2 2001 46(1)	4784	GC3 2002 66(2)	3086	AMB11 2001 6	8395
3	BSCQ5 2001 46(4)	3583	GC3 2003 67 (1)	4982	AMB14 2001 6	4121
4	RPQ2 2000 3(1)	2300	RCHA4 2001 19(3)	3150	AD14 2002 5	6156
5	ACV3 1999 50(1)	10009	RCHA5 2001 19(3)	2735	CHU2 2002 34 (1)	13419
6			RCHA5 2002 20(2)	1780	INE4 2001 25(2)	8876
7			RCHA2 2002 20(3)	4251	INE2 2002 26(2)	8034
8			RCHN6 2000 73(4)	5010	NS3 2003 184	6701
9			RCHN7 2000 73(4)	4407	NS7 2003 186	7150
10			RCHN1 2002 75(2)	5522	NS10 2003 188	5436
11			AMV2 2001 33(1)	2609		
12			AMV15 2002 34(2)	3519		

Suma		24246		45959		73838
promedio		4849,2		3829,9		7383,8

Tabla 12. Muestra de artículos para análisis semántico latente.

A partir, de esta muestra se realizaron los siguientes procedimientos:

1. Se procedió a conformar con cada artículo cuatro archivos diferentes, correspondientes a cada una de las variables en estudio (palabras clave.txt, resumen.txt, introducción.txt, contenido.txt)

2. Utilizando el espacio semántico ES-ARTICO se procedió a calcular para cada artículo la similitud léxico-semántica entre las variables, obteniendo para cada relación de variables un promedio de similitud léxico-semántica expresado por valores entre 0 y 1, donde 0= sin similitud y 1= similitud absoluta. Para llevar a cabo este procedimiento se utilizó el programa escrito en Java denominado *LSACompareFiles.java*.

a. De este modo, se calculó la similitud léxico-semántica de cada una de las palabras clave respecto del resumen del artículo. La similitud de cada una de las palabras clave respecto de cada párrafo de la introducción del artículo y la similitud de cada una de las palabras clave respecto de cada párrafo del contenido del artículo de investigación científica.

b. Así también, se calculó la similitud léxico-semántica del resumen respecto de cada párrafo de la introducción y la similitud léxico-semántica del resumen respecto de cada párrafo del contenido de cada artículo de investigación científica (para un ejemplo ver Tabla 13).

Tabla 13. Ejemplo de cálculo de similitud léxico-semántica según variables en el espacio ES-ARTICO para el artículo AMV 3-2002.

ARTÍCULO	Párrafos	P + R	R+C	P1+ C	P2+ C	P3+ C	P4+ C	P + C	R + I	P + I
AMV 3-2002	1	0,3016349	0,527947	0,19735056	0,16572621	-0,013318104	0,19113438	0,135223262	0,527947	0,19735056
	2	0,4684747	0,47576445	0,30936727	0,15460828	-0,033237576	0,15611221	0,146712546	0,47576445	0,30936727
	3	0,015647093	0,589801	0,2814851	0,17166096	0,036741566	0,023186414	0,12826851	0,589801	0,2814851
	4	0,11782002	0,59167826	0,18858676	0,21653175	-0,03000281	0,0451018	0,105054375	0,59167826	0,18858676

5		0,46615788	0,03261172	0,17169383	0,03662898	0,04013718	0,070267928	0,46615788	0,03261172
6		0,3743995	0,14473695	0,16209954	0,030697308	-0,008408682	0,082281279	0,3743995	0,14473695
7		0,6770533	0,24300376	0,33377996	0,05375653	0,10600246	0,184135678	0,6770533	0,24300376
8		0,5700686	0,10336868	0,1584651	0,028528502	0,07259285	0,090738783		0,16572621
9		0,5527166	0,2199834	0,26968065	-0,006086713	-0,002961422	0,120153979		0,15460828
10		0,75782055	0,25242776	0,26784515	-0,03101719	0,11406286	0,150829645		0,17166096
11		0,6033305	0,18971322	0,3345815	0,035661932	0,04997319	0,152482461		0,21653175
12		0,57471305	0,2412779	0,40208307	-0,017743135	0,11622367	0,185460376		0,17169383
13		0,74202	0,17704092	0,2658535	0,011844111	0,07823965	0,133244545		0,16209954
14		0,7732731	0,17791227	0,31447673	-0,031162728	0,09825178	0,139869513		0,33377996
15		0,73955435	0,20445311	0,29180256	-0,023265146	0,083340734	0,139082815		-0,013318104
16		0,7354442	0,16981824	0,28826734	-0,04737847	0,10387353	0,12864516		-0,033237576
17		0,7785111	0,21892409	0,3501919	-0,03828263	0,09738836	0,15705543		0,036741566
18		0,7548616	0,20808345	0,30844584	-0,041209433	0,0732299	0,137137439		-0,03000281
19		0,64736605	0,16030738	0,3376131	-0,013205729	0,070519	0,138808438		0,03662898
20		0,59711874	0,08276514	0,15775816	-0,039600287	0,053678647	0,063650415		0,030697308
21		0,5703473	0,196734	0,16942258	-0,020448755	0,037547823	0,095813912		0,05375653
22		0,4798455	0,11303359	0,22020026	-0,05367424	0,09790657	0,094366545		0,19113438
23		0,8232034	0,21157958	0,33331862	-0,020207563	0,095903926	0,155148641		0,15611221
24		0,5501515	0,22918494	0,36999947	-0,002158021	0,29919434	0,224055182		0,023186414
25		0,4666385	0,19386926	0,2790249	-0,018455246	0,2944894	0,187232079		0,0451018
26		0,56478024	0,24159458	0,35474193	0,001239838	0,29804584	0,223905547		0,04013718
27		0,16677566	0,0796004	0,059388753	0,041250695	0,241374	0,105403462		-0,008408682
28		0,6984569	0,22284096	0,18301542	0,009220193	0,053389218	0,117116448		0,10600246

29		0,76674426	0,20031434	0,27586073	-0,033314276	0,08791639	0,132694296		
30		0,5633902	0,23619421	0,37002924	0,041345786	0,28533524	0,233226119		
31		0,6242381	0,21816057	0,49118543	-0,017038366	0,06388496	0,189048149		
32		0,7037167	0,29832378	0,36813715	-0,002017319	0,10905859	0,19337555		
33		0,6473779	0,1743574	0,25746247	0,026274137	0,13939098	0,149371247		
34		0,7418502	0,25897545	0,38623223	0,020314679	0,058178402	0,18092519		
35		0,5967524	0,32858634	0,22599016	-0,013292456	0,06670948	0,151998381		
36		0,775571	0,2530961	0,28174287	-0,004041097	0,13143818	0,165559013		
37		0,71086276	0,2908565	0,34796968	0,046720915	0,01809799	0,175911271		
38		0,31093436	0,11502402	0,1309673	-0,009772942	0,16607402	0,1005731		
39		0,8734857	0,21497656	0,41090757	-0,005861448	0,08559046	0,176403286		
40		0,546839	0,2133386	0,14835428	0,097110346	0,060939513	0,129935685		
Promedio	0,225894	0,619608	0,202065	0,272789	-0,003732	0,1074668	0,144647	0,528972	0,1217062

En la Tabla 13 se ejemplifica el modo de sistematizar los datos obtenidos de los cálculos de similitud léxico-semántica de las variables en el artículo AMV 3-2002 (tercer artículo de la revista Archivos de Medicina Veterinaria del año 2002 del área de ciencias biológicas). De esta forma, en primer lugar, se calculó la relación entre cada palabra clave y el resumen (P+R), teniendo el artículo 4 palabras clave. En segundo lugar, se calculó la relación del resumen con cada párrafo del contenido (40 párrafos en total). En tercer lugar, se calculó la similitud léxico-semántica de cada palabra clave (P1..P4) con cada párrafo del contenido del artículo y se promediaron estas similitudes en P+C. En cuarto lugar, se calculó el resumen con cada párrafo de la introducción (R+I), teniendo la introducción 7 párrafos en total. En quinto lugar, se calculó la similitud entre cada palabra clave con cada párrafo de la introducción (P+I). Finalmente, se calculó para cada relación entre variables el promedio de similitud semántica. Esto se realizó para cada uno de los 27 artículos que conformaron la muestra.

En la Tabla 13 se puede observar que el mayor promedio de similitud léxico-semántica se presenta en la relación R+C, seguido por R+I. Luego, con promedios mucho más bajos, aparecen las relaciones P+R, P+C y P+I, respectivamente.

3 . Se procedió a realizar pruebas estadísticas de diferencias de medias para comprobar si existía o no diferencias significativas por área de la ciencia y por relación de variables según los promedios de similitud léxico-semántica obtenidos según el espacio semántico ES-ARTICO.

CAPÍTULO 4

“Análisis de resultados y conclusiones”

4.1 Objetivo del capítulo

El objetivo de este capítulo es presentar el análisis de los resultados de nuestra investigación, obtenidos a partir de la cuantificación de las similitudes léxico-semánticas en la muestra de artículos de investigación científica. Los cálculos de similitud léxico-semántica se obtuvieron utilizando un espacio semántico especializado (ES-ARTICO) de 296 dimensiones. Este se construyó utilizando el Análisis Semántico Latente, según se especificó en el Capítulo 3. Para cumplir con el objetivo de este capítulo, presentaremos, en primer lugar, las interpretaciones de los resultados descriptivos generales de los análisis de similitud léxico-semántica entre las variables textuales, según los artículos usados en la muestra para cada área de la ciencia. En segundo lugar, daremos a conocer los resultados de los análisis estadísticos que nos permitirán probar o rechazar nuestra hipótesis general y nuestras hipótesis específicas (ver apartado 3.3.4). Con relación a la hipótesis general, analizaremos los índices de las similitudes léxico-semánticas existentes entre las variables textuales investigadas en los artículos de investigación científica de las tres áreas de la ciencia, incluidos en la muestra de investigación. Acorde con las hipótesis específicas, analizaremos los resultados de la comparación estadística de los índices de similitud semántica de las variables en cada uno de los tres dominios de la ciencia. En tercer lugar, se presentan las conclusiones a la luz del análisis de resultados obtenidos. Por último, se presentan tanto las posibles debilidades y limitaciones como las fortalezas y proyecciones de la investigación desarrollada.

4.2 Análisis de resultados generales

4.2.1 Análisis de los índices de similitudes semánticas entre las variables

En los siguientes tres apartados (4.2.1, 4.2.2 y 4.2.3) describiremos e interpretaremos los resultados obtenidos al comparar cada una de las variables textuales en estudio (palabras clave-resumen, palabras clave-introducción, palabras clave-contenido, resumen-introducción y resumen-contenido). Cabe señalar, como se planteó en el apartado 3.6.1, que para el estudio se seleccionó una muestra estratificada representativa (con un error de 3%) correspondiente a 27 artículos de investigación científica, distribuidos de la siguiente manera: 5 artículos de ciencias exactas, 12 artículos de ciencias biológicas y 10 artículos de ciencias sociales.

Recordemos que el análisis de tipo semántico que se realiza en esta investigación está fundado en una semántica de tipo asociativa, particularmente colocacional, que es implementada computacionalmente. Este tipo de semántica, como hemos planteado en el Capítulo 2, se operacionaliza a través del concepto estadístico-lingüístico denominado similitud semántica. Según el cual, la medición de la similitud semántica es conceptualizada como una medida vectorial, que permite para determinar la semejanza de dos palabras que son representadas como vectores en un espacio multidimensional. Este procedimiento estadístico es llevado a cabo, en nuestro caso, por la herramienta Análisis Semántico Latente. Esta herramienta, gracias a la implementación de la técnica estadística de reducción dimensional, denominada descomposición en valores singulares, permite representar eficientemente las relaciones entre palabras y párrafos de corpus textuales originales, lo que nos faculta para comparar el significado de unidades textuales diversas a partir de la cuantificación de estas similitudes léxico-semánticas.

Los resultados que se consideran para el análisis corresponden a los promedios totales de las similitudes léxico-semánticas entre variables, es decir, el índice de similitud semántica entre variables textuales según cada una de las áreas, obtenidos usando el LSA. Cabe señalar que los puntajes de similitud léxico-semántica, teóricamente, se ubican entre los valores 0 (sin similitud semántica) y 1 (similitud semántica absoluta).

Ahora bien, debemos hacer notar que en una comparación entre dos textos reales distintos, por muy alta que sea su similitud, esta nunca corresponderá a 1. Si se diera este valor, significaría que se está repitiendo léxicamente la misma información. Debido a lo anterior, se realizó un procedimiento estadístico de agrupación de los índices de similitud, considerando las similitudes léxico-semánticas de todas las variables de la muestra con el fin de establecer un parámetro que nos permitiera evaluar cada uno de sus índices, de acuerdo a una segmentación en cuartiles. La Tabla 6 muestra los resultados obtenidos a través de este procedimiento.

Tabla 14. Valores de segmentación en cuartiles para evaluación de los índices de similitud léxico-semántica

Cuartiles	1	2	3	4
Rango Porcentual	0-25%	25-50%	50-75%	75-100%
Grado de similitud léxico-semántica	Bajo	Medio Bajo	Medio Alto	Alto
Valor de similitud léxico-semántica para cada rango	0,0134	0,1781	0,2915	0,4976

La Tabla 6, además, muestra la relación entre cada rango porcentual del cuartil con el grado y valor umbral obtenido, a partir de todos los índices de similitud léxico-semántica entre variables textuales de la muestra de artículos (135 índices de similitud léxico-semántica en total). Estos datos nos permiten evaluar los índices de similitud obtenidos según la comparación de las variables textuales en cada una de las áreas. Así, aquellos índices de similitud que se encuentren en el rango que va de 0,0134 a 0,1780 (para la presentación de los resultados se consideran cuatro posiciones decimales) corresponderán a un grado bajo de similitud léxico-semántica. Aquellos que se encuentre entre 0,1781 y 0,2914 tendrán un grado medio bajo de similitud léxico semántica. Los índices de similitud que se ubiquen

entre 0,2195 y 0,4975 presentarán un grado medio alto de similitud léxico-semántica. Por último, aquellos que sobrepasen el valor 0,4976 corresponderán a un grado alto de similitud léxico-semántica.

Más adelante, en el apartado 4.2.2 se presentará el resultado de las pruebas estadísticas que nos permitirá, a partir del análisis de la diferencia de media de las similitudes léxico-semántica de los artículos de la muestra investigada por área, comprobar si nuestra hipótesis general se cumple o se rechaza.

4.2.1.1 Análisis de los índices de similitudes semánticas entre las variables en ciencias exactas

A continuación presentamos los resultados de las similitudes léxico-semánticas realizadas entre las variables textuales de los artículos en ciencias exactas.

Como se observa en la Tabla 15, los promedios de similitud léxico-semántica para la comparación entre palabras clave y resumen fluctúan entre el valor mínimo 0,121 en ACV3 1999 (50)1, correspondiendo este índice a un grado bajo de similitud léxico-semántica, y el valor máximo 0,453 en BSCQ1 2000 45(3) interpretándose como un grado medio alto de similitud léxico-semántica. Esta comparación entre palabras clave y resumen presenta un índice promedio de similitud léxico semántica de 0,299, situándola en un rango superior al 50% y menor que al 75% de similitud, lo que según nuestra Tabla 14 corresponde a un grado medio alto de similitud léxico-semántica entre palabras clave y resumen en esta área.

Tabla 15. Índices de similitud léxico-semántica entre variables textuales en ciencias exactas

Artículos Ciencias exactas	Similitud entre variables				
	P + R	R + C	P + C	R + I	P + I
BSCQ1 2000 45(3)	0,4530	0,6177	0,2473	0,5123	0,2473
BSCQ2 2001 46(1)	0,4028	0,5962	0,3306	0,6806	0,3821
BSCQ5 2001 46(4)	0,2186	0,5805	0,1517	0,4889	0,1272
RPQ2 2000 3(1)	0,3022	0,4289	0,1501	0,4815	0,1745
ACV3 1999 50(1)	0,1208	0,3790	0,2200	0,4185	0,2127
Promedio	0,2995	0,5205	0,2199	0,5164	0,2288

Si observamos los datos correspondientes a la comparación entre resumen y contenido de los artículos de investigación científica en esta área, podemos observar que es también el artículo ACV3 1999 50(1) el que presenta la menor similitud entre estas variables, con un grado medio alto de similitud semántica, en tanto, que el artículo BSCQ1 2000 45(3) es el que presenta la mayor similitud semántica (0,6177 = grado alto de similitud léxico-semántica) de la muestra en esta comparación de variables. El promedio de esta comparación en todos los artículos del área es de 0,5205, permitiéndonos aseverar que entre el resumen y el contenido en los artículos de la muestra existe un alto grado de similitud léxico-semántica. Cabe señalar que el valor promedio de similitud entre estas dos variables es el más alto del área, por lo que es posible afirmar que el resumen, en los artículos de la muestra de ciencias exactas, macrosemantiza en un grado mayor el contenido que las palabras clave y que, en general, el resumen da cuenta en un alto grado del significado expresado en el contenido del artículo.

Este primer dato resulta muy revelador, ya que confirma la intuición que todo escritor experto tiene al considerar al resumen como un texto que sintetiza mejor que las palabras clave el significado de un artículo. De hecho, por extensión, organización retórica y cantidad de ideas incluidas en el artículo es esperable que esto sea así. Lo destacable en este caso es que la herramienta utilizada permite detectar este fenómeno y logra discriminar las relaciones léxico-semánticas que se generan entre el resumen y las palabras clave en relación con el contenido semántico del artículo. Lo anterior, sin lugar a dudas, es un primer acercamiento a las potencialidades que el LSA ofrece en el análisis comparativo, trabajando solo con datos de co-ocurrencia léxica y sin ningún otro tipo de información de entrada (sintaxis, segmentación retórico-estructural, temática, etc.).

En cuanto al análisis de las similitudes entre palabras clave y contenido en las ciencias exactas, es posible observar que el artículo RPQ2 2000 3(1) es el que presenta el índice de similitud más bajo (0,1501), presentando un leve grado de similitud léxico-semántica. Al contrario, el artículo que mayor similitud presenta en esta comparación es el artículo BSCQ2 2001 46(1) con un índice de 0,3306, siendo esto un grado medio alto de similitud léxico-semántica. El promedio de las similitudes de las variables alcanza un índice de 0,2199, siendo el valor promedio más bajo de todas comparaciones entre variables en esta

área y correspondiendo a un grado medio bajo de similitud léxico-semántica. En consecuencia, es posible afirmar que, en ciencias exactas, la menor similitud semántica entre variables textuales es la que corresponde a la similitud semántica entre palabras clave y contenido. Por lo tanto, las palabras clave macrosemantizan muy poco el significado correspondiente al contenido de los artículos de investigación científica seleccionados en ciencias exactas. Dado estos resultados, es muy posible argumentar que las palabras clave resultan ser un débil predictor del contenido semántico de un texto en comparación con el resumen.

Si se observan los datos correspondientes a los índices de similitud entre resumen e introducción, notamos que nuevamente el artículo AVC 1999 50(1) es el que presenta el menor índice (0,4185), correspondiendo a un grado medio alto de similitud léxico-semántica. Por el contrario, el artículo BSCQ2 2001 46(1) presenta un índice de similitud de 0,6806, representando este valor un grado alto de similitud léxico-semántica. En general, las variables resumen e introducción presentan un índice promedio de 0,5164, correspondiendo a un alto grado de similitud léxico-semántica. Además, este valor es bastante cercano al índice promedio de la similitud léxico-semántica existente entre resumen y contenido. Cabe señalar, por tanto, que el resumen, en esta área, macrosemantiza en alto grado el significado tanto del contenido como del apartado retórico introducción. Por lo mismo, es posible argumentar que la relación léxico-semántica entre resumen e introducción sería en alguna medida un buen predictor de la relación existente entre el resumen y el contenido, pudiéndose así plantear que a mayor relación entre resumen e introducción, mayor similitud entre el resumen y el contenido del artículo, al menos como se presenta en nuestros datos.

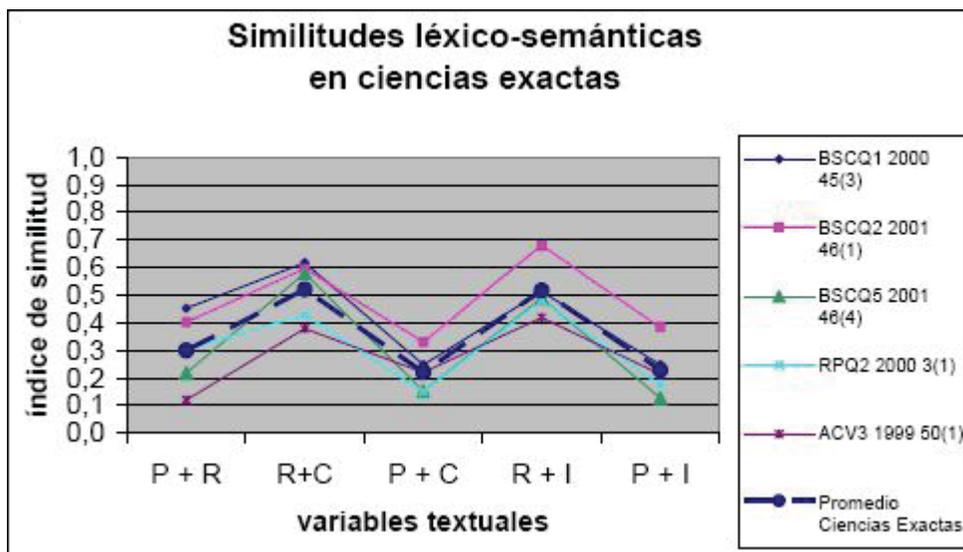
En cuanto a la similitud entre palabras clave e introducción en ciencias exactas, observamos que el artículo con menor similitud semántica corresponde a BSCQ5 2001 46(4) con apenas un 0,1272, correspondiendo a un grado bajo de similitud léxico-semántica. Por otra parte, el artículo que mayor similitud presenta entre estas variables es BSCQ2 46(1) con un índice de 0,3821, presentando por ello un grado medio alto de similitud léxico-semántica. En cuanto al valor promedio de los índices de similitud entre estas variables es posible afirmar que corresponde a un grado medio bajo (0,2288) de similitud léxico-semántica.

En suma, en el área de ciencias exactas los mayores índices de similitud corresponden a las variables resumen-contenido y a resumen-introducción, cuyos valores promedios se corresponden con un alto grado de similitud léxico-semántica. En tanto, las variables palabras clave-resumen, palabras clave-contenido y palabras clave-introducción presentan en general promedios de índices de similitud semántica correspondientes a grados que van desde medio bajo a medio alto en cuanto a la similitud léxico-semántica.

Como ya se ha esbozado, se hace aquí más evidente que la herramienta LSA permite distinguir adecuadamente grados pertinentes de relación léxico-semántica, comprobándose, a través de ella, que el resumen y el contenido están mucho más vinculados entre sí léxico-semánticamente que las palabras clave y el contenido.

A continuación, con el objeto de visualizar mejor los datos, graficamos los resultados recién expuestos y comentados.

Gráfico 3. Promedio de las similitudes semánticas entre variables en ciencias exactas.



En este gráfico se presentan los índices de similitud léxico-semántica obtenidos al comparar las variables textuales de los artículos usados como muestra para las ciencias exactas en el espacio ES-ARTICO, utilizando el LSA. Observando el Gráfico 3, podemos comprobar que, en promedio, la similitud semántica entre resumen-contenido es la que presenta un mayor índice de similitud léxico-semántica, seguida por resumen-introducción. También, es posible observar que la similitud semántica entre las variables que incluyen a las palabras

clave es, en el promedio de todos los casos, menor que las que incluyen al resumen. Además, cabe señalar que en promedio la similitud entre las variables palabras clave-resumen es más alta que palabras clave-contenido e introducción. Es interesante destacar que el artículo BSCQ2 2001 46(1) presenta un comportamiento bastante regular en cuanto a las similitudes semánticas entre las variables, presentando en la mayoría de las comparaciones los índices más altos de similitud léxico-semántica en la muestra de artículos del área de las ciencias exactas.

Acorde con lo anterior, es interesante destacar que los resultados, obtenidos utilizando una herramienta computacional como el LSA y un espacio semántico especializado, nos permite diferenciar claramente la medida en que se producen las relaciones entre las variables textuales estudiadas, particularmente entre el resumen y las palabras clave. En función de lo anterior podemos argumentar que el LSA discrimina adecuadamente, a partir de procedimientos estadísticos-matemáticos, las diferentes relaciones léxico-semánticas existentes entre las variables textuales consideradas en la investigación. En este sentido, podemos pensar que el LSA entrega como producto una discriminación muy similar a la que obtendría a través del juicio de escritores expertos ante una tarea de evaluación de las similitudes entre estas variables textuales, tal como ha sido probado para el inglés por diversos autores (Landauer, Foltz & Laham, 1998; Kintsch, Steinhart, Stahl, LSA Research Group, Matthews & Lamb, 2000; Kintsch, 2000, 2001).

4.2.3.1 Análisis de las similitudes semánticas de las variables en ciencias biológicas

En la Tabla 16 se sintetizan los resultados obtenidos para la comparación de similitudes léxico-semánticas entre las variables en estudio en ciencias biológicas.

Con relación a las primeras variables a comparar en esta área, cabe señalar que el artículo RCHA5 2001 19(3) es el artículo que presenta el menor índice de similitud léxico-semántica entre las palabras clave y el resumen (0,1468), correspondiendo este índice a un bajo grado de similitud léxico-semántica. Por el contrario, el artículo que presenta mayor índice de similitud semántica entre estas variables es GC3 2003 67(1) (0,3569), valor que según nuestra segmentación en cuartiles corresponde a un grado medio alto de similitud

léxico-semántica. En cuanto al índice promedio entre las variables, su valor alcanza a 0,2595 correspondiendo a un grado medio bajo de similitud léxico-semántica.

Tabla 16. Índices de similitud léxico-semántica entre variables textuales en ciencias biológicas

Artículos Ciencias biológicas	Similitudes entre variables				
	P + R	R+C	P + C	R + I	P + I
AMV3-2002 34(1)	0,2259	0,6196	0,1446	0,5290	0,1217
GC3 2002 66(2)	0,2928	0,6218	0,2054	0,5863	0,2464
GC3 2003 67 (1)	0,3569	0,3472	0,1412	0,5061	0,1398
RCHA4 2001 19(3)	0,2856	0,5635	0,2003	0,6009	0,1737
RCHA5 2001 19(3)	0,1468	0,4022	0,0661	0,3386	0,0134
RCHA5 2002 20(2)	0,2058	0,4324	0,1025	0,4613	0,1290
RCHA2 2002 20(3)	0,2698	0,5155	0,1845	0,4664	0,1440
RCHN6 2000 73(4)	0,2180	0,4207	0,0834	0,5386	0,1436
RCHN7 2000 73(4)	0,2670	0,6366	0,2059	0,6151	0,2512
RCHN1 2002 75(2)	0,2915	0,5062	0,1602	0,5217	0,1699
AMV2 2001 33(1)	0,2546	0,4601	0,1546	0,3931	0,1549
AMV15 2002 34(2)	0,2988	0,4350	0,1418	0,4934	0,1768
Promedio	0,2595	0,4967	0,1492	0,5042	0,1554

En cuanto a la comparación entre las variables resumen y contenido en esta área es posible establecer que el artículo GC3 2003 67(1), que anteriormente había presentado la mayor similitud léxico-semántica entre sus palabras clave y el resumen, presenta el índice más bajo de similitud léxico-semántica (0,3472), correspondiendo a un grado medio alto de similitud entre las variables. Por el contrario, el artículo de investigación que presenta el mayor índice entre estas variables es RCHN7 2000 73(4) obteniendo un valor de 0,6366, correspondiendo a un alto grado de similitud léxico-semántica. En cuanto al valor promedio de los índices de similitud de estas variables en ciencias biológicas, este alcanza a 0,4967 correspondiendo a un grado medio alto de similitud léxico-semántica. Cabe destacar que el índice de similitud promedio de esta comparación entre resumen y contenido es el que se presenta más bajo, comparativamente, entre las tres áreas de la ciencia en estudio. Esto último indica que en los artículos investigados en esta área existen resúmenes que no dan

cuenta significativamente del contenido semántico global del artículo que macrosemantizan. Un argumento posible para esta situación es que tanto los escritores como el proceso editorial al cual están afectos estos artículos no logran dar cuenta, de la misma manera como se hace en los artículos de ciencias exactas, en el resumen el significado global del contenido del artículo. Otro argumento posible es que los artículos originales (sin procesamiento computacional) usados en esta muestra presentan una gran variedad de recursos no lingüísticos como imágenes, mapas, dibujos, etc. Así, puede darse que exista mayor cantidad de descripciones asociadas a estos recursos y que no son recuperadas léxico-semánticamente en el resumen, haciendo, por ello, que la similitud léxico-semántica sea más baja que en los artículos investigados en las otras áreas.

Como podemos observar en la comparación entre palabras clave y contenido en esta área, el artículo RCHA5 2001 19(3) es el que presenta el menor índice de similitud léxico-semántica entre las variables con un valor de apenas 0,0661, correspondiendo a un grado bajo de similitud semántica. Por otra parte, el artículo RCHN7 2000 73(4) presenta el índice más alto en esta comparación con 0,2059, siendo esto un grado de similitud medio bajo. En promedio los artículos de esta área presentan entre las palabras clave y el resumen un índice de 0,1492, lo cual correspondería a un grado bajo de similitud léxico-semántica. Este valor promedio entre estas variables ubica a los artículos de ciencias biológicas como los de menor índice de similitud léxico-semántica entre los artículos investigados para todas las áreas de la ciencia. Este resultado es interesante, puesto que confirma el índice obtenido para la relación resumen y contenido. En este sentido, tanto resumen como palabras clave se relacionan menos en los artículos de esta área que en los de las otras ciencias investigadas. Esto puede deberse a que existe menor rigurosidad en la redacción del resumen, no considerando algunos significados relevantes del contenido y a que la elección de las palabras clave se utilizarían menos para macrosemantizar el contenido semántico global del artículo que dar cuenta de aspectos más contextuales de la investigación (por ejemplo, referencia al país donde se efectuó la investigación).

En cuanto a la comparación de similitud entre las variables resumen e introducción, el artículo RCHA5 2001 19(3) presenta un índice de 0,3386, siendo el más bajo del grupo de artículos y correspondiendo a un grado medio alto de similitud léxico-semántica. Por el

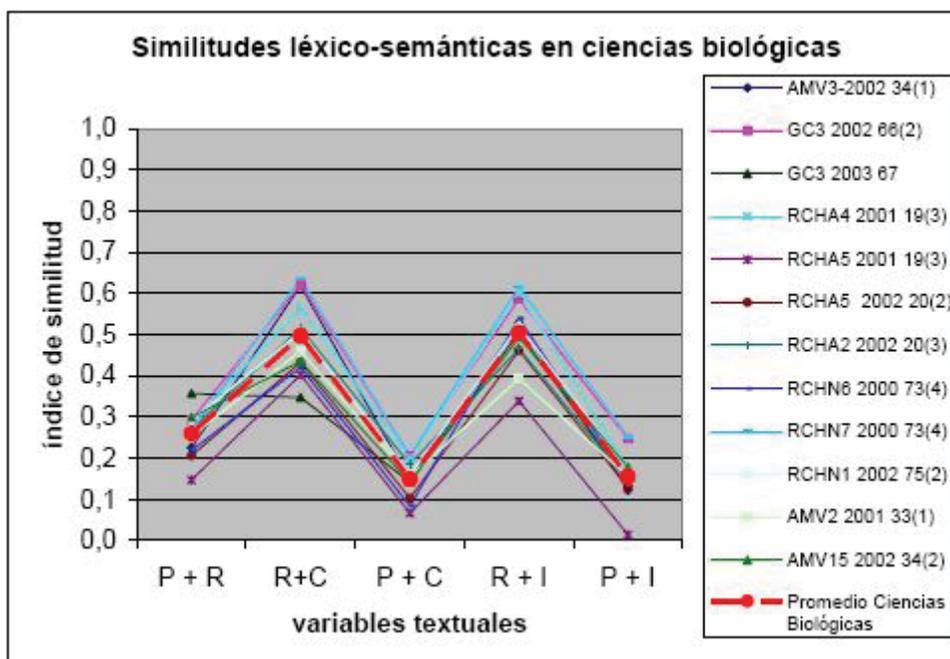
contrario, el artículo RCHN7 2000 73(4) es el que presenta el índice más alto de similitud entre estas variables (0,6151), correspondiendo a un grado alto de similitud léxico-semántica. El valor promedio de los índices de similitud entre estas variables alcanza a 0,5042, lo que nos permite establecer que el grado de similitud léxico-semántica entre resumen e introducción en esta área es alto. Cabe hacer notar que este índice promedio entre variables textuales es el más alto de esta área.

Este resultado fortalece la idea de que el contenido posee o bien muchos procesos descriptivos no reflejados en el resumen o bien mayores disgregaciones temáticas que no son posibles de incluir en el resumen. Aparentemente esto no se daría respecto de la introducción, debido a su función retórica, en la que existiría una tendencia a una mayor consistencia temática y por ello se presentaría una mayor relación léxico-semántica con el resumen, identificada por la herramienta computacional.

Por último, la comparación de similitudes entre los artículos de esta área según las variables palabras clave e introducción nos permite establecer que es nuevamente el artículo RCHA5 2001 19(3) el que presenta el índice de similitud más bajo respecto del área con un valor de 0,0134. Cabe señalar, que este índice es el valor más bajo que se presenta al analizar todas las similitudes léxico-semánticas entre las variables textuales realizadas en esta investigación. El artículo de RCHN7 2000 73(4) es el que presenta el mayor índice de similitud entre estas variables en ciencias biológicas con un valor de 0,2512, correspondiendo a un grado medio bajo de similitud léxico-semántica. En cuanto al promedio (0,2288) de las similitudes entre las variables palabras clave e introducción en esta área es posible establecer que corresponden a un grado bajo de similitud léxico-semántica.

A continuación, presentamos en el Gráfico 4 una visualización de los resultados comentados anteriormente.

Gráfico 4. Promedio de las similitudes semánticas entre variables en ciencias biológicas



Como se observa en el Gráfico 4, en esta área de las ciencias solo la relación entre las variables resumen e introducción presentan un alto grado de similitud léxico-semántica, siendo un valor que apenas sobrepasa el valor umbral del cuartil más alto, en tanto que entre resumen y contenido, la segunda más alta, se presenta solo un grado medio alto de similitud léxico-semántica. Las relaciones entre variables que consideran a las palabras claves se mantienen, en general, entre los grados medio bajo y bajo de similitud léxico-semántica. Cabe señalar que este patrón de similitudes entre las variables textuales es el mismo que se produce en ciencias exactas, aunque con valores promedio más bajos.

Estos bajos índices de similitud léxico-semántica en los artículos de la muestra en ciencias biológicas, así como también en los otros artículos de la muestra, puede deberse a que el método de cuantificación de las similitudes no capta las estrategias de macrosemantización que excede los márgenes de los textos. Es posible pensar, en este sentido, que las palabras clave sí macrosemantizan el significado global del texto, pero no en términos de las relaciones léxicas intratextuales, sino que en función de relaciones de macrosemantización más abstractas aún, de tipo exógenas al texto y de carácter eminentemente intertextual e incluso interdiscursivo.

Un ejemplo, de lo anterior se da en el artículo GC3-2003 67(1) de ciencias biológicas, en el que de entre sus cinco palabras clave dos de ellas ‘nuevo hospedero’ y ‘nuevo registro’ solo aparecen una vez cada una en la introducción, en tanto que las otras aparecen reiteradas veces a lo largo del texto. La primera de ellas aparece en el texto refiriéndose a la caracterización de una rata (*Rattus rattus*), en cuanto a que en ella habitarían 3 especies de pulgas, que son investigadas en el estudio. De este modo, ‘nuevo hospedero’ es un concepto muy importante, porque es un lugar no descrito previamente en la literatura en esta disciplina en donde habitan estos insectos, sin embargo, esto no es mencionado nuevamente en el texto. La segunda frase clave ‘nuevo registro’ indica el aporte de la investigación, mostrando la importancia de la investigación, sin embargo, no es retomada nuevamente en el texto. Por último, cabe señalar que, la mayoría de las palabras clave se encuentran concentradas en el último párrafo de la introducción. La diferencia entre estas frases clave con las otras palabras o frases clave, es que estas últimas se manifiestan reiterativamente de modo literal o semánticamente relacionadas con otras palabras en el resto del texto.

Ejemplo:

Palabras claves: Pulgas, VIII Región, nuevo hospedero, nuevo registro, roedores

“Teniendo en consideración que los **roedores** han sido el principal sustrato de investigación en Chile para éstos *insectos*, proporcionando en su mayor parte la información taxonómica disponible sobre *sifonápteros*, lo que ha dado origen a catálogos y descripciones de especies (Smit 1968; Alarcón 2000), y que aspectos ecológicos de las **pulgas** como por ejemplo su fenología y especificidad son pobremente entendidos, es que este artículo pretende contribuir al escaso conocimiento de las **pulgas**, en Chile. Se discute además a *Rattus rattus* como un **nuevo hospedero** para 3 especies de **pulgas** *Ctenoparia inopinata*, *Sphinctopsylla ares* y *Tetrapsyllus rhombus* y la presencia de *Neotyphloceras crassispina* como un **nuevo registro** para la **VIII Región**” (CG3 2003 67(1): 16)

Para una mejor visualización hemos ennegrecido las palabras clave que aparecen en el texto y hemos puesto en cursivas las palabras que tienen relación léxico-semántica con alguna de las palabras clave.

Como podemos observar, la palabra clave que más se repite es pulga, estando además muy relacionada con las palabras *insectos*, *sifonápteros* (nombre científico de la especie) y con las distintas especies mencionadas. Luego, aparece roedor relacionado con *Rattus rattus* (nombre de un tipo de roedor). Las frases clave ‘nuevo hospedero’, ‘nuevo registro’ y ‘VIII Región’ ocurren solo una vez en el texto del ejemplo, sin embargo, la última aparece en múltiples ocasiones en el texto completo. Reiteramos entonces que las palabras clave, según nuestros datos y el ejemplo, cumplen una función de señalización disciplinar, temática y/o metodológica, e incluso, en este caso la función parece más bien resaltar el aporte de la investigación que el contenido de la misma.

Como hemos destacado en los resultados descritos previamente, los artículos usados en la muestra en ciencias biológicas se distinguen de los artículos de ciencias exactas y, como veremos de ciencias sociales, en cuanto a sus bajos valores de similitud semántica entre las variables textuales en estudio. Planteamos que estas diferencias pueden ser atribuidas tanto al uso de recursos visuales que son descritos en términos específicos y que no se reflejan en el resumen o en las palabras clave, como a la posibilidad de que en esta área existan referencias temáticas o focalizaciones hacia los aportes de la investigación que no se presentan de ninguna forma ni en el resumen ni en las palabras clave, haciendo que la relación sea comparativamente más baja que en las otras áreas.

Otro argumento posible que puede explicar esta menor similitud léxico-semántica, particularmente en cuanto al resumen, es que los artículos escogidos al azar en la muestra presentan una mayor tendencia hacia una construcción descriptiva más que informativa del resumen. De este modo, se ofrecería en estos resúmenes los enunciados fundamentales del trabajo original, pero no en cuanto a resultados concretos de las reflexiones o de los estudios expuestos en el artículo.

4.2.3.1 Análisis de las similitudes semánticas de las variables en ciencias sociales

En la Tabla 17 se presentan los resultados obtenidos para la comparación de similitudes léxico-semánticas entre las variables en estudio en ciencias sociales.

Los resultados para la comparación entre las variables textuales palabras clave y resumen muestran que el artículo AD14 2002 5 es el que obtiene el menor índice de similitud léxico-semántica (0,0891), correspondiendo a un grado bajo de similitud entre la variables. Por el contrario, el artículo de investigación INE2 2002 26(2) es el que presenta el mayor índice de similitud (0,4762), correspondiendo a un grado medio alto de similitud léxico-semántica. Cabe hacer notar, que este artículo es el que presenta el mayor índice de similitud léxico-semántica entre las variables palabras clave y resumen de todos los artículos investigados en las tres áreas de la ciencia. El resultado promedio de los índices de similitud entre estas variables alcanza a un valor de 0,2397, correspondiendo a un grado medio bajo de similitud léxico-semántica.

Artículos Ciencias sociales	Similitudes entre variables				
	P + R	R+C	P + C	R + I	P + I
AMB1 2001 6	0,2479	0,1693	0,0507	0,1051	0,0741
AMB11 2001 6	0,2256	0,6572	0,2593	0,7088	0,2867
AMB14 2001 6	0,1020	0,3830	0,1266	0,5214	0,0902
AD14 2002 5	0,0891	0,5106	0,0707	0,5587	0,0591
CHU2 2002 34 (1)	0,3580	0,4831	0,2003	0,4239	0,2213
INE4 2001 25(2)	0,3224	0,5900	0,2608	0,5990	0,2446
INE2 2002 26(2)	0,4762	0,6307	0,3762	0,6534	0,4733
NS3 2003 184	0,2183	0,5019	0,1470	0,5457	0,1704
NS7 2003 186	0,2405	0,5965	0,2128	0,6694	0,2957
NS10 2003 188	0,2517	0,7268	0,2730	0,7678	0,2010
Promedio	0,2532	0,5249	0,1977	0,5553	0,2116

Tabla 17. Índices de similitud léxico-semántica entre variables textuales en ciencias sociales.

En cuanto a la comparación entre resumen y contenido en los artículos de ciencias sociales, podemos observar que el artículo que presenta el menor índice de similitud léxico-semántica es el artículo AMB1 2001 6 con 0,1693, lo cual corresponde a un grado bajo de similitud léxico-semántica. Por otra parte, el artículo NS10 2003 188 presenta el mayor índice de similitud entre las variables resumen y contenido en esta área, alcanzando un

valor de 0,7628. En promedio, los índices de similitud entre las variables alcanzan un valor de 0,5249, lo que corresponde a un alto grado de similitud léxico-semántica promedio.

Con relación a la comparación entre las variables palabras clave y contenido es posible observar que el artículo AMB1 2001 6 presenta el menor índice de similitud entre las variables (0,0507), correspondiendo a un grado medio bajo de similitud léxico-semántica. Por el contrario, el artículo INE2 2002 26(2) es el que presenta el mayor índice de similitud (0,3762) entre las variables palabras clave y contenido en ciencias sociales. En cuanto al valor promedio de los índices de similitud semántica entre estas variables es posible señalar que alcanzan un valor correspondiente a 0,1977, lo que implica un grado medio bajo de similitud léxico-semántica.

Como es posible observar en los datos correspondientes a la comparación de similitudes entre las variables resumen e introducción, es nuevamente el artículo AMB1 2001 6 el que presenta la menor similitud léxico-semántica con un índice de 0,1051, es decir, con un bajo grado de similitud léxico-semántica. Por otra parte, es nuevamente el artículo NS10 2003 188 el que presenta el mayor índice de similitud léxico-semántica entre el resumen y la introducción (0,7678). Cabe hacer notar que este valor, que corresponde a un alto grado de similitud, es el índice más alto de similitud léxico-semántica registrado en esta investigación. Por lo tanto, podemos plantear que este artículo presenta una alta consistencia semántica entre las variables retóricas que lo constituyen, particularmente entre el resumen y la introducción.

Por último, en cuanto a la comparación entre las palabras clave y la introducción se observa en la Tabla 17 que el artículo AD14 2002 5 es el que presenta el menor índice de similitud léxico-semántica (0,0591), correspondiendo este valor a un bajo grado de similitud. Por el contrario, el artículo INE2 2002 26(2) es el que presenta el mayor índice de similitud (0,4733), correspondiendo esto a un grado medio alto de similitud léxico-semántica. Cabe señalar que este artículo es el que presenta la mayor similitud léxico-semántica entre las variables palabras clave e introducción de todos los artículos investigados en la muestra en estudio.

En síntesis, los artículos de investigación científica del área de las ciencias sociales presentan, como se representa en el Gráfico 5, un patrón regular en el promedio de los índices de similitud semántica entre las variables textuales que incluyen al resumen y entre los que incluyen a las palabras clave.

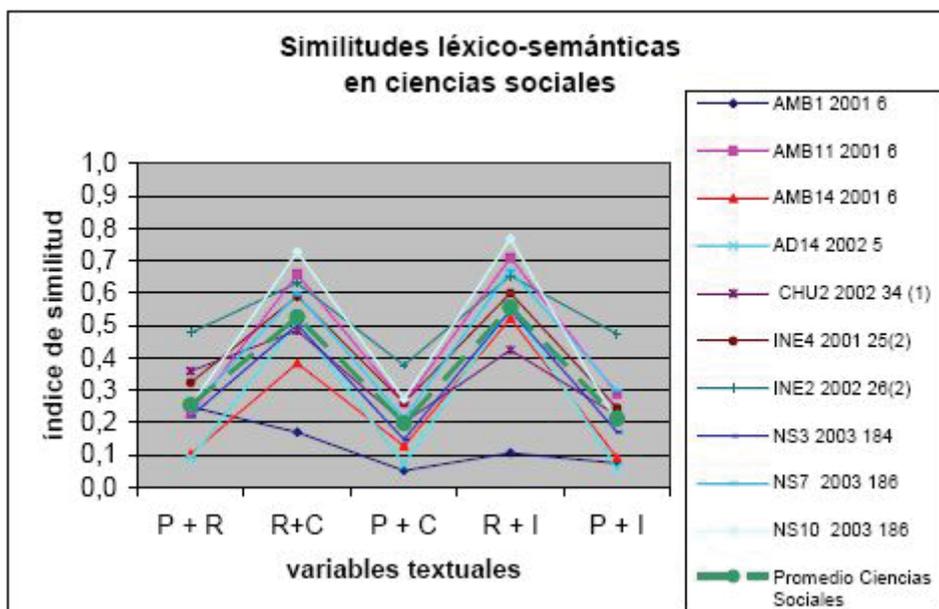


Gráfico 5. Promedio de las similitudes semánticas entre variables en ciencias sociales

El Gráfico 5 muestra que el mayor índice de similitud léxico-semántica se encuentra en la relación entre las variables resumen e introducción, siendo además el mayor índice de similitud encontrado en la investigación. Asimismo, observamos que la similitud entre resumen y contenido presenta un alto grado de similitud léxico-semántica. En tanto, la comparación de variables que incluyen a las palabras clave son, en el promedio de todos los casos, menores que las que incluyen al resumen. Es interesante destacar que el artículo INE2 2002 26(2) presenta un comportamiento bastante regular en cuanto a las similitudes semánticas entre todas las variables, presentando valores muy altos para las variables palabras clave-resumen, palabras clave-contenido y palabras clave-introducción (siendo estos índices los más altos para las comparaciones entre estas variables en todo los artículos investigados). Además, se presentan altos grados de similitud semántica entre las variables resumen-contenido y resumen-introducción (0,6307 y 0,6534 respectivamente). Por otra parte, el artículo que presenta menores similitudes semánticas entre las variables textuales

es el artículo AMB1 2001 6, el cual presenta como similitud semántica más alta la relación entre palabras clave y resumen.

En cuanto a estos resultados, podemos establecer que la alta similitud entre las variables textuales en el artículo INE2 2002 26(2) se puede deber, principalmente, a que, en el ámbito de las ciencias sociales, la economía presenta una forma aparentemente más estandarizada y sistemática en la construcción de su conocimiento, en comparación con las otras disciplinas de las ciencias sociales; estando más cercana a la matemática y a la estadística, y con un claro sistema terminológico, que le da consistencia y que hace que se perciba a esta disciplina como la más lejana al núcleo central de las ciencias sociales y más próxima a las ciencias exactas. En este sentido, es interesante el resultado, ya que el LSA también estaría dando cuenta a través de las cuantificaciones de las relaciones léxico-semánticas de esta percepción. Esta idea se fortalece al observar los resultados obtenidos por el artículo AMB1 2001 6, cuyo foco es la temática cultural y televisiva. En este sentido, las bajas relaciones de similitud semántica se podrían deber a que en estas disciplinas no hay un cuerpo terminológico sistematizado y que hay mayores opciones léxicas que hacen que las relaciones varíen grandemente debilitando las fuerzas de relación de significado entre las variables.

4.2.2 Comparación de las áreas a partir de las similitudes léxico-semánticas (H_1)

A continuación, presentaremos comparativamente los resultados de similitud léxico-semántica entre las variables y los artículos de investigación científica de las tres áreas de la ciencia en estudio.

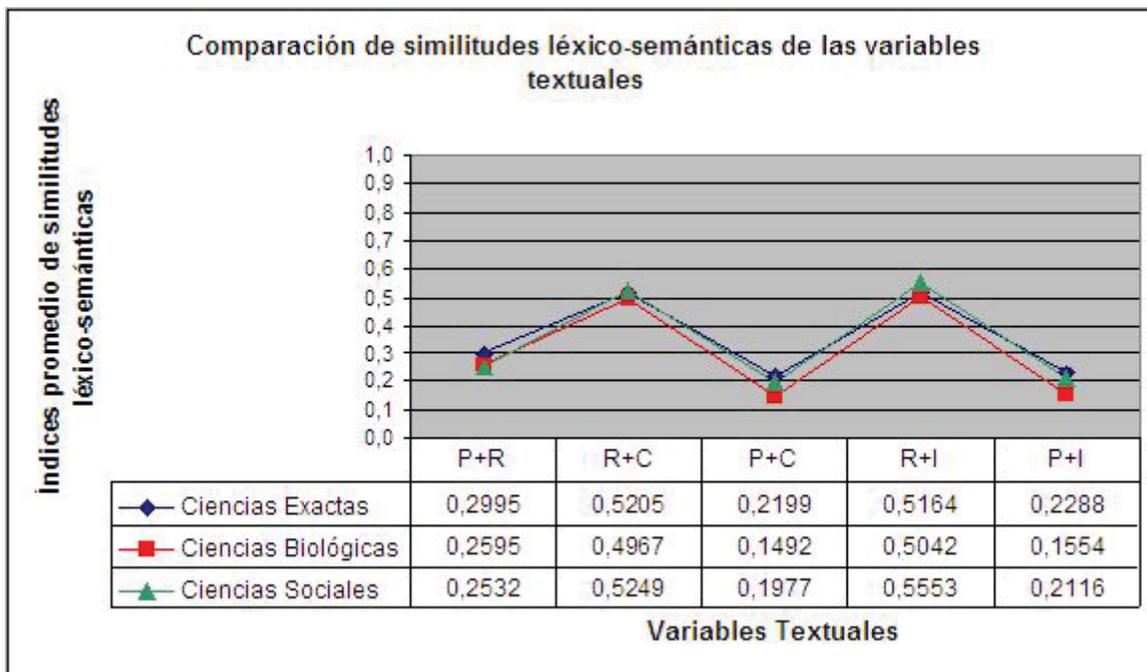


Gráfico 6. Comparación de los índices promedios de similitud léxico-semántica entre las variables de las tres áreas.

Como se puede observar en el Gráfico 6 los promedios de similitud léxico-semántica para todas las áreas son muy similares. Notamos que entre los valores en comparación entre las variables textuales se configura un patrón de relaciones de similitud semántica común a todas las áreas científicas estudiadas, así las variables palabras clave-resumen, palabras clave-contenido y palabras clave-introducción presentan en todos los artículos investigados, independientemente de la ciencia a la cual pertenecen, un valor promedio menor que las relaciones de similitud semántica entre resumen-contenido y resumen-introducción. En términos más específicos, se observa un índice promedio levemente mayor en la relación entre las palabras clave y el resumen que en las relaciones entre palabras clave-contenido y palabras clave-introducción, lo que permite suponer una tendencia a relacionar las palabras claves más con el resumen que con las otras variables textuales estudiadas. En cuanto a la comparación entre resumen-contenido y resumen-introducción, se observa que no existe mayor diferencia en los valores promedio entre las áreas. Este resultado permite suponer que, en general, el resumen macrosemantiza en igual medida tanto el contenido como la introducción de los artículos de investigación científica estudiados, independiente del área.

Al analizar con algo más de detalle cada una de las descripciones, nos percatamos que existe un patrón común en los valores de similitud semántica entre los artículos utilizados en esta investigación. De este modo, podemos afirmar que las áreas no se distinguen en las relaciones léxico-semánticas de las variables estudiadas, no, al menos, a través de la metodología computacional utilizada. Un argumento posible en torno a este resultado es que los artículos seleccionados, independiente del área de la ciencia a la cual representan, responden a procesos similares de producción, esto es, son artículos que dan cuenta de investigaciones o reflexiones científicas, producidos por escritores científicos y publicados en revistas indexadas. Esto supone un exigente proceso de estandarización en cuanto a su contenido y su forma, y por ello en cuanto a sus relaciones semánticas, a través de un proceso de evaluación editorial que tiene por fin publicar artículos coherentes e informativos.

Ahora bien, como se observa en el Gráfico 7, al utilizar la segmentación en cuartiles, como estrategia para identificar los grados de similitud semántica asociados a los índices promedio de similitud léxico-semántica entre las variables en las tres áreas, identificamos algunas diferencias, aunque no significativas estadísticamente, que no se presentan al observar solo los índices promedios.

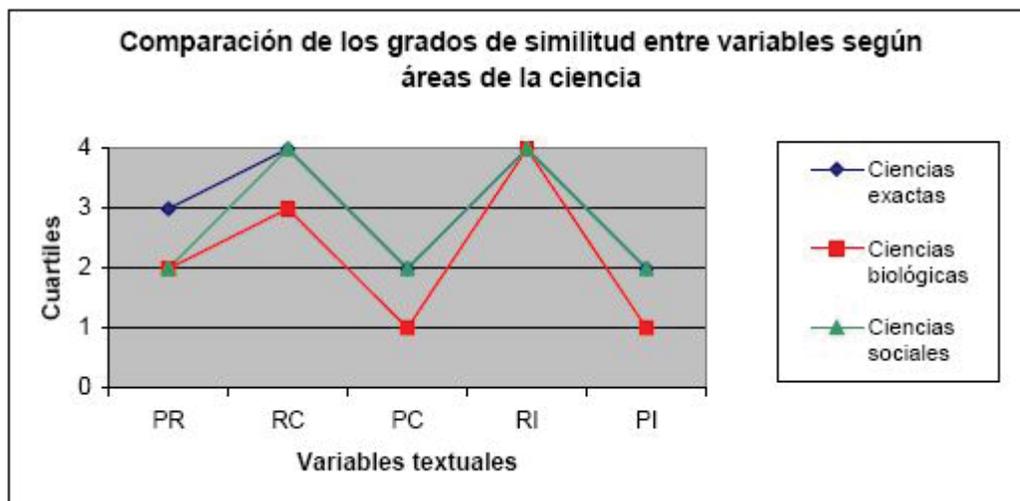


Gráfico 7. Comparación de los grados de similitud entre las variables a partir de la segmentación en cuartiles.

En el Gráfico 7 observamos que, al considerar la comparación entre palabras clave y resumen, existe diferencia en el grado de similitud semántica entre las áreas, favoreciendo a ciencias exactas, área que presenta un grado medio alto de similitud léxico-semántica, en tanto que las ciencias biológicas y las ciencias sociales presentan un grado medio bajo de similitud léxico-semántica. Este resultado nos permite establecer que en los artículos de investigación estudiados en el área de ciencias exactas existe, por lo general, una mayor cercanía semántica entre las palabras clave y el resumen que en las otras áreas, probablemente debido a la mayor especialización y relaciones terminológicas, las que tenderían a concentrarse en las estructuras retóricas que requieren mayor compactación de significado.

También es posible observar que ciencias exactas y ciencias sociales obtienen los mismos grados de similitud semántica al comparar las variables resumen-contenido, resumen-introducción, palabras clave-contenido y palabras clave-introducción, estando las primeras comparaciones en el grado más alto de similitud léxico-semántica, entre el 75 y el 100% de similitud, según los índices obtenidos en la muestra total. En cambio, respecto de las mismas comparaciones, ciencias biológicas presenta el mismo grado de similitud entre el resumen y la introducción que las otras dos ciencias, pero con un grado de similitud media alta para la comparación entre resumen-contenido.

El resultado anterior, si bien sorprende en alguna medida, pues se esperaban mayores diferencias entre las áreas, permite afirmar que los artículos de ciencias exactas y sociales presentan mayor grado de similitud léxico-semántica entre el resumen y el contenido del artículo de investigación que ciencias biológicas. Esta menor similitud entre las variables en ciencias biológicas podría deberse a que en el contenido de los artículos en esta área existen muchos procesos descriptivos particulares de los objetos en estudios, apoyados fuertemente con recursos gráficos (dibujos, imágenes, gráficos, mapas, tablas) y no necesariamente lexicalizados, lo que no permite que estos procesos sean recuperados en la reconstrucción léxico-semántica del resumen.

Cabe hacer notar, que las variables palabras clave-contenido y palabras clave-introducción en esta área también presentan un grado de similitud léxico-semántica menor que en ciencias exactas y en ciencias sociales, presentando ambas variables un bajo grado de

similitud léxico-semántica. Este resultado se puede explicar también en función de la mayor cantidad de recursos gráficos desplegados en los artículos de ciencias biológicas.

Cabe señalar, como ya hemos visto más arriba, que las similitudes entre las palabras clave y las variables resumen, contenido e introducción son, para las tres áreas de la ciencia aquí estudiadas, muy bajas en comparación con la similitud entre el resumen y las mismas variables. Esta diferencia es posible de ser explicada en cuanto a que existe mayor dificultad en macrosemantizar un contenido textual a través de cuatro o cinco palabras clave que a través de un texto mayor como el resumen. En este sentido, recordemos que las palabras clave tienen una doble función acorde con la cual, de una forma altamente abstracta, permiten resumir los significados más relevantes del contenido de un documento y, por otra, permiten indexar artículos según una temática determinada, facilitando una búsqueda rápida de un artículo relevante para el lector, cuando hay una necesidad específica (Turney, 1999). Acorde con nuestros resultados parece ser que la primera función no se cumple empíricamente, al menos en los artículos estudiados, dando pie a pensar que de estas funciones, la principal es la de permitir la recuperación del artículo por parte de un lector, lo que no necesariamente requiere una alta similitud semántica con el contenido, sino que con variables que incluso pueden ser extratextuales, cumpliendo más con una función de señalización o identificación de la disciplina, de la temática y/o de los procesos metodológicos que se relacionan con la investigación, que con la función de macrosemantización del contenido semántico global de los artículos.

Por otra parte, como señalamos en el Capítulo 1, el resumen de los artículos presenta la ventaja de poder construir un esquema cerrado semánticamente capaz, según la habilidad del productor y de sus múltiples correcciones editoriales, de referir abstractamente, usando entre 200 y 250 palabras, todos los enunciados importantes del documento primario, incluyendo aspectos teóricos que sustentan la investigación, los datos relativos a los métodos de investigación, los porcentajes, el empleo de equipos y la esfera de aplicación, pudiendo por ello captar una mayor cantidad de significados del texto (Ratteray, 1985; López, 1997; Moyano, 2000). Ahora bien, lo más relevante de este dato es que, utilizando el método que hemos seguido y la herramienta computacional LSA, podemos confirmar estas ideas. Así, no solamente la cantidad de palabras es la que produce la diferencia sino

que la relación entre las posiciones que ellas ocupan en un amplio cuerpo textual, lo que permite establecer las semejanzas en su uso y por ello las relaciones de significado entre los trozos textuales. Así, el LSA sin más información que los textos mismos nos permite confirmar la función establecida para el resumen y la medida en la que esta función se realiza en nuestros artículos de investigación científica. Dado lo anterior, podemos afirmar empíricamente que el resumen en todas las áreas científicas estudiadas permite macrosemantizar mejor el contenido semántico global que las palabras clave en el artículo de investigación científica

El Gráfico 8 presenta los valores promedio de los índices de similitud semántica obtenidos por cada uno de los artículos de investigación científica en cada una de las áreas de la ciencia investigadas. Cabe recordar, que este promedio se obtuvo a través de los índices de similitud léxico-semántica entre cada una de las variables.

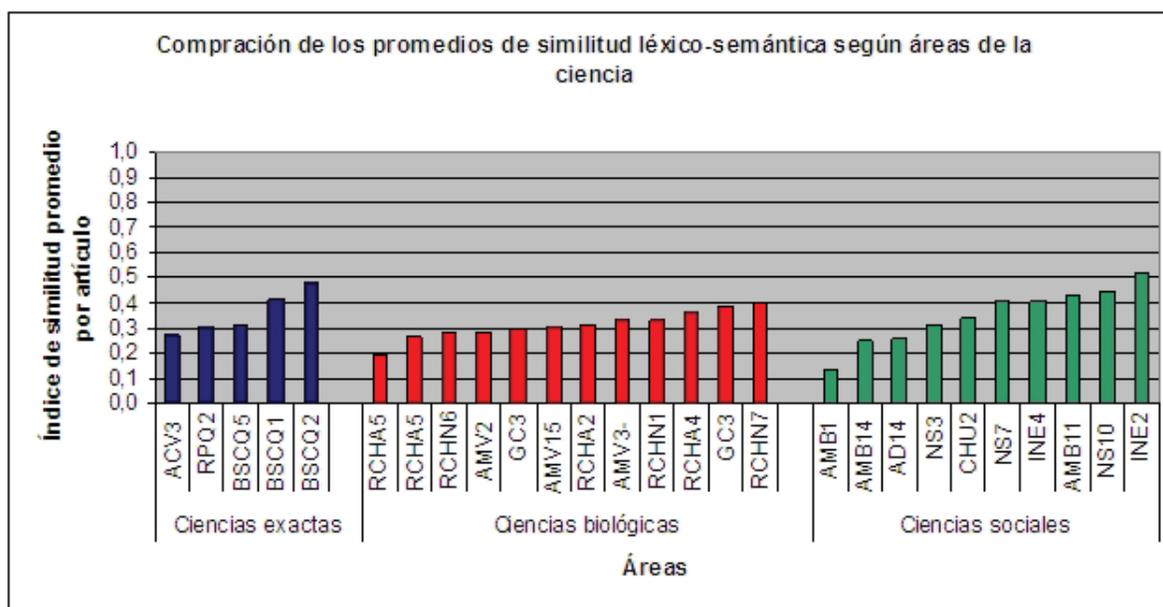


Gráfico 8. Comparación de la similitud léxico-semántica promedio de los artículos de investigación científica según las tres áreas de la ciencia.

Como es posible observar, en el área de ciencias exactas, 3 de los 5 artículos de la muestra se ubican en un rango promedio de similitud léxico-semántica menor a 0,4, en tanto que los otros dos artículos presentan un promedio de similitud léxico-semántica mayor a 0,4 pero menor a 0,5. En otras palabras, el 60% de los artículos se concentra en un rango medio bajo

de similitud semántica, en tanto que el 40% de los artículos se encuentra en un rango medio alto de similitud léxico-semántica, según los promedio totales de todos los artículos investigados. Por otra parte, en el área de ciencias biológicas, podemos observar que de los 12 artículos que constituyen la muestra, 5 se ubican bajo el promedio de similitud léxico-semántica 0,3, en tanto que los otros 7 artículos se ubican entre los promedios 0,3 y 0,4. En términos porcentuales, podemos establecer que el 41,6% de los artículos se encuentran en un rango bajo de similitud léxico-semántica global, en tanto que el 58,3% restante se concentraría en un rango medio bajo acorde a los otros promedios de similitud de los artículos. Por último, en ciencias sociales observamos que 3 de los 10 artículos presentan un promedio de similitud léxico-semántica menor a 0,3; por otra parte, 4 de los 10 artículos tienen un promedio entre 0,3 y 0,4 de similitud. Se observa también que 3 de los 10 artículos se ubican en un rango mayor a 0,4, pero menor a 0,5 y que sólo un artículo obtiene un promedio de similitud semántica superior a 0,5. Según estos resultados, podemos plantear que el 30% de los artículos se ubica en un rango bajo de similitud semántica global, que el 40% de los artículos se ubica en un rango medio bajo de similitud léxico-semántica, que el 30% de los artículos del área presenta una similitud léxico semántica global medio alta y que solo el 10% presentaría un grado de similitud léxico-semántica alta, considerando los promedios de todos los artículos de investigación científica estudiados.

De acuerdo con los resultados presentados en el Gráfico 8, podemos sostener que existe un fuerte tendencia de las ciencias biológicas a obtener resultados promedio muy similares, existiendo poca variabilidad intragrupal (desv. estándar de 0,0560). En el caso de ciencias exactas esta variabilidad es algo mayor (desv. estándar de 0,0867), en tanto que en ciencias sociales se presenta una alta variabilidad intragrupal (desv. estándar de 0,1152). Esto nos permite inferir que en el área de ciencias biológicas, los artículos seleccionados presentan mayor relación entre ellos desde el punto de vistas de sus relaciones lexico-semánticas, a pesar de que existen mayoritariamente promedios medios bajos de similitud léxico-semántica. En tanto, en ciencias sociales las relaciones léxico-semánticas son más diversas a través de los distintos artículos, pudiendo encontrarse artículos con baja similitud semántica, así como también algún artículo con alta similitud semántica en la muestra. Esto último debido, muy probablemente, a la presencia de disciplinas muy poco relacionadas al interior de las ciencias sociales (por ejemplo, economía versus ciencias de la comunicación)

o también debido a la menor estandarización terminológica o retórico estructural de los artículos de investigación científica en esta área. Como excepción a lo anterior, aparecen los artículos de la revistas Investigaciones Económicas, cuyo artículos presentan, en un caso, el mayor promedio de similitud semántica global (0,5219) y en el segundo, un puntaje correspondiente a (0,4033) considerado como de similitud léxico-semántica medio alta. Sin embargo, a pesar de la variabilidad que se observa en los datos, los promedios generales por área, presentados en el Gráfico 9, resultan ser muy similares no diferenciándose significativamente desde el punto de vista estadístico. De esta manera, se confirma que, en general, no habría mayor diferencia entre las áreas científicas acorde a las similitudes semánticas promedio entre las variables textuales de los artículos de investigación científica en estudio.

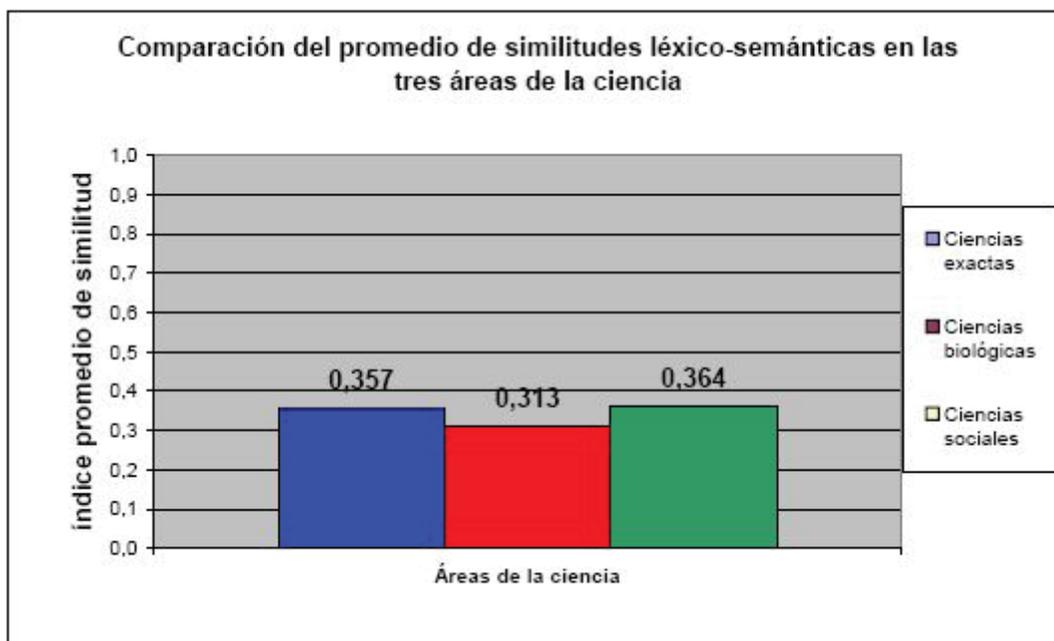


Gráfico 9. Comparación del promedio de similitud léxico-semántica de las tres áreas según ES-ARTICO.

Ahora bien, retomando la hipótesis general en la cual se planteó que:

H₁: Al analizar los índices de similitudes léxico-semánticas entre las variables textuales (palabras clave-resumen, palabras clave-introducción, palabras clave-contenido, resumen-introducción y resumen-contenido) de los artículos de investigación científica en un

espacio semántico especializado se presentan diferencias significativas entre las tres áreas del conocimiento investigadas.

Se procedió a realizar una prueba estadística de comparación de medias, que permita confirmar o refutar esta hipótesis general. De este modo, se aplicó una prueba no paramétrica, debido a que los datos de la muestra no cumplen con el supuesto de normalidad, denominada test de Kruskal-Wallis para tres muestras independientes, según la cual tenemos que:

$$H_{0,1}: \bar{X}_1 = \bar{X}_2 = \bar{X}_3 (\alpha < 0,05)$$

Es decir, se refuta la hipótesis si la media de los datos obtenidos no permite distinguir ninguna diferencia estadística entre las áreas de la ciencia investigadas (\bar{X}_1 = media de ciencias exactas, \bar{X}_2 = media de ciencias biológicas y \bar{X}_3 = media de ciencias sociales) con un error estimado de un 5%. En cambio, si al menos un par es diferente se acepta la hipótesis general de investigación H_1 ($\bar{X}_1 \neq \bar{X}_2 \neq \bar{X}_3$ ($\alpha < 0,05$)).

Finalmente, el test entrega un valor alfa igual a 0,4510, según el cual, y con un error de un 1%, nuestra hipótesis de investigación se refuta y se acepta la hipótesis nula. Esto quiere decir que, al analizar los índices de similitudes léxico-semánticas entre las variables textuales (palabras clave-resumen, palabras clave-introducción, palabras clave-contenido, resumen-introducción y resumen-contenido) de los artículos de investigación científica, en un espacio semántico especializado, no se presentan diferencias significativas entre las tres áreas del conocimiento investigadas.

Una interpretación posible de este resultado es que, en términos generales, los artículos de investigación científica usados en la muestra en estudio son el producto material de un complejo proceso de producción científica, en el que múltiples agentes conjugan sus competencias discursivas y disciplinares para lograr un artículo de investigación científica que tenga la calidad suficiente para ser publicado en una revista científica indexada, acorde con estándares internacionales de calidad científica. Así, pareciera que los artículos de investigación tienden hoy en día a una mayor homogeneización retórico-estructural y, por

ello, también a una mayor similitud léxico-semántica entre sus partes. Todo lo anterior en función de una mayor calidad comunicativa de las investigaciones realizadas en las distintas áreas de la ciencia.

Así, por ejemplo, un artículo producido por uno o varios autores, está sometido primero a un proceso recursivo de escritura y reescritura, en el cual no sólo se toma en cuenta la adecuada descripción (y argumentación) de la investigación, sino que también se consideran las normas de estandarización de la revista indexada a la cual los autores enviarán su manuscrito. Luego, dicho manuscrito es recepcionado por la correspondiente revista e ingresado al proceso de evaluación de pares científicos, quienes acorde a las normas e instrucciones propias de la revista evalúan el contenido (acorde a su experticia en la disciplina y su competencia discursiva) y la forma (según las normas de la revista y las normas de la lengua). Luego, con los comentarios emitidos por los evaluadores, si el artículo es aceptado y fuera el caso de ajustes o enmiendas, se devuelve a sus autores con el fin de reescribir nuevamente el artículo y reenviarlo para una nueva evaluación. De este modo, si los autores aceptan las sugerencias realizadas por los evaluadores y enmiendan el artículo, lo reenvían a la revista en donde el artículo nuevamente es sometido a evaluación por parte de los evaluadores y, finalmente, por parte del equipo editorial o editor, quienes en definitiva toman la decisión de enviar el artículo a imprenta en el caso que cumpla con todas las normas establecidas para su publicación o a una nueva evaluación de pares científicos según corresponda.

Ahora bien, el proceso de publicación científica de un artículo se ha presentado de modo bastante simple. Sin embargo, en las revistas indexadas, este proceso puede durar muchos meses, incluso años, y considerar muchas evaluaciones y reevaluaciones, cumpliéndose con un proceso de estandarización y mejoramiento continuo de la calidad del artículo de investigación científica.

Por lo tanto, la no diferenciación entre las áreas puede ser atribuible a que todos los artículos de investigación estudiados han sido sometidos a procesos de evaluación y ajuste de sus contenidos y formatos, de modo que se produce una cierta homogeneización en la cual se pierden las diferencias específicas propias, al menos en cuanto a las diferencias en

los procesos de macrosemantización entre las variables investigados, en estas tres áreas científicas estudiadas.

Cabría señalar, finalmente, que la única forma de comprobar si el argumento anterior es válido, sería llevando a cabo una comparación de similitudes léxico-semánticas entre artículos que hayan sido incluidos en procesos de evaluación de pares y otros que no, investigación que queda proyectada a partir de los resultados mostrados anteriormente y que podría ser desarrollada en un futuro cercano.

4.3 Resultados específicos

4.3.1 Análisis de las diferencias entre los índices de similitud léxico-semántica

A continuación se presentan los resultados obtenidos en función de las hipótesis específicas relacionadas con la comparación de los índices de similitud léxico-semántica, obtenidos entre las variables textuales estudiadas para cada una de las áreas de la ciencia.

A través de estas hipótesis se pretende confirmar o refutar estadísticamente los datos descriptivos presentados en la primera parte de este capítulo, según los cuales, las similitudes léxico-semánticas existentes entre el resumen-contenido y resumen-introducción son mayores que las similitudes léxico-semánticas entre las palabras clave-resumen, palabras clave-contenido y palabras clave-introducción, conformándose así un patrón característico de relaciones de similitud léxico-semántica en los artículos de investigación científica de todas las áreas en estudio.

Cabe señalar que se han agrupado en tres hipótesis específicas, las subhipótesis estadísticas que corresponden a todas las comparaciones posibles entre los índices de similitud léxico-semántica entre variables, según cada área investigada.

Para determinar si cada uno de estos pares de relaciones se diferencian estadísticamente se aplicó, para todas las subhipótesis de todas las áreas, la prueba estadística, no paramétrica, denominada U de Mann-Whitney. De este modo, en cada recuadro se considera el resultado

cuantitativo de la prueba, así como la comprobación (\neq) o rechazo ($=$) de las subhipótesis estadísticas particulares.

4.3.1.1 Análisis de las diferencias entre los índices de similitud léxico-semánticas en ciencias exactas (H_2)

Según la primera hipótesis específica (H_2), planteamos que al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, se esperan diferencias significativas entre los índices de las siguientes relaciones:

Tabla 18. Hipótesis específica de investigación en ciencias exactas

CIENCIAS EXACTAS ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		\neq	\neq	\neq	\neq
P+I			\neq	\neq	\neq
P+C				\neq	\neq
R+I					\neq
R+C					

Luego de aplicada la prueba estadística, los resultados obtenidos son los siguientes:

Tabla 19. Resultados de la hipótesis específica en ciencias exactas

CIENCIAS EXACTAS ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		= 0,420	= 0,547	\neq 0,015	\neq 0,047
P+I			= 1,000	\neq 0,007	\neq 0,0158
P+C				\neq 0,007	\neq 0,007
R+I					= 1,000
R+C					

Estos resultados permiten afirmar, con un porcentaje de error menor al 5%, que de las diez comparaciones entre los índices de similitud léxico-semántica entre las variables textuales en ciencias exactas, cuatro no se diferencian estadísticamente. Así, no existe diferencia

estadística entre los índices de similitud léxico-semántica entre palabras clave-resumen y palabras clave-introducción ni palabras clave-contenido, como tampoco existe diferencia estadística entre palabras clave-introducción y palabras clave-contenido. Los resultados muestran, también, que no existe diferencia estadística entre los índices de similitud léxico-semántica entre resumen-introducción y resumen-contenido.

Como vemos, la hipótesis específica se rechaza en un 40%, sin embargo, observamos que estos resultados confirman que la similitud léxico-semántica que se produce entre las palabras clave y las otras variables textuales es la misma, confirmándose el patrón presentado en el análisis descriptivo de los datos (ver Gráfico 9). Este argumento se fortalece si se observan que en todos los casos en que se compara el índice de similitud léxico-semántica entre el resumen respecto de las otras variables y el de las palabras clave respecto de las otras variables, sí existe diferencia estadística, comprobándose las hipótesis de investigación para estas comparaciones.

Los resultados en esta área permiten establecer que la hipótesis específica se comprueba en un 60% de los casos, siendo esto coherente con la determinación de un patrón de similitud léxico-semántica en el que el resumen macrosemantiza mejor que las palabras clave cualquiera de las variables textuales investigadas, particularmente, el contenido y la introducción del artículo de investigación científica en ciencias exactas.

4.3.1.2 Análisis de las diferencias entre los índices de similitud léxico-semánticas en ciencias biológicas (H₃)

La segunda hipótesis específica (H₃) plantea que, al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, se esperan diferencias significativas entre los índices de las siguientes relaciones:

Tabla 20. Hipótesis específica de investigación en ciencias biológicas

CIENCIAS BIOLÓGICAS ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		≠	≠	≠	≠
P+I			≠	≠	≠
P+C				≠	≠
R+I					≠
R+C					

La Tabla 21 presenta los resultados obtenidos según la prueba estadística para la comparación entre cada uno de los pares de relaciones. Estos resultados indican que en el 80% de los casos se presentan diferencias significativas entre los índices de similitud léxico-semántica entre las variables textuales de los artículos de ciencias biológicas.

Tabla 21. Resultados de la hipótesis específica en ciencias biológicas

CIENCIAS BIOLÓGICAS ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		≠ 0,000	≠ 0,000	≠ 0,000	≠ 0,000
P+I			= 0,932	≠ 0,000	≠ 0,000
P+C				≠ 0,000	≠ 0,000
R+I					= 0,842
R+C					

En esta área, a diferencia de ciencias exactas, sí existe diferencia estadística entre los índices de palabras clave-resumen y palabras clave-contenido, así como con palabras clave-introducción. Esto implica que en esta área existe una mayor similitud léxico-semántica entre las palabras clave y el resumen del artículo, a diferencia de lo que sucede entre las mismas palabras clave y los otras variables textuales. Esta diferencia, si bien nos sorprende, puede deberse a que en esta área el contenido del artículo contiene en muchos de los casos representaciones gráficas que permiten dar cuenta de la descripción de los objetos investigados, lo que a su vez no sería representado lingüísticamente, perdiéndose, por ello, valor en relación de similitud léxico-semántica. Por otra parte, en cuanto a la diferencia entre las palabras clave-resumen y las palabras clave-introducción sería posible suponer que la menor relación de las variables palabra clave-introducción se debería a que en la

introducción existiría información temática que no ha sido representada mayormente en las palabras clave. Sin embargo, no se debe olvidar que los índices de similitud entre estas variables son muy bajos en comparación con los que incluyen al resumen.

Es probable, también, que la mayor similitud entre las palabras clave y el resumen que entre las palabras clave y las otras variables textuales, dependa de una especie de condensación terminológica, esto es que se privilegie la mayor densidad terminológica en el resumen y en las palabras clave, con el fin de presentar la investigación a los especialistas acorde a un lenguaje altamente especializado. Tal densidad, en la introducción y el contenido iría perdiendo consistencia acorde un mayor desarrollo descriptivo y explicativo de los fenómenos investigados.

Si se observa el resto de los datos, notamos que existe congruencia con los resultados expuestos en ciencias exactas, produciéndose nuevamente el patrón de relaciones entre las variables descritas a partir del Gráfico 9. Así, observamos que existe diferencia estadística en las relaciones que involucran al resumen con el contenido y la introducción versus las relaciones que involucran a las palabras clave y el contenido, la introducción y el resumen. Además, se repite el resultado de la comparación entre el índice de similitud léxico-semántica del resumen-contenido y resumen-introducción, no existiendo diferencia estadística entre estos índices. Esto quiere decir que el resumen, de manera clara y sistemática, se presenta como la unidad textual que da cuenta semánticamente del contenido del artículo de investigación de la misma manera como lo hace con la introducción. Esto permitiría decir, al menos para la muestra estudiada, que la similitud léxico-semántica existente entre el resumen y la introducción es igual a la que se da entre el resumen y el contenido, siendo por ello la similitud con la introducción una especie de predictor de la relación de similitud semántica con el contenido del artículo de investigación científica. De este modo, si quisiéramos conocer la medida en la cual se relaciona el resumen con el contenido, bastaría con establecer el índice de similitud semántica del resumen con la introducción para plantear que la similitud sería muy similar, o al menos, que no se diferenciaría significativamente.

En síntesis, la hipótesis específica H_3 se confirma en un 80%, pudiendo establecerse que se mantiene el patrón de similitudes léxico-semánticas observado en ciencias exactas, con la

leve diferencia que existe diferencia en la similitud entre palabras clave y resumen versus palabras clave-contenido e introducción.

4.3.1.3 Análisis de las diferencias entre los índices de similitud léxico-semánticas en ciencias sociales (H₄)

La última hipótesis específica (H₄), al igual que las anteriores, plantea que al comparar los índices de similitud léxico-semántica entre las variables en estudio de los artículos de investigación científica en las ciencias exactas, obtenidos utilizando un espacio semántico especializado, se esperan diferencias significativas entre los índices de las siguientes relaciones:

Tabla 22. Hipótesis específica de investigación en ciencias sociales

CIENCIAS SOCIALES ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		≠	≠	≠	≠
P+I			≠	≠	≠
P+C				≠	≠
R+I					≠
R+C					

Los resultados obtenidos a través de la prueba estadística, presentados en la Tabla 23, permite establecer que, de las 10 subhipótesis estadísticas en estudio para el área de ciencias sociales, 6 de ellas confirman la hipótesis específica. Por lo tanto, es posible afirmar que la hipótesis H₄ se acepta en un 60%.

Tabla 23. Resultados de la hipótesis específica en ciencias sociales

CIENCIAS SOCIALES ($\alpha < 0,05$)	P+R	P+I	P+C	R+I	R+C
P+R		= 0,630	= 0,164	≠ 0,000	≠ 0,000
P+I			= 0,739	≠ 0,000	≠ 0,000
P+C				≠ 0,000	≠ 0,000
R+I					= 0,970
R+C					

En cuanto a los resultados particulares de las comparaciones estadísticas en esta área de la ciencia, notamos que no existe diferencia significativa entre índices de similitud léxico-semántica entre las variables palabras clave-resumen y palabras clave-contenido e introducción, así como tampoco existe diferencia estadística entre los índices de palabras clave-introducción y palabras clave-contenido. En suma, la relación que se establece entre las palabras clave y las otras variables textuales investigadas no presentan diferencia estadística. Por otra parte, tampoco existe diferencia estadística en los índices de las variables resumen-introducción y resumen-contenido.

Como se desprende de lo expuesto, en el 60% de las comparaciones estadísticas sí determinamos diferencias significativas. Estas diferencias se producen, al igual que en ciencias exactas, al comparar algún índice de similitud léxico-semántica que incluya a la variable resumen con alguna otra variable textual respecto de la variable palabras clave con alguna otra variable textual. Esto confirma, también para esta área, el patrón de similitudes anteriormente identificado para ciencias exactas y ciencias biológicas, en el que las mayores similitudes entre variables son las que incluyen el resumen y las menores las que incluyen a las palabras clave. Además, podemos afirmar que en todas las áreas existe una tendencia a que la comparación entre los índices de similitud entre el resumen-contenido no presente diferencias con el índice de similitud entre el resumen-introducción. De este modo, en ciencias sociales podemos percatarnos que el resumen cumple bastante bien la función informativa, según la cual condensa el significado global del contenido de un artículo mejor que las palabras clave. Por lo anterior, podemos afirmar que en los artículos escritos en las distintas disciplinas de las ciencias sociales, el resumen cumple, del mismo modo que en ciencias sociales y ciencias biológicas, la función de macrosemantizar el contenido semántico global del artículo de investigación, confirmándose los resultados y las interpretaciones expuestas anteriormente.

4.4 Conclusiones

En este apartado presentaremos, en primer lugar, una breve síntesis de los principales aspectos tratados en los capítulos anteriores con el fin de contextualizar en el marco de esta investigación los resultados obtenidos. En segundo lugar, plantaremos las conclusiones

pertinentes acorde con las hipótesis investigadas. En tercer lugar, revisaremos críticamente los resultados obtenidos, identificando algunas limitaciones de esta investigación. Por último, indicaremos aquellos aspectos, que a nuestro juicio, constituyen los aportes y posibles proyecciones de este trabajo.

Desde el punto de vista teórico, esta investigación se caracteriza por su naturaleza multidisciplinar, en tanto hemos construido un marco teórico a partir de diversos estudios afines, tales como la lingüística textual, la semántica, las ciencias cognitivas (en términos de la representación de conocimiento) y el tratamiento informatizado del lenguaje natural. Así, la lingüística textual nos ha permitido, en primer, lugar situar y caracterizar bastante exhaustivamente nuestro objeto de estudio, el artículo de investigación científica escrito en español. La semántica nos ha ayudado a establecer el foco de investigación, según el cual nos ha interesado estudiar las relaciones semánticas identificables en el texto a partir de una semántica de corte asociativo, más específicamente, desde una semántica colocacional basada en co-ocurrencias. Las ciencias cognitivas nos han permitido relacionar este tipo de semántica con una forma de representación del conocimiento de tipo conexionista, ya que este enfoque posibilita reflejar en forma de pautas de activación la estructura y regularidades del ambiente (en nuestro caso textual). Por último, el tratamiento informatizado del lenguaje natural entrega herramientas para el tratamiento del lenguaje natural que puede ser modelado en computadores, congruentemente con la semántica y la perspectiva representacional utilizadas. De este modo, hemos usado el LSA en cuanto modelo algebraico-computacional, que da cuenta de las relaciones semánticas léxicas existentes en un corpus textual, a partir del fortalecimiento de las relaciones semánticas débiles existentes entre las palabras y los párrafos de todos los textos incluidos en el procedimiento matemático-estadístico. El producto obtenido con este modelo vectorial es un espacio semántico multidimensional, el cual permite cuantificar respecto de él las similitudes léxico-semánticas existentes entre palabras, entre párrafos y entre palabras aisladas y párrafos de un corpus de textos especializados.

Ahora bien, según los resultados obtenidos en este trabajo podemos afirmar que la hipótesis general de investigación no se confirmó. De este modo, acorde con la prueba estadística realizada para los datos de similitud léxico-semántica obtenidos a través del Análisis

Semántico Latente, concluimos que no se presentan diferencias significativas, con un 1% de error, entre los textos de la muestra provenientes de las áreas de ciencias exactas, las ciencias biológicas y las ciencias sociales. Ello, a partir del análisis comparativo de los índices de similitudes léxico-semánticas obtenidos a través de la comparación de las variables textuales de los artículos de investigación científica, utilizando el espacio semántico especializado ES-ARTICO.

Como se comprende, este resultado no responde a lo proyectado, debido a que, nacionalmente, se esperaría que en ciencias exactas existiera una notoria rigurosidad y estandarización de la escritura, fundamentalmente entre las partes del artículo científico y en el uso terminológico del léxico, en contraste con los textos provenientes de las ciencias sociales, donde suponemos una menor estandarización retórico-estructural, variabilidad léxica e inestabilidad conceptual. Incluso, recordemos que en el Capítulo 1 habíamos aceptado la generalización realizada por varios autores que planteaba que los artículos varían de una disciplina a otra en grados de estandarización y estilo, De este modo, aquellos producidos en las ciencias conocidas como “duras”, “exactas” o “físicas” cuyo modelo de producción textual tiende a ser más rígido, en comparación con aquellos provenientes desde las ciencias sociales en que existen grupos que han intentado adaptarse al modelo de las ciencias exactas, con diferentes grados de éxito, mientras que otros se resisten a establecer reglas fijas para sus textos (Moyano, 2000; Mogollón, 2003; Swales, 2004).

Una explicación posible para esta homogeneidad detectada a partir de los datos es que los escritores científicos, independiente del área disciplinar en la cual participan, construyen relaciones de similitud léxico-semántica al interior de sus textos de manera similar. Esto significa que, si bien es posible reconocer en las diferentes disciplinas, distintos campos temáticos cuya textualización se realiza por medio de recursos léxicos y terminológicos, e incluso en muchos casos retórico-estructurales, propios de la disciplina, en lo que concierne a las relaciones semánticas de tipo léxico asociativo que se establecen en el nivel textual global de los artículos no existirían diferencias significativas. De este modo, los escritores de las tres disciplinas de la ciencia, según los resultados presentados en los Gráficos 6 y 9, macrosemantizarían el contenido semántico global de sus artículos en igual medida. Lo anterior nos permite pensar que los escritores científicos, autores de los artículos

investigados, pueden presentar una competencia textual similar, la que les permitiría construir relaciones léxico-semánticas al interior de sus textos en términos más o menos parecido. Recordemos que por similitud léxico-semántica entendemos aquí las relaciones de significado de tipo asociativo colocacional de un texto, que se establecen a partir de las regularidades y las fuerzas de relación entre las unidades léxicas de significado, esto, es a partir de la relación cotextual entre las palabras y los párrafos de un texto. Así, la similitud léxico-semántica se explica por el grado de intercambiabilidad contextual o el grado en el cual una palabra puede ser substituida por otra en un contexto textual dado. También, debemos recordar que esta definición de similitud semántica ha sido implementada computacionalmente en una herramienta que solo utiliza las palabras y la frecuencia de co-ocurrencia de tales palabras en los distintos párrafos. Por ello, si bien no podemos saber nada respecto del significado asociado al orden sintáctico en el cual tales palabras aparecen o en cuanto a las relaciones estructurales o funcionales del léxico, como las que propone Mayor (2000), sí podemos asegurar empíricamente, por medio del Análisis Semántico Latente, que existen relaciones de similitud léxico-semántica en los textos estudiados. De manera más específica, esto sucede entre las partes retórico-estructurales (por ejemplo entre el resumen y la introducción, el resumen y el contenido, etc.) de los textos y que, más aun, tales relaciones entre las variables textuales investigadas obtienen índices de similitud léxico-semántica global muy semejantes en los textos producidos por los escritores científicos.

Otro argumento complementario al anterior es que los artículos de investigación que constituyen el corpus y la muestra corresponden a artículos publicados en revistas científicas indexadas acorde con exigentes estándares de calidad internacional. En este sentido, el artículo producido por el autor o autores está sometido, como planteábamos más arriba, a un complejo proceso editorial. A través de él, muchos productores/comprendedores científicos conjugan sus conocimientos disciplinares y sus competencias textuales-discursivas para co-construir un artículo de investigación científica que tenga no solamente calidad en el contenido, sino que también presente calidad en su organización, acorde con la estructura retórica exigida por la revista. Así, es posible que los artículos de investigación tiendan hoy en día, en la mayoría de las disciplinas, a una mayor homogeneización retórico-estructural, y, por ello, también a una mayor similitud léxico-

semántica entre sus partes, en función de un énfasis en la calidad comunicativa de las investigaciones realizadas en las distintas áreas de la ciencia. Por lo tanto, la semejanza entre las áreas, distinguida por la implementación del Análisis Semántico Latente en esta investigación, se justifica en la medida que todos los artículos de investigación de la muestra han sido sometidos a procesos rigurosos de evaluación de pares científicos y a una estandarización de sus contenidos y formatos. De este modo, se tiende hacia una paulatina semejanza en la cual se pierden las diferencias disciplinares propias, al menos en cuanto a las diferencias en los procesos de macrosemantización entre las variables investigadas en las muestras de textos de ciencias exactas, ciencias biológicas y ciencias sociales.

Ahora bien, si retomamos la pregunta planteada en la problematización, respecto de ¿quién o quiénes deben enseñar a los investigadores no iniciados a escribir artículos de investigación científica con la calidad que se requiere hoy en día?, la respuesta más acorde con los argumentos expuestos anteriormente, se inclina hacia la comunidad científica, particularmente a través de los procesos de publicación en revistas de calidad internacional. En este sentido, un investigador que quiera integrarse a una comunidad discursiva científica debe aprender, entre otras cosas, a textualizar su investigación según las normas semántico-textuales asociadas a este tipo de texto y a las normas disciplinares propias de las revistas de su especialidad, asistido en este proceso de aprendizaje por las instancias editoriales de la revista en la cual se desea publicar. Este proceso incluye múltiples evaluaciones y sugerencias que hacen tanto los pares científicos como el comité editorial o editor de la revista, en quienes recae la responsabilidad final del artículo publicado, constituyéndose por ello todo el proceso de escritura científica en una co-construcción semántico-textual, orientada hacia un producto discursivo constructor de conocimiento disciplinar. De este modo, resulta interesante la idea de que el artículo de investigación científica, finalmente publicado, en algunos casos, pueda ser un producto de gran interacción. Esto quiere decir que, para llegar al formato final, el escritor debe atender a múltiples voces que en definitiva le pueden hacer cambiar no su solo su formato y contenido, sino que también su propósito comunicativo original.

En función del argumento anterior, desde una perspectiva crítica, cabe llamar la atención sobre el hecho que este procedimiento de estandarización contribuye a la pérdida de la

personalidad, originalidad y de la expresión auténtica del investigador en pos de la homogeneización de la comunicación de la ciencia. Tal homogeneización se percibe en nuestros resultados, y por ello, cabría preguntarse en qué medida y de qué modo la comunicación científica de hoy, que se sustenta preferentemente en los artículos de investigación, contribuye a la pérdida de identidad de los escritores, tal como lo plantea Ivanic (1998).

Nuestras hipótesis específicas, por otra parte, nos permiten afirmar la presencia de un patrón característico entre las relaciones que se producen entre las variables textuales. En este sentido, hemos explorado estadísticamente cada posible relación entre pares de variables, a través de sus índices de similitud semántica, obteniendo como resultado que existe una tendencia a que en las relaciones en las que aparecen las palabras clave con alguna otra variable textual (resumen, contenido o introducción) no se diferencien significativamente. Esto implica que las palabras clave condensan de manera muy similar el significado de los apartados retóricos. Este resultado, junto al resultado más descriptivo presentado en el Gráfico 9, nos permite afirmar, además, que en la mayoría de las áreas estas relaciones entre variables no solo son similares estadísticamente sino que también, en todas las áreas los valores promedios de similitud semántica son los más bajos obtenidos en esta investigación.

Otra conclusión, a partir de estos datos, señala que las palabras clave, al menos en la muestra de artículos estudiados en esta investigación, no tienen una clara función macrosemantizadora del significado global del contenido. Siendo esto así, se puede argumentar que lo más probable es que las palabras clave funcionen situando el artículo de investigación en un campo disciplinar, temático o procedimental, haciendo muy poca referencia de ello en el interior del artículo. Cabe señalar que el método de cuantificación de las similitudes se funda en la colocación cotextual de las palabras en el texto, siendo, de este modo, altamente probable que esta función de las palabras clave no sea captada por los valores de similitud léxico-semántica que entrega el LSA. Así, se puede claramente dar el caso de que las palabras clave sí macrosemanticen el significado global del texto, pero no lo hagan en términos de las relaciones léxicas intratextuales, sino que en función de relaciones

de macrosemantización más abstractas aún, de tipo exógenas al texto y de carácter eminentemente intertextual e incluso interdiscursivo.

Cabe también argumentar que las palabras clave cumplirían con otra función específica. Ella sería de corte más persuasivo que informativo, esto es, los escritores o incluso los comités editoriales utilizan o sugieren palabras o frases clave que no se relacionan fuertemente con el texto, pero que sí despiertan el interés de un posible lector, cumpliéndose con ello un primer paso de acercamiento a la lectura del texto por parte de otros científicos e inclusión de tales ideas en el campo disciplinar particular. Por último, debemos señalar a este respecto, que estas explicaciones deben ser, por una parte, estudiadas cualitativamente con la finalidad de describir con más detalle esta función de las palabras clave y, por otra, empíricamente con el fin de establecer algún algoritmo que pueda dar cuenta de estas relaciones aún más débiles al interior de los textos, pero fuertes en términos de la práctica discursiva.

Otro resultado interesante que se desprende de las hipótesis específicas es que no existe diferencia significativa, en todos los artículos investigados para las tres áreas de la ciencia, entre resumen-contenido y resumen-introducción. Este resultado nos permite afirmar que el “género incrustado” resumen del artículo de investigación científica, como lo plantea Swales (1990), cumple su función de dar cuenta del significado global del artículo, de la misma manera cómo da cuenta del significado global de la introducción para todas las áreas. Cabe mencionar que en todas las áreas, estas relaciones son las que obtienen los mayores valores de similitud semántica, con lo cual se confirma esta función macrosemantizadora del resumen. Además, nuestros datos confirman la noción de resumen orientada al lector, propuesta por van Dijk (1983), según la cual el hecho de que el resumen se encuentre al inicio de la lectura ayuda al lector a formarse una hipótesis sobre el tópico del discurso del episodio. En nuestro caso particular, hemos comprobado que el grado de macrosemantización del resumen es alto para contenido e introducción, por lo que, para este tipo de textos en las áreas investigadas, las hipótesis de los lectores serían cumplidas probablemente en un alto grado.

En este sentido, nos cabe plantear nuevamente el valor de este hallazgo, puesto que, si bien es nocionalmente compartido este resultado, aquí se ha probado empíricamente a

través de una herramienta computacional capaz de obtener resultados semejantes a los planteados por quienes se han dedicado a la investigación de esta relación textual. De este modo, el LSA, utilizando datos matemáticos-estadísticos a partir de los textos seleccionados, permite de una manera bastante precisa y económica, en términos computacionales, establecer las fuerzas de relación entre los componentes léxicos (palabras y párrafos fundamentalmente), asignándoles valores de similitud semántica y entregando como producto el grado de similitud léxico-semántica entre componentes textuales e incluso retórico-estructurales, como ha sido lo realizado en nuestra investigación. De este modo, resaltamos el valor de la herramienta que hemos utilizado, ya que, comparativamente con modelos simbólicos o basados en análisis lógico-proposicionales desde una perspectiva semántica computacional (como los presentados en el Capítulo 2), es mucho más económica en términos de procesamiento y eficiente en términos de tiempo y costos de programación. Ello, puesto que no se necesita programar cada una de las reglas de restricción y selección léxico-semánticas y/o retórico-estructurales, obteniendo, además, como hemos visto en esta y otras investigaciones mencionadas en la primera parte de nuestro trabajo, resultados bastante plausibles con lo esperado teóricamente. En este sentido, desde el punto de vista del lingüista (u otro experto sin experticia computacional avanzada) que desea hacer investigaciones en torno a las relaciones de similitud léxico-semánticas entre unidades lingüísticas particulares (incluyendo palabras, párrafos y textos completos), con esta herramienta solo requiere un corpus de investigación y la sistematización de las comparaciones que desea realizar (dando por hecho que se tiene el programa y equipo adecuado para usar el LSA y los programas para las consultas, además de las licencias correspondientes (al respecto ver <http://lsi.research.telcordia.com/>).

En relación con este resultado, también hemos planteado que la similitud semántica existente entre resumen-introducción y resumen-contenido no es significativa. Esto nos permite sugerir una notoria relación entre contenido e introducción que puede ser identificada a partir del resumen. Ello significa que existe un patrón de similitud léxico-semántica entre la introducción y el contenido que es captado por el resumen, permitiéndonos de este modo plantear que las introducciones están muy relacionadas con el contenido en nuestros artículos de investigación, y que, por ello, la relación resumen-introducción se convierte en un buen predictor de la relación resumen-contenido. En otras

palabras, conociendo la similitud que existe entre el resumen y la introducción en artículos de investigación científica es posible proyectar una similitud muy similar respecto del contenido. Esto puede ser de mucha utilidad en términos de evaluación de artículos de investigación científica en el contexto de las publicaciones científicas, ya que contando con la introducción y el resumen se podría conocer el grado de similitud entre el resumen y el texto, teniendo, con ello, una información complementaria respecto de la calidad de las relaciones léxico-semánticas del texto. En este sentido, en nuestra muestra hemos tenido, para esta comparación de variables textuales, un resultado promedio de 0,5, valor que puede constituirse, luego de mayores comprobaciones empíricas, en un parámetro de evaluación de la cantidad de similitud semántica que debieran presentar los artículos que postulan a ser incluidos en publicaciones científicas indexadas.

A continuación nos referiremos a las limitaciones que, a nuestro juicio, se presentan en el desarrollo de esta investigación. Pensamos que esta información es relevante para comprender mejor los resultados obtenidos y descritos en esta investigación, así como para quienes deseen emprender investigaciones similares en el futuro.

En cuanto a los aspectos teóricos, hemos desarrollado aquí una visión panorámica de algunos problemas teóricos centrales respecto de la investigación. Particularmente, hemos presentado la noción de texto y discurso desde múltiples orientaciones disciplinares, las que nos permitieron comprender que el texto es parte integrante y fundamental del proceso discursivo. En este sentido, hemos decidido, debido a nuestra aproximación computacional y convencimiento de que a partir del texto se pueden estudiar relaciones de significado, optar por una noción restringida de texto en la que hemos definido al texto prioritariamente en cuanto a producto (*ergon*) de la actividad psicosociodiscursiva. Sin embargo, debemos señalar que no se desconocen las potencialidades teóricas del texto de ser algo más que materialidad, particularmente en cuanto a su íntima relación con el contexto social y psíquico de los individuos que lo producen y lo comprenden.

Con relación a lo anterior, estamos conscientes de que este estudio, a pesar de usar un amplio corpus textual que permite configurar un espacio semántico especializado respecto del cual se hacen las comparaciones entre las variables textuales, es restringido en cuanto al estudio del texto sin su contexto de producción y circulación. Ello nos obliga a declarar que

las generalizaciones realizadas respecto de los datos se aplican a los textos estudiados según el establecimiento de las variables presentadas en el apartado metodológico y sin mayor atención a la situación psicosociodiscursivas que lo afectan. A pesar de lo anterior, algunos resultados nos permiten acercarnos a interpretaciones de este tipo, pero siempre en el entendido que son interpretaciones *ad hoc*.

Otro aspecto que podría surgir como debilidad es la noción de significado colocacional aquí utilizada. Hemos presentado aquí diversas teorías del significado y de las relaciones de significado que ofrece la literatura, acogiendo, de entre ellas, una noción más bien clásica de corte empirista respecto del significado. Tal noción de significado, que sabemos no es la más aceptada, sobre todo por quienes defienden una postura más deductiva simbólica, sin embargo, nos ha sido útil en el desarrollo de esta investigación particular, aceptando que, con ello, no hemos podido afirmar nada respecto del significado en términos proposicionales, acorde con una semántica lógica; o respecto de las relaciones metafóricas, según una semántica más cognitiva; o incluso respecto de las relaciones estructurales o funcionales, siguiendo una semántica de corte estructuralista o composicional. En este sentido, solo nos cabe decir que, acorde al avance de la computación y las herramientas computacionales existentes, estas semánticas han sido escasamente implementadas y, si lo han sido, responden a sistemas computacionales de acceso muy restringido y, fundamentalmente, realizados en base a reglas que no necesariamente pueden ser aplicadas a corpus en español como es nuestro caso.

Por último, de nuestra presentación realizada respecto del LSA, podría desprenderse que adherimos a la postura más fuerte de los psicólogos de Boulder (Landauer & Dumais, 1996, 1997; Landauer, Foltz & Laham, 1998; Kintsch, 1998, 2000, 2001, 2002; Rehder, Schreiner, Wolfe, Laham, Landauer & Kintsch, 1998; Wolfe, Schreiner, Rehder, Laham, Foltz, Kintsch & Landauer, 1998; Landauer & Psozka, 2000; Landauer, 2002; Quesada, Kintsch & Gómez, 2002; Quesada, 2003) quienes plantean que este método de identificación y cuantificación de similitudes semánticas sería una teoría plausible de la adquisición y representación del lenguaje, pudiendo a través de sus resultados explicar el modo en que los individuos construyen relaciones en el lenguaje a través de sistemas estadísticos incorporados en nuestro sistema cognitivo. Sin embargo, en esta investigación

nos hemos planteado escépticos ante tal posibilidad, particularmente porque creemos que los procesos cognitivos del individuo no pueden ser modelados por las computadoras y que, solo en algunos aspectos más bien prácticos, la noción de computación (en cuanto metáfora explicativa) asociada a los modelos cognitivos es permisible.

En otras palabras, pensamos que el LSA no da cuenta de cómo se lleva a cabo ningún tipo de proceso cognitivo, sino que creemos en la utilidad del LSA como herramienta que permite obtener resultados de corte empírico cercanos a los que los seres humanos han obtenido para tareas similares, siendo tales resultados útiles, pero absolutamente dependientes de la interpretación humana.

En cuanto a los aspectos metodológicos, podemos plantear que una limitación de esta investigación está en la representatividad de los datos estudiados. Por una parte el acceso a los datos en formato electrónico fue más difícil de lo que se planificó en un principio, fundamentalmente porque muchas de las revistas que tradicionalmente publicaban en español han incluido en sus publicaciones artículos de investigación en inglés, especialmente en el área de ciencias exactas. De este modo, la cantidad de artículos de investigación científica en ciencias exactas que estuviera acorde a los criterios de selección establecidos fue mucho menor de lo esperado. Debido a lo anterior, debimos, incluso, incorporar artículos que no se encontraban en el rango de tiempo estimado en primera instancia. A su vez, esto influyó en la muestra de artículos de esta área para la investigación, ya que acorde con la muestra estratificada representativa del área se tuvo que seleccionar al azar solo 5 textos de esta área. Lo cual, en comparación a las otras áreas, representaba cerca del 50%. Sin embargo, optamos por confiar en el método estadístico y mantener el 97% de confiabilidad entregado por los cálculos. Esta baja cantidad de textos pudo haber afectado en alguna medida los resultados, sin embargo, tal como observamos en el análisis de los resultados, los índices de similitud semántica se mantienen similares y constantes con las otras áreas, lo que nos permite asumir que, de incorporar mayor cantidad de textos, las diferencias no serían cuantitativamente significativas.

Otro aspecto que debemos mencionar en este apartado, es la dificultad que constituyó comprender las potencialidades reales del LSA. Primeramente porque responde a un desarrollo tecnológico difundido para y en inglés, en un campo disciplinar completamente

ajeno al que estamos ambientados. Por lo anterior, se requirió de un buen esfuerzo para comprender las potencialidades del LSA en cuanto herramienta y en cuanto a los aspectos lingüísticos involucrados (no siempre expresamente comentados por los autores). En este sentido, comprender, al menos en términos generales, la formalización matemático-algebraica del LSA fue un desafío que sigue estando presente. Por otra parte, la implementación computacional de esta herramienta fue desarrollada en el Laboratorio de Ciencias Cognitivas del Instituto de Sistemas Inteligentes de la Universidad de Memphis, quienes generosamente nos permitieron utilizar la versión del LSA con la cual ellos han desarrollado sus investigaciones, limitándonos a adaptar nuestros procedimientos de investigación a lo que ellos han avanzado en la implementación computacional de la herramienta y confiando en la experticia de quienes han escrito los programas que se han utilizado en esta investigación.

Con relación a las fortalezas de esta investigación podemos plantear, en primer lugar, que nuestro objeto de estudio, el artículo de investigación científica, si bien ha sido estudiado profundamente por diversas orientaciones de los estudios del discurso especializado, todas ellas han sido llevadas a cabo fundamentalmente desde perspectivas lingüístico-textuales, retóricas y sociocognitivas y, la mayoría de ellas, desde un enfoque cualitativo comparado interlenguas, a partir de muestras ejemplares de textos y con resultados basados en estadísticas descriptivas muy básicas. A diferencia de lo anterior, en nuestro trabajo hemos asumido, ante nuestro objeto de estudio, una actitud multidisciplinaria orientada hacia la aplicación de herramientas tecnológicas innovadoras. Esto nos ha posibilitado trabajar con un gran corpus digitalizado no marcado de artículos de investigación científica, respecto de los cuales se han podido realizar procedimientos de tipo multivariante, utilizando, por ello, una estadística inferencial implementada computacionalmente en el LSA y en el cálculo de los índices de similitud léxico-semántica. De este modo, hemos podido estudiar las relaciones de similitud léxico-semántica entre las variables textuales de artículos de investigación científica, siendo este tipo de estudio un nuevo aporte al conocimiento de este tipo textual, particularmente en español.

En relación con lo anterior, esta investigación ha abordado un aspecto teórico relativo al estudio del artículo científico, que es muy poco conocido en lingüística textual y en

semántica. Permitiendo abrir un nuevo ámbito de investigación que puede aportar nuevo conocimiento a los estudios del discurso y a la construcción del significado en textos especializados. En particular destacamos los resultados que han demostrado la existencia de una alta homogeneidad en las similitudes léxico-semánticas en los artículos utilizados en la muestra de investigación. Estos datos pueden fortalecer las discusiones teóricas y político-científicas, respecto de la conveniencia de la estandarización de los productos científicos o de la pérdida de identidad del productor de los textos científicos en función de una supuesta modernización y mejoramiento de la calidad comunicativa a través de sus productos científicos. Por último, esto mismo conlleva el cuestionamiento respecto de cómo debe ser enseñando el artículo de investigación científica.

Otro aspecto destacable en este trabajo dice relación con el desarrollo de un espacio semántico especializado para el español y los programas de consulta de comparaciones textuales, según los cuales se puede hoy realizar investigaciones de corte semántico asociativo en nuestra lengua. Es probable que en el desarrollo de esta investigación no se haya percibido el tiempo, la dedicación y el esfuerzo que implicó la implementación del LSA para el español, sin embargo, se debe señalar que este espacio semántico, correspondiente a un dominio particular del discurso y a un tipo textual determinado, es, para nuestra lengua, un avance de considerable importancia, teniendo en cuenta que quienes han trabajado en este ámbito con corpus en español lo han hecho con corpus muy limitados y con fines relacionados más con la comprobación de herramientas computacionales que con objetivos investigativos centrados en aspectos lingüístico-textuales como lo hemos desarrollado aquí.

En cuanto a las proyecciones, visualizamos como de absoluta necesidad complementar esta investigación con estudios de corte cualitativo que nos permitan confirmar los resultados obtenidos y caracterizar con mayor detalle los tipos de similitudes semánticas que el LSA ha cuantificado. Tal investigación no se ha realizado aquí, fundamentalmente, porque representaría un esfuerzo que sobrepasa los límites temporales y de focalización que para esta investigación se han planteado.

Otra proyección interesante está en relación con la implementación de este tipo de herramienta en combinación con otras herramientas computacionales que permitan dar

cuenta más eficazmente de aspectos semánticos, retórico-estructurales y/o de producción y circulación de los textos. Así, se podría pensar, como ya se ha realizado en algunos casos para el inglés, en combinaciones de algoritmos que permitan construir modelos híbridos orientados a la determinación de los significados de distintos tipos y niveles al interior del texto, por ejemplo niveles de coherencia textual, esto es significados asociados a las estructuras y funciones de los párrafos en el texto, identificación de significados asociados al orden de las oraciones o a las relaciones interpersonales y a las actitudes de los productores presentes en el texto. Lo anterior, sin embargo, requeriría de un fuerte equipo multidisciplinar y un foco de trabajo más orientado hacia la lingüística computacional que hacia la lingüística del texto o del discurso.

Por último, algunas aplicaciones prácticas de este conocimiento y herramientas computacionales podrían orientarse hacia la construcción de programas computacionales que permitan asistir en la evaluación de los artículos de investigación para congresos, en la creación de programas que evalúen respuestas abiertas, basándose en la similitud léxico-semántica, en la implementación de programas que determinen la similitud entre un resumen y un texto de origen, en la construcción de mejores buscadores de información en internet, etc. Cabe hacer notar que, si bien, muchas de estas aplicaciones han sido creadas por equipos académico internacionales, todas han sido realizadas para la lengua inglesa. Por lo mismo, su construcción en español no solo es una buena proyección, sino que también un buen desafío para quienes quieran emprender tales proyectos.

REFERENCIAS BIBLIOGRÁFICAS

- Adam, J.M. (1992). *Les textes, types et prototypes: Récit, description, argumentation, explication et dialogue*. Paris: Nathan.
- Adam, J.M. (1999). *Linguistique Textuelle. Des genres de discours aux textes*. Paris: Nathan.
- Albentosa, J. & Moya, A. (2000). La reducción del grado de transitividad de la oración en el discurso científico en lengua inglesa. *Revista Española de Lingüística*, 2, 445-468.
- Allen, J. (1995). *Natural language understanding*. Reedwood City: Benjamin Cummings.
- Álvarez, G. (1995). Competencia discursiva y textual del hablante nativo. *Revista de Lingüística Teórica y Aplicada*, 33, 5-14.

- Álvarez, G. (2001). Textos y discursos. Introducción a la lingüística del texto. Concepción: Editorial Universidad de Concepción.
- Anthony, L. (1999). Writing research article introductions in software engineering: How accurate is the standar model? IEE Transactionsof Professional Communication, 42, 38-46.
- 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.
- Aristóteles (1966). Del sentido y lo sensible. De la memoria y el recuerdo. Buenos Aires: Aguilar.
- Austin, J. (1982). Cómo hacer cosas con palabras: Palabras y acciones. Barcelona: Paidós.
- Baayen, H. & Lieber, R. (1997). Word frequency distribution and lexical semantics. Computers and the Humanities, 30, 281-291.
- Bajtín, M. (1982). Estética de la creación verbal. México D.F.: Siglo XXI.
- Baldinger, K. (1970). Teoría semántica. Hacia una semántica moderna. Madrid: Romania.
- Bally, Ch. & Sechehaye, A. (1983). Prefacio a la primera edición francesa. En F. Saussure, Curso de lingüística general (pp.59-62). Madrid: Alianza.
- Bally, Ch. (1940). L'arbitraire du signe. Valeur et signification. Le Français Moderne, VII, 193-206.
- Baykoucheva, S. (2001). Internet resources for scientific writing [en línea]. Disponible en: http://pubs.acs.org/subscribe/journals/ci/31/special/02sb_inet.html
- Bazerman, Ch. (1988). Shaping written knowledge: The genre and activity of the experimental article in science. Madison: The University of Wisconsin Press.
- Benítez, R. & Velásquez, M. (2000). Un análisis microestructural de la producción escrita de alumnos de 6º y 8º año de educación general básica. Lenguas Modernas, 26-27, 95-117.
- Berkenkotter, C. & Huckin, T.N. (1995). Genre knowledge in disciplinary communication: Cognition/Culture/Power. Hillsdale, N.J.: Lawrence.
- Bernárdez, E. (1982). Introducción a la lingüística del texto. Madrid: Espasa-Calpe.
- Bernárdez, E. (1995). Teoría y epistemología del texto. Madrid: Cátedra.
- Bhatia, V.K. (1993). Analysing genre: Language use in professional settings. London: Longman.

- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus linguistics: Exploring language structure and use*. Cambridge: Cambridge University Press.
- Biber, D. (2005). Paquetes léxicos en textos de estudio universitario. Variación entre disciplinas académicas. *Revista Signos*, 38(57), 19-29.
- Blei, D., Ng, A. & Jordan, M. (2002). Latent dirichlet allocation. Informe Técnico CSD-02-1194. Berkley: Division of Computer Science, University of California,
- Bod, R. (2003). Introduction to elementary probabilistic theory and formal stochastic language theory. En R. Bod, J. Hay & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 11-37). London: MIT Press.
- Bolívar, A. (2000). Homogeneidad versus variedad en la estructura de los resúmenes de investigación para congresos. *Akadosmos*, 2, 121-138.
- Bonilla, S. (1997). Estudio preliminar. En R. De Beaugrande & W. Dressler (Eds.), *Introducción a la lingüística del texto* (pp. 7-25). Barcelona: Ariel.
- Breál, M. (1883). Les lois intellectuelles du langage. Fragment de sémantique. *Annuaire de l'Association pour l'encouragement des études grecques en France*, 132-142 [en línea]. Disponible en: http://www.chass.utoronto.ca/epc/langueXIX/breal/breal_lois.htm
- Breál, M. (1897). *Essai de sémantique*. (Sciences de significations) [en línea]. Disponible en : <http://www.chass.utoronto.ca/epc/langueXIX/breal/>
- Brinker, K. (1979). Zur Gegenstandbestimmung und Aufgabenstellung der Textlinguistik. En J. Petöfi (Ed.), *Text vs. sentence. Basic questions of textlinguistics* (pp. 3-12). Hamburg: Buske.
- Brinker, K. (1988). *Linguistische Textanalyse*. Berlin: E. Schmidt.
- Bunge, M. (1981). *La ciencia, su método y su filosofía*. Buenos Aires: Ediciones Siglo Veinte.
- Budanitsky, A. & Hirst, G. (2001). Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures. En *Workshop on WordNet and other lexical resources: Second Meeting of the North American Chapter of the Association for Computational Linguistics*, 29-34.

Cabré, M.T. (1999). El discurs especialitzat o la variació funcional determinada per la temàtica: Noves perspectives. En T. Cabré (Ed.), *La terminología. Representación y comunicación. Una teoría de base comunicativa y otros artículos* (pp. 151-173). Barcelona: IULA.

Cabré, M.T. (2002). Textos especializados y unidades de conocimiento: Metodología y tipologización. En J. García & M. Fuentes (Eds.), *Texto, terminología y traducción* (pp. 122–187). Barcelona: Almar.

Cabré, M. & Feliu, J. (2001). *Terminología científico-técnica: Reconocimiento, análisis y extracción de información formal y semántica*. Barcelona: Universitat Pompeu Fabra.

Cademártori, Y. & Venegas, R. (2004). El uso de la nominalización en manuales de tres áreas de especialización técnico-científica: Un acercamiento terminológico. Ponencia presentada en el III Congreso de la Asociación Latinoamericana de Estudios del Discurso, Valdivia, Chile.

Cademártori, Y, Parodi, G. & Venegas, R. (En prensa). El discurso escrito y especializado: Caracterización y función de las nominalizaciones en los manuales técnicos. *Literatura y Lingüística*.

Calsamiglia, H. (Coord.) (1998). *Análisis discursivo de la divulgación científica* [en línea]. Disponible en: <http://www.upf.es/dtf/personal/danielcass/anali.htm>.

Calsamiglia, H., Bonilla, S., Cassany, D., López, C. & Martí, J. (2001). Análisis discursivo de la divulgación científica. En J. de Bustos, P. Charadeau, J. Girón, S. Iglesias & C. López (Coords.), *Lengua, discurso, texto. I Simposio Internacional de Análisis del Discurso*. Madrid: Visor Libros.

Campanario, J. (2004). El enfoque conexionista en Psicología Cognitiva y algunas aplicaciones sencillas en Didáctica de las Ciencias. *Enseñanza de las Ciencias*, 22(1), 93-104.

Caravedo, R. (1999). *Gramática española: Enseñanza e investigación. Apuntes Metodológicos*. Salamanca: Universidad de Salamanca.

Carnap, R. (1958). *Meaning and necessity: A study in semantics and modal logic*. Chicago: Phoenix Books in Philosophy.

- Cassany, D., López, C. & Martí, J. (2000). La transformación divulgativa de redes conceptuales científicas: Hipótesis, modelo y estrategias. *Discurso y Sociedad*, 2(2), 73-103.
- Chafe, W. (1994). *Discourse, consciousness and time*. Chicago: The University of Chicago Press.
- Charaudeau, P. & Maingueneau, D. (2002). *Dictionnaire d'analyse du discours*. Paris: Éditions du Seuil.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton & Co.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA.: MIT Press.
- Chomsky, N. (1991). Linguistics and cognitive science: Problems and mysteries. En A. Kasher (Ed.), *The chomskyan turn* (pp. 26-53). Oxford: Blackwell Publishers.
- Chu, B. (1996). Introductions in state-of-the-art, argumentative, and teaching tips TESL journal articles: Three possible sub-genre of introduction? Hong Kong: City University of Hong Kong.
- Ciapuscio, G. (1994). *Tipos textuales*. Buenos Aires: EUDEBA.
- Ciapuscio, G. (2000). Hacia una tipología del discurso especializado. *Discurso y Sociedad*, 2(2), 39-71.
- Ciapuscio, G. (2003). *Textos especializados y terminología*. Barcelona: IULA.
- Ciapuscio, G. (2005). La noción de género en la lingüística sistémico funcional y en la lingüística textual. *Revista Signos*, 38(57), 31-48.
- Ciapuscio, G. & Otañi, I. (2002). Las conclusiones de los artículos de investigación desde una perspectiva contrastiva. *RILL*, 15, 117-133.
- Collins, A. & Loftus, E. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Collins, A. & Quillian, R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.
- CONICYT (2004). *Indicadores científicos y tecnológicos [en línea]*. Disponible en: <http://www.conicyt.cl/bases/indicadores/>
- Conrad, S. & Biber, D. (1998). Multi-dimensional methodology and the dimensions of register variation in English. En S. Conrad & D. Biber (Ed.), *Variation in English: Multi-dimensional studies* (pp.13-42). London: Longman.

- Corominas, J. & Pascual, J. (1986). Diccionario crítico etimológico castellano e hispánico. Madrid: Gredos.
- Corredor, C. (1999). Filosofía del lenguaje: Una aproximación a las teorías del significado del siglo XX. Madrid: Visor.
- Coseriu, E. (1986). Introducción a la lingüística. Madrid: Gredos.
- Coseriu, E. (1987). Gramática, semántica, universales: Estudios de lingüística funcional. Madrid: Gredos.
- Coseriu, E. (1991). El hombre y su lenguaje. Estudios de teoría y metodología lingüística. Madrid: Gredos.
- Coulthard, M. (1977). An introduction to discourse analysis. London: Longman.
- Crombie, W. (1985). Discourse and language learning: A relational approach to syllabus design. Oxford: Oxford University Press.
- Cruse, D. (1986). Lexical semantics. Cambridge: Cambridge University Press.
- Daneš, F. & Viehweger, D. (1976). Probleme der Textgrammatik. Berlin: Akademie-Verlag.
- De Beaugrande, R. (1997). New Foundations for a science of text and discourse: Cognition, communication, and the freedom of access to knowledge and society. New Jersey: Ablex.
- De Beaugrande, R. (2005). Text, discourse, and process. Toward a multidisciplinary science of texts [en línea]. Disponible en: <http://beaugrande.bizland.com/TDPOpening.htm>
- De Beaugrande, R. & Dressler, W. (1997). Introducción a la lingüística del texto. Barcelona: Ariel.
- De Vega, M. (1984). Introducción a la psicología cognitiva. Madrid: Alianza.
- De Vega, M. (2002). Comprensión del discurso: Representación del significado y métodos de investigación. Seminario dictado por la Universidad de Concepción, Chile.
- Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.
- Dennis, S. (2003). A comparison of statistical models for the extraction of lexical information from text corpora. Artículo presentado en The Twenty-fifth Conference of the Cognitive Science Society.

- Dennis, S. & Harrington, M. (2001). The syntactic paradigmatic model: A distributed instance-based model of sentence processing. Artículo presentado en el 2nd Workshop on Natural Language Processing and Neural Networks, Tokyo.
- Dessus, Ph. (2000). Construction de connaissances par exposition à un cours avec LSA. In *Cognito*, 18, 27-34.
- Dressen, D. & Swales, J. (2000). "Geological setting/cadre géologique" in English and French petrology articles: Muted indications of explored places. En A. Trosberg (Ed.), *Analysing professional genres*. Amsterdam: Benjamins.
- Dressler, W. (1972). Textgrammatische invarianz in Übersetzungen? En E. Gülich & W. Raible (Eds.), *Textsorten: Differenzierungskriterien aus linguistischer Sicht* (pp. 98-106). Frankfurt: Athenäum.
- Dressler, W. (1978). *Current trends in textlinguistics*. Berlin: W. de Gruyter.
- Dressler, W. & Schmidt, S. (1973). *Textlinguistik: Kommentierte Bibliographie*. München: Fink.
- Dudley-Evans, T. (1986). Genre analysis: An investigation of the introduction and discussion sections of MSc. dissertation. En M. Coulthard (Ed.), *Talking about text* (pp. 128-145). Birmingham: English Language Research, Birmingham University.
- Dumais, S. (1997). Using latent semantic indexing (LSI) for information retrieval, information filtering and other things [en línea]. Disponible en: <http://lsa.colorado.edu>
- Dumais, S., Landauer, T. K. & Littman, M. (1996). Automatic cross language information retrieval using latent semantic analysis [en línea]. Disponible en: <http://lsa.colorado.edu>
- EAGLES (1996a). *Text corpora working group reading guide*. Pisa: EAGLES (Expert Advisory Group on Language Engineering).
- EAGLES (1996b). *Preliminary recommendations on corpus typology*. Pisa: EAGLES (Expert Advisory Group on Language Engineering).
- Eggins, S. & Martin, J.R. (2003). El contexto como género: Desde la perspectiva de la lingüística funcional. *Revista Signos*, 36(54), 185-205.
- Elejalde, A. (1998). *Discurso literario y discurso académico* [en línea]. Disponible en: <http://macareo.pucp.edu.pe/~elejalde/ensayo/dlitdacad.html>.
- Escandell, M.V. (2004). *Fundamentos de semántica composicional*. Barcelona: Ariel.
- Fairclough, N. (1992). *Discourse and social change*. Cambridge: Polity.

- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. London: Longman.
- Fauconier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge: Cambridge University Press.
- Fernández, A., Hervás, S. & Báez, V. (1979). *Introducción a la semántica*. Madrid: Cátedra.
- Fillmore, Ch. (1971). *Hacia una teoría moderna de los casos*. En H. Contreras (Comp.), *Los fundamentos de la gramática transformacional* (pp. 45-65). México, D.F.: Siglo XXI.
- Fodor, J.A. (1977). *Semantics*. New York: Crowell.
- Fodor, J.A. (1984). *Observation reconsidered*. *Philosophy of science*, 51, 23-43.
- Foltz, P. (1990). *Using latent semantic indexing for information filtering*. En R.B. Allen (Ed.), *Actas de The Conference on Office Information Systems* (pp. 40-47). Cambridge, MA.: MIT Press.
- Foltz, P., Kintsch, W. & Landauer, T.K. (1998). *The measurement of textual coherence with latent semantic analysis*. *Discourse Processes*, 25, 285-307.
- Foucault, M. (1972). *The archaeology of knowledge*. London: Routledge.
- Foucault, M. (1999). *Orden del discurso*. Barcelona: Tusquets.
- Frege, G. (1972). *Lógica y semántica*. Valparaíso: EUVSA.
- Fries, U. (1972). *Textlinguistik*. *Linguistik und Didactic*, 7, 219-234.
- García-Madruga, J. (1992). *Introducción a la edición española*. En D. Rumelhart & J. McClelland (Eds.), *Introducción al procesamiento distribuido en paralelo* (pp. 11-30). Madrid: Alianza.
- Gardner, H. (1988). *La nueva ciencia de la mente: Historia de la revolución cognitiva*. Barcelona: Paidós.
- Garrido, M. (1977). *Lógica simbólica*. Madrid: Tecnos.
- Ge, X. & Iwata, S. (2002). *Learning the part of objects by auto-association*. *Neural Networks*, 15, 285-295.
- Geisler, C. (1994). *Academic literacy and the nature of expertise*. Hillsdale, N.J.: Erlbaum.
- Gernsbacher, M.A. (1990). *Language comprehension as structure building*. Hillsdale, N.J.: Erlbaum.

- Gernsbacher, M.A. (1996). The structure-building framework: What it is, what it might also be, and why. En B. Britton & A. Graesser (Eds.), *Models of understanding text* (pp. 289-311). Mahwah, N.J.: Erlbaum.
- Gläser, R. (1982). The problem of style classification in LSP (ESP). Ponencia presentada en el 3rd European Symposium on LSP, Copenhagen.
- Gläser, R. (1993). A multi-level model for a typology of LSP genres. *Fachsprache. International Journal of LSP*, 15(1-2), 18-26.
- Gnutzmann, C. & Oldenburg, H. (1991). Contrastive text linguistics in LSP-research: Theoretical considerations and some preliminary findings. En H. Schröder (Ed.), *Subject-oriented texts. Language for special purposes and text theory* (pp. 103-136). Berlin: W. de Gruyter.
- Gómez de Silva, G. (1995). *Breve diccionario etimológico de la lengua española*. México D.F.: Fondo de Cultura Económica.
- Gómez, L. (1997a). Dimensión social de la comprensión verbal (propuesta teórica para un mejoramiento cualitativo del aprendizaje escolar). En M. Peronard, L. Gómez, G. Parodi & P. Núñez (Comps.), *Comprensión de textos escritos: De la teoría a la sala de clases* (pp. 43-53). Santiago: Andrés Bello.
- Gómez, L. (1997b). Elementos para un marco teórico de la comprensión textual. En M. Peronard, L. Gómez, G. Parodi & P. Núñez (Comps.), *Comprensión de textos escritos: De la teoría a la sala de clases* (pp. 95-117). Santiago: Andrés Bello.
- Gotti, M. (2003). *Specialized discourse. Linguistic features and changing conventions*. Bern: Peter Lang.
- Graesser, A.C., Trabasso, T. & Singer, M. (1994). Constructing inference during narrative text comprehension. *Psychological Review*, 101, 371-395.
- Graesser, A.C., Millis, K. & Zwaan, R. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163-189.
- Graesser, A.C. & Wiemer-Hastings, K. (1999). Situation models and concepts in story comprehension. En S. Goldman, A. Graesser & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence* (pp. 77-92). Mahwah, N.J.: Erlbaum.

- Graesser, A.C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N. & the Tutoring Research Group (2000). Using latent semantic analysis to evaluate the contributions of students in Autotutor. *Interactive Learning Environments*, 8, 149-169.
- Graesser, A.C., Person, N.K., Harter, D. & Tutoring Research Group (2001). Teaching tactics and dialog in Autotutor. *International Journal of Artificial Intelligence in Education*, 12, 257-279.
- Graesser, A.C., Wiemer-Hastings, P. & Wiemer-Hastings, K. (2001). Constructing inferences and relations during text comprehension. En T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 249-271). Amsterdam: Benjamins.
- Graesser, A.C., Hu, X., Olde, B.A., Ventura, M., Olney, A., Louwerse, M., Franceschetti, D.R. & Person, N. (2002). Implementing latent semantic analysis in learning environments with conversational agents and tutorial dialog. *Actas de The Twenty-fourth Annual Meeting of the Cognitive Science Society* (pp. 37). Mahwah, N.J.: Erlbaum.
- Greimas, A. (1966). *Semántica estructural. Investigación metodológica*. Madrid: Gredos.
- Grice, H.P. (1975) *Logic and conversation*. En P. Cole & J.L. Morgan (Eds.), *Syntax and semantics: Speech acts* (pp. 41-58). New York: Academic Press.
- Griffiths, T.L. & Steyvers, M. (2002). A probabilistic approach to semantic representation. Artículo presentado en *The Twenty-fourth Annual Conference of Cognitive Science Society*, Georg Mason University, Fairfax, VA.
- Gülich, E. & Raible, W. (1977). *Linguistische Textmodelle. Grundlagen und Möglichkeiten*. München: Fink.
- Gunnarsson, B.L. (2000). Análisis aplicado del discurso. En T. van Dijk (Ed.), *El discurso como interacción social* (pp. 405-441). Barcelona: Gedisa.
- Hair, J., Anderson, R., Tatham, R. & Black, W. (1999). *Análisis multivariante*. Madrid: Prentice Hall.
- Halliday, M.A.K. (1978). *Language as a social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M.A.K. (1989). *An Introduction to functional grammar*. London: Edward Arnold.

- Halliday, M.A.K. (1991). Corpus studies and probabilistic grammars. En K. Aijmer & Altenberg, B. (Eds.), *English corpus linguistics. Studies in honor of Jan Svartvik* (pp. 31-43). London: Longman.
- Halliday, M.A.K. (1992). Language as a system and language as an instance: The corpus as a theoretical construct. En J. Svartvik (Ed.), *Directions in corpus linguistics* (pp. 61-77). New York: W. de Gruyter.
- Halliday, M.A.K. (1993). On language and physical science. En M.A.K. Halliday & J. Martin (Eds.), *Writing Science: Literacy and discursive power* (pp. 54-68). London: Falmer.
- Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London: Logman.
- Halliday, M.A.K. & Martin, J.R. (1993). *Writing Science: Literacy and discursive power*. London: Falmer.
- Harman, D. (1992). Ranking algorithms. En W.B. Frakes & R. Baeza-Yates (Eds.), *Information retrieval: Data structures and algorithms* (pp.363-392). Englewood Cliffs, N.J.: Prentice Hall.
- Harris, Z. (1952). Discourse analysis. *Language*, 28, 1-30.
- Hartley, J. & Kosstoff, R. (2003). How useful are 'key words' in scientific journals? *Journal of information Science*, 29(5), 433-438.
- 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: Facultad de Filosofía y Humanidades. Universidad Austral de Chile.
- Harweg, R. (1968). *Pronomina und Textkonstitution*. München: Fink.
- Harweg, R. (1974). Deiktische und adeiktische Zeitstufen. *Zeitschrift für romanische Philologie*, 90, 499-525.
- Harweg, R. (1978). Substitutional text linguistics. En W. Dressler (Ed.), *Current trends in textlinguistics* (pp. 247-260). Berlin: W. de Gruyter.
- Heinemann, W. & Viehweger, D. (1991). *Textlinguistik. Eine Einführung*. Tübingen: Niemeyer.
- Hernández, R., Fernández, C. & Baptista, P. (1998). *Metodología de la investigación*: México D.F.: Mc Graw Hill.
- Hjelmslev, L. (1974). *Prolegómenos a una teoría del lenguaje*. Madrid: Gredos.

- Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. In *Cognitio*, 18, 27-34.
- Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: Benjamins.
- Hyland, K. (1999). Disciplinary discourses: Writer stance in research articles. En C. Candlin & K. Hyland (Eds.), *Writing texts, processes and practice* (pp. 99-121). London: Longman.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London: Longman.
- Iñíguez, L. (1993). De discursos, estructuras y análisis: ¿Qué prácticas?, ¿en qué contextos? [en línea]. Disponible en: <http://antalya.uab.es/liniguez/Materiales/003.pdf>
- Ivanic, R. (1998). *Writing and identity: The discorsal construction of identity in academic writing*. Amsterdam: Benjamins.
- Iza, M. & Esquerro, J. (1999). Representación conexionista y procesamiento del discurso. *Anales de Psicología*, 15(2), 303-318.
- Jackobson, R. & Halle, M. (1956). *Fundamentals of language*. The Hague: Mouton & Co.
- Jackson, P. & Moulinier, I. (2002). *Natural language processing for online applications. Text retrieval, extraction and categorization*. Amsterdam: Benjamins.
- Jeanneret, Y. (1994). *Écrire la science. Formes et enjeux de la vulgarisation*. Paris: PUF.
- Jurafsky, D. (2003). Probabilistic modelling in psycholinguistics: Linguistic comprehension and production. En R. Bod, J. Hay & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 38-95). London: MIT Press.
- Jurafsky, D. & Martin, J. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. New Jersey: Prentice Hall.
- Kanejiya, D., Kumar, A. & Prasad, S. (2003). Automatic evaluation of students' answers using syntactically enhanced LSA. Ponencia presentada en The Human Language Technology Conference (HLT-NAACL 2003) Workshop on Building Educational Applications using NLP. Edmonton, Canadá.
- Katz, J. & Fodor, J. (1976). *Estructura de una teoría semántica*. México D.F.: Siglo XXI.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension construction-integration model. *Psychological Review*, 95, 163-182.

- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Kintsch, W. (2000). Metaphor comprehension: A computational Theory. *Psychonomic Bulletin & Review*, 7-2, 257-266.
- Kintsch, W. (2001). Predication. *Cognitive Science*, 25(2), 173-202.
- Kintsch, W. (2002). On the notions of theme and topic in psychological process models of text comprehension. En M. Louwerse & W. van Peer (Eds.), *Thematics: Interdisciplinary Studies* (pp. 151-170). Amsterdam: Benjamins.
- Kintsch, E., Steinhart, D., Stahl, G., LSA Research Group, Matthews, C. & Lamb, R. (2000). Developing summarization skills through the use of LSA-based feedback. *Interactive learning environments*, 8(2), 87-109.
- Kintsch, W. & Bowles, A.R. (2002). Metaphor comprehension: What makes a metaphor difficult to understand? *Metaphor and Symbol*, 17, 249-262.
- Kintsch, W. & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Kolda, T. & O'Leary, D.A. (1998). Semidiscrete matrix decomposition for latent semantic indexing in information retrieval. *ACM Trans. Information Systems*, 16, 322-346.
- Lakoff, G. (1970). Global rules. *Language*, 46, 627-639.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: Chicago University Press.
- Landauer, T.K. (2002). On the computational basis of learning and cognition: Arguments from LSA. En N. Ross (Ed.), *The psychology of learning and motivation*, 41, 43-84.
- Landauer, T.K. & Psofka, J. (2000). Simulating text understanding for educational applications with latent semantic analysis: Introduction to LSA. *Interactive Learning Environments*, 8(2), 73-86.
- Landauer, T.K. & Dumais, S.T. (1996). How come you know so much? From practical problem to theory. En D. Hermann, C. McEvoy, M. Jonson & P. Hertel (Eds.), *Basic and applied memory: Memory in context* (pp. 105-126). Mahwah, N.J.: Erlbaum.
- Landauer, T.K. & Dumais, S.T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104, 211-240.

- Landauer, T.K., Laham, D., Rehder, B. & Schreiner, M.E. (1997). How well can passage meaning be derived without using word order? A comparison of latent semantic analysis and humans. En M.G. Shafto & P. Langley (Eds.), *Actas de The 19th annual meeting of the Cognitive Science Society* (pp. 412-417). Mahwah, N.J.: Erlbaum.
- Landauer, T.K., Foltz, P.W. & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25, 259-284.
- Lappin, S. (2003). Semantics. En R. Mitkov (Ed.), *The Oxford handbook of computational linguistics* (pp. 91-111). New York: Oxford University Press.
- Latorres, E. (2004). Similitud semántica: Comparación y crítica a los modelos actuales. En M. Solar, D. Fernández-Baca & E. Cuadros-Vargas (Eds.), *30ma. Conferencia Latinoamericana de Informática* (pp. 833-844). Arequipa: Sociedad Peruana de Computación.
- Lebart, L., Salem, A. & Bécue, M.. (2000). *Análisis estadístico de textos*. Lleida: Milenio.
- Lee, D. & Sung, H. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401, 788-791.
- Leech, G. (1977). *Semántica*. Madrid: Alianza.
- 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.
- Lemaire, B., Bianco, M., Sylvestre, E. & Noveck, I. (2001). Un modèle de compréhension de textes fondé sur l'analyse de la sémantique latente. En H. Paugam-Moisy, V. Nyckees & J. Caron-Pargue (Eds.), *La cognition entre individu et société* (pp. 309-320). Grenoble: Hermès.
- León, J.A. (2003) (Coord.). *Conocimiento y discurso: Claves para inferir y comprender*. Madrid: Pirámide.
- Leroy, M. (1969). *Las grandes corrientes de la lingüística*. Madrid: Gredos.
- Lin, D. & Pantel, P. (2001). Induction of semantic classes from natural language text. Artículo presentado en el SIGKDD-01, San Francisco, California, Estados Unidos.
- Lindsay, P. & Norman, D. (1972). *Human information processing. An introduction to psychology*. New York: Academic Press.
- Lo Cascio, V. (1998). *Gramática de la argumentación*. Madrid: Alianza.

- Locke, J. (1956). *Ensayo sobre entendimiento humano*. México D.F.: Fondo de Cultura Económica.
- López, C. (2002). Aproximaciones al análisis de los discursos profesionales. *Revista Signos*, 35(51-52), 195-215.
- López, J.A. (1997). El resumen como fuente de información y medio de comunicación. *RESUMED*, 10(3), 103-106.
- Louwerse, M. (2004). Un modelo conciso de cohesión en el texto y coherencia en la comprensión. *Revista Signos*, 37(56), 41-58.
- Louwerse, M., McCarthy, P., McNamara, D. & Graesser, A. (2004). Variation in language and cohesion across written and spoken registers. En K.D. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 843-848). Mahwah, N.J.: Erlbaum.
- Lyons, J. (1980). *Semántica*. Barcelona: Teide.
- Maineuneau, D. (1989). *Introducción a los métodos de análisis del discurso*. Buenos Aires: Hachette.
- Maldonado, M.F. (2002). *Hermes: Servidor y biblioteca de modelos de recuperación de información*. Tesis profesional, no publicada, Universidad de las Américas, Puebla, México.
- Manning, C. & Schütze, H. (2003). *Foundations of statistical natural language processing*. Cambridge: MIT Press.
- Marinkovich, J. & Cademártori, Y. (2004). Foco narrativo y foco informativo: Dos dimensiones para una descripción de los manuales en la formación técnico-profesional. *Revista Signos*, 37(55), 31-40.
- Marinkovich, J. & Morán, P. (1995). La calidad del texto argumentativo en estudiantes de 4º medio: Algunos rasgos significativos. *Estudios Filológicos*, 30, 78-89.
- Marinkovich, J. (1999). Un estudio exploratorio acerca de los tipos de textos escritos que se producen en el aula. En G. Parodi (Ed.), *Discurso, cognición y educación. Ensayos en honor a Luis Gómez Macker* (pp. 179-194). Valparaíso: EUVSA.
- Martin, J.R. (1997). Analysing genre: Functional parameters. En F. Christie & J. Martin (Eds.), *Genre and institutions* (pp. 3-39). London: Continuum.

- Martin J.R. & Rose, D. (2003). *Working with Discourse. Meaning beyond the clause*. London: Continuum.
- Martin, J.R. & Veel, R. (1998). *Reading Science. Critical and functional perspectives on discourses of science*. London: Routledge.
- Martín, P. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, 22, 25-43.
- Martinet, A. (1965). *Elementos de lingüística general*. Madrid: Gredos.
- Martinet, A. (1971). *El lenguaje desde el punto de vista funcional*. Madrid: Gredos.
- Martínez, J. (1997). La ciencia cognitiva, una investigación interdisciplinar. *Etapas de su construcción (I y II)*. *Contextos*, XV(29-30), 179-217.
- Maruhenda, Y. (2003). Estructura de un artículo científico [en línea]. Disponible en: <http://gidovi.umh.es/doctorado/internetInvestigacion/estructuraArticulo.html>.
- Matsumoto, Y. (2003). Lexical knowledge acquisition. En R. Mitkov (Ed.), *The Oxford handbook of computational linguistics* (pp. 396-409). New York: Oxford University Press.
- Mayor, J. (2000). Taxonomías de las relaciones semánticas en el marco de la actividad lingüística. En J. Beltrán, V. Bermejo, L. Pérez, M.D. Prieto, D. Vence, R. González (Coords.), *Intervención psicopedagógica y currículo escolar* (pp. 80-101). Madrid: Pirámide.
- McClelland, J.L., Rumelhart, D.E. & Hinton, G.E. (1992). El atractivo del procesamiento distribuido en paralelo. En D.E. Rumelhart, J. McClelland & el grupo P.D.P. (Eds.), *Introducción al procesamiento distribuido en paralelo* (pp. 39-80). Madrid: Alianza.
- McEnery, T & Wilson, A. (1996). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- McNamara, D. (2004). Aprender del texto: Efectos de la estructura textual y las estrategias del lector. *Revista Signos*, 37(55), 19-30.
- Miller, G.A. & Charles, W.G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6(1), 1-28.
- Mogollón, G. (2003). Paradigma científico y lenguaje especializado. *Revista de la facultad de Ingeniería de la Universidad Central de Venezuela*, 18(3), 5-14.
- Mora, J.P. (2001). Directed motion in English and Spanish. *Estudios de Lingüística Española*, 11 [en línea]. Disponible en: <http://elies.rediris.es/elies11/>

- Morán, P. (1999). Hacia el perfil de una competencia textual argumentativa: Una perspectiva pragma-retórica. En G. Parodi (Ed.), *Discurso, cognición y educación. Ensayos en honor a Luis Gómez Macker* (pp. 121-134). Valparaíso: EUVSA.
- Moreira, V. & Huyck, Ch. (2002). Portuguese-English experiments using latent semantic indexing [en línea]. Disponible en: <http://clef.iei.pi.cnr.it/workshop2002/WN/9.pdf>
- Moreno, A. (1998). *Lingüística computacional: Introducción a los modelos simbólicos, estadísticos y biológicos*. Madrid: Síntesis.
- Moyano, E. (2000). *Comunicar ciencia*. Buenos Aires: Secretaría de Investigaciones. Universidad Nacional de Lomas de Zamora.
- Narvaja de Arnoux, E., Di Stefano, M. & Pereira, C. (2000). *La lectura y la escritura en la universidad*. Buenos Aires: EUDEBA.
- Nida, E. (1975). *Exploring semantic structures*. München: Fink.
- Núñez, P. (1999). La construcción de textos argumentativos escritos en los inicios de la adolescencia de jóvenes chilenos. En G. Parodi (Ed.), *Discurso, cognición y educación. Ensayos en honor a Luis Gómez Macker* (pp. 159-178). Valparaíso: EUVSA.
- Nyrop, K. (1913). *Grammaire historique de la langue française*. Copenhague: Nordisk Forlag.
- Oakes, M. (1998). *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.
- OECD & Statistics Canada (2000). *Literacy in information age: Final report of the International Literacy Survey*. Paris: OECD Publication Service.
- OECD (2001). *Proyecto PISA (Programme for International Student Assessment)* [en línea]. Disponible en: <http://www.pisa.oecd.org/>
- Ogden, C.K. & Richards, I.A. (1984). *El significado del significado*. Barcelona: Paidós.
- Osgood, Ch. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49, 197-237.
- Osgood, Ch., Suci, G. & Tannenbaum, P. (1976). *La medida del significado*. Madrid: Gredos.
- Oyanedel, M. (2004). Programa de desarrollo de habilidades comunicativas. Ponencia presentada en el II Congreso Internacional sobre Lengua Escrita y Textos Académicos, Tlaxcala, México.
- Palmer, F. (1980). *La semántica*. Madrid: Siglo XXI.

- Paltridge, B. (1997). *Genres, frames, and writing in research settings*. Amsterdam: Benjamins.
- Parodi, G. (1998). Conexiones entre comprensión y producción de textos escritos: Estudio exploratorio en alumnos de educación básica. *Revista de Lingüística en el Aula*, 2(2), 7-18.
- Parodi, G. (1999). *Relaciones entre lectura y escritura: Una perspectiva cognitiva discursiva*. Valparaíso: EUVSA.
- Parodi, G. (2000). La evaluación de la producción de textos escritos argumentativos: Una alternativa cognitivo discursiva. *Revista Signos*, 33(47), 151-167.
- Parodi, G. (2001). Comprensión y producción del discurso escrito: Estudio empírico en escolares chilenos. *Revista Iberoamericana de Discurso y Sociedad*, 3(1), 75-101.
- Parodi, G. (2003). *Relaciones entre lectura y escritura: Una perspectiva cognitiva discursiva. Bases teóricas y antecedentes empíricos*. Valparaíso: EUVSA.
- Parodi, G. (2004). Textos de especialidad y comunidades discursivas técnico-profesionales: Una aproximación basada en corpus computarizado. *Revista Estudios Filológicos*, 39(39), 7-36.
- Parodi, G. (2005a). Lingüística de corpus y análisis multidimensional: Exploración de la variación en el corpus PUCV-2003: Una aproximación multiniveles. En G. Parodi (Ed.), *Discurso especializado e instituciones formadoras* (pp. 83-125). Valparaíso: EUVSA.
- Parodi, G. (2005b). Textos de especialidad y comunidades discursivas técnico-profesionales: Una aproximación a los manuales técnicos. En G. Parodi (Ed.), *Discurso especializado e instituciones formadoras* (pp. 61- 82). Valparaíso: EUVSA.
- Parodi, G. (2005c). *Comprensión de textos escritos*. Buenos Aires: EUDEBA.
- Parodi, G. & Núñez, P. (1998). Desarrollo de estrategias de lectura comprensiva: Una aplicación experimental del Programa L & C: Leer y Comprender. En M. Peronard, L. Gómez, G. Parodi & P. Núñez (Comps.), *Comprensión de textos escritos: De la teoría a la sala de clases* (pp. 249-264). Santiago: Andrés Bello.
- Parodi, G. & Núñez, P. (1999). El desarrollo de estrategias de lectura comprensiva: Una aplicación experimental. *Anales del Instituto de Lingüística*, XXI, 11-39.
- Parodi, G. & Gramajo, A. (2003). Los tipos textuales del corpus técnico-profesional PUCV 2003: Una aproximación multiniveles. *Revista Signos*, 36(54), 207-223.

- Parodi, G. & Venegas, R. (2004). BUCÓLICO: Aplicación computacional para el análisis de textos. Hacia un análisis de rasgos de la informatividad. *Revista de Lingüística y Literatura*, 15, 223-251.
- Pérez, C. (2002). Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento [en línea]. Disponible en: <http://elies.rediris.es/elies18/index.html>.
- Pérez, D., Gliozzo, A., Strapparava, C., Alfonseca, A., Rodríguez, P. & Magnini, B. (2005). Automatic assessment of student's free-text answers underpinned by the combination of a Bleu-inspired algorithm and latent semantic analysis. *American Association for Artificial Intelligence* [En prensa].
- Perfetti, Ch. (1998). The limits of co-occurrence: Tools and theories in language research. *Discourse Processes*, 25, 365-377.
- Peronard, M. (1989). Estrato social y estrategias de comprensión de lectura. *Lenguas Modernas*, 16, 19-32.
- Peronard, M. (1997). ¿Qué significa comprender un texto escrito? En M. Peronard, L. Gómez, G. Parodi & P. Núñez (Comps.), *Comprensión de textos escritos: De la teoría a la sala de clases* (pp. 55-78). Santiago: Andrés Bello.
- Peronard, M. (1998). La evaluación de la comprensión de textos escritos: El problema del resumen. En M. Peronard, L. Gómez, G. Parodi & P. Núñez (Comp.), *Comprensión de textos escritos: De la teoría a la sala de clases* (pp. 213-226). Santiago: Andrés Bello.
- Peronard, M. & Gómez, L. (1985). Reflexiones acerca de la comprensión lingüística: Hacia un modelo. *Revista de Lingüística Teórica y Aplicada*, 23, 19-32.
- Peronard, M., Gómez, L., Parodi, G. & Núñez, P. (1997). *Comprensión de textos escritos: De la teoría a la sala de clases*. Santiago: Andrés Bello.
- Petöfi, J. & Reiser, H. (1974). *Probleme der modeltheoretischen Interpretation von Texten*. Hamburg: Buske.
- Pinker, S. (1994). *The language instinct*. New York: Harper Collins.
- Pinker, S. (1997). *How the mind works*. New York: W. W. Norton.
- Platón (1981). *Diálogos*. Madrid: Gredos.

- Pomerantz, A. & Fehr, B. (1997). Conversation analysis: An approach to the study of social action as sense making practises. En T. van Dijk (Ed.), *Discourse as social interaction* (pp. 64-91). London: Sage.
- Pott, A. (1974). *Einleitung in die Allgemeine Sprachwissenschaft*. Amsterdam: Benjamins.
- Portier, B. (1974). *Linguistique générale. Théorie et description*. Paris: Klincksieck
- Pottier, B. (1983). *Semántica y lógica*. Madrid: Gredos.
- Pozo, J. (2001). *Humana mente. El mundo, la conciencia y la carne*. Madrid: Morata.
- Puente, A. (1998). *Cognición y aprendizaje. Fundamentos psicológicos*. Madrid: Pirámide.
- Pylyshyn, Z. (1989). Computing in cognitive science. En M. Posner (Ed.), *Foundations of cognitive science* (pp. 49-91). Cambridge: MIT Press.
- Quesada, J. (2003). *Latent problem solving analysis (LPSA): A computational theory of representation in complex, dynamic problem solving tasks*. Tesis doctoral no publicada, Universidad de Granada, España [en línea]. Disponible en: <http://www.andrew.cmu.edu/user/jquesada//dissertation/>
- Quesada, J.F., Kintsch, W. & Gómez, E. (2002). A theory of complex problem solving using latent semantic analysis. En W.D. Gray & C.D. Schunn (Eds.), *Actas de The 24th Annual Conference of the Cognitive Science Society* (pp. 750-755). Mahwah, N.J.: Erlbaum.
- Quine, W. (1968). *Palabra y objeto*. Barcelona: Labor.
- Ratteray, O. (1985). Expanding roles for summarized information. *Written Communication*, 2(4), 457-472.
- Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Sciences*, 4, 28-34.
- Rehder, B., Schreiner, M.E., Wolfe, M.B., Laham, D., Landauer, T.K., & Kintsch, W. (1998). Using latent semantic analysis to assess knowledge: Some technical considerations. *Discourse Processes*, 25, 337-354.
- Reisig, Ch. (1839). *Vorlesungen über lateinische Sprachwissenschaft*. Berlin: Calvary.
- Rodríguez, H. (2004). Lingüística y estadística, ¿incompatibles? En Martí, M.A. & Llisterri, J. (Eds.), *Tecnologías del texto y del habla* (pp. 89-117). Barcelona: Edicions Universitat de Barcelona.

- Rojo, G. (2002). Sobre la lingüística basada en análisis de corpus [en línea]. Disponible en: http://www.uzei.com/corpusajardunaldia/01_grojo.pdf
- Roweis, S. & Saul, L. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500), 2323-2326.
- Rumelhart, D.E. (1975). Notes on a schema for stories. En D.G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 185-210). New York: Academic Press.
- Rumelhart, D.E., McClelland, J. & el Grupo P.D.P. (1992). Introducción al procesamiento distribuido en paralelo. Madrid: Alianza.
- Rusell, B. (1973). *Essays in analysis*. London: George Allen & Unwin.
- Russell, B. (1937). *Principles of mathematics*. London: George Allen & Unwin.
- Sabaj, O. (2004). Especificidad, especialización y variabilidad verbal: Una aproximación computacional en estadística léxica. *Revista Signos*, 37(56), 75-89.
- Sager, J., Dungworth, D. & McDonald, P. (1980). *English special languages*. Wiesbaden: Oscar Brandstetter.
- Salager-Meyer, F. (1991). A text-type based discourse analysis of medical English. Abstracts internal structuring. Ponencia presentada en el Second Latin American ESP Colloquium, Santiago de Chile, Chile.
- Salager-Meyer, F. (1992). A text-type and move analysis study of verb tense and modality distribution in medical English abstracts. *ESP Journal*, 11(2), 93-113.
- Salton, G. & Buckley, C. (1988). Term-wighting approaches in automatic text retrieval. *Information Processing and Management*, 24(5), 513-523.
- Salton, G. & McGill, M. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.
- Salton, G. (1968). *Automatic information organization and retrieval*. New York: McGraw-Hill.
- Sánchez, M. (1993). *Los textos expositivos. Estrategias para mejorar su comprensión*. Madrid: Santillana.
- Sanders, T. & Spooren, W. (2001). Text representation as an interface between language and its users. En T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text representation. linguistic and psycholinguistics aspects* (pp. 1-25). Amsterdam: Benjamins.

- Saussure, F. (1983). *Curso de lingüística general*. Madrid: Alianza.
- Schaff, A. (1974). *Introducción a la Semántica*. México D.F.: Fondo de Cultura Económica.
- Schank, R. & Abelson, R. (1977). *Scripts, plans, goals and understanding*. New Jersey: Erlbaum.
- Schmidt, S. (1973). *Texttheorie*. München: Fink.
- Schmidt, S. (1976). *Texttheorie. Probleme einer Linguistik der sprachlichen Kommunikation*. München: Fink.
- Schröder, H. (1991). Linguistic and text-theoretical research on languages for special purposes. A thematic and bibliographical guide. En H. Schröder (Ed.), *Subject-oriented texts. Languages for special purposes and text theory* (pp. 1- 48). Berlin: W. de Gruyter.
- Searle, J. (1986). *Actos de habla: Ensayo de filosofía del lenguaje*. Madrid: Cátedra.
- Serafin, R., Di Eugenio, B. & Glass, M. (2003). *Latent semantic analysis for dialogue act classification*. Artículo presentado en HLT-NAACL03 2003, Human Language Technology Conference, Edmonton, Canada.
- Sigurd, B., Eeg-Olofsson, M. & van de Weijer, J. (2004). Word length, sentence length and frequency-Zipf revisited. *Studia linguistica*, 58(1), 37-52.
- SIMCE (2005a). *Evaluaciones internacionales* [en línea]. Disponible en: <http://www.simce.cl/paginas/evaluaciones.htm>
- SIMCE (2005b). *Resultados SIMCE* [en línea]. Disponible en: <http://www.simce.cl/>
- Sinclair, J.M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sitta, H. & Brinker, K. (1973). *Studien zur Texttheorie und zur deutschen Grammatik*. Düsseldorf: Schwann.
- Sperber, D. & Wilson, D. (1992). *La relevancia*. Madrid: Visor.
- Stempel, W. (1971). *Beiträge zur Textlinguistik*. München: Fink.
- Steyvers, M., Shiffrin, R.M. & Nelson, D.L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. En A. Healy (Ed.), *Experimental cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer*. Washington D.C.: American Psychological Association.
- Stubbs, M. (1996). *Text and corpus analysis. Computer-assisted studies of language and culture*. Malden, Massachusetts: Blackwell Publishers.

- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell Publishers.
- Swales, J. (1990). *Genre analysis. English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. (2004). *Research genres. Explorations and applications*. Cambridge: Cambridge University Press.
- Tapia, M., Burdiles, G. & Arancibia, B. (2003). Aplicación de una pauta diseñada para evaluar informes académicos universitarios. *Revista Signos*, 36(54), 249-257.
- Titscher, S., Meyer, M., Wodak, R. & Vetter, E. (2000). *Methods of text and discourse analysis*. London: Sage.
- Todorov, T. (1975). *Poética*. Buenos Aires: Losada.
- Tognini- Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: Benjamins.
- Trier, J. (1931). *Der deutsche Wortschatz im Sinnbezirk des Verstandes*. Heidelberg.
- Trosborg, A. (2000). *Analysing professional genres*. Amsterdam: Benjamins.
- Turney, P. (1999). *Learning to extract keyphrases from text*. Ottawa: National Research Council, Institute for Information Technology. Technical Report ERB-1057.
- Turney, P. (1997). *Extraction of keyphrases from text: Evaluation of four algorithms*. Ottawa: National Research Council Canada. Technical Report ERB-1051.
- Tzoukermann, E., Klavans, J. & Strzalkowski, T. (2003). Information retrieval. En R. Mitkov (Ed.), *The Oxford handbook of computational linguistics* (pp. 530-544). New York: Oxford University Press.
- Ullman, S. (1967). *Semántica. Introducción a la ciencia del significado*. Madrid: Aguilar.
- UNESCO-OREALC (1999). *Laboratorio latinoamericano de la calidad de la educación (LLECE)[en línea]*. Disponible en: <http://llece.unesco.cl/documentosdigitales/>
- Valiña, M. & Martín. M. (2000). *Psicología cognitiva. Perspectiva histórica, métodos y metapostulados*. Madrid: Pirámide.
- van den Broek, P. & Gustafson, M. (1999). Comprehension and memory for text. Three generation of reading research. En S. Goldman, A. Graesser & P. van den Broek (Eds), *Narrative comprehension, causality, and coherence* (pp. 15-34). Mahwah, N.J.: Erlbaum.
- van den Broek, P., Rinden, K., Fletcher, C. & Thurlow, R. (1996). A “Landscape” view of reading: Fluctuating patterns of activation and the construction of a stable memory

representation. En B. Britton & A. Graesser (Eds.), *Models of understanding text* (pp. 165-188). Mahwah, N.J.: Erlbaum.

van den Broek, P., Young, M., Tzeng, Y. & Linderholm, T. (1999) The landscape model of reading: Inferences and online construction of a memory representation. En H. van Oostendorp & S. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-89). Mahwah, N.J.: Erlbaum.

van Dijk, T. (1980). *Estructuras y funciones del discurso. Una introducción interdisciplinaria a la lingüística del texto y a los estudios del discurso*. México D.F.: Siglo XXI.

van Dijk, T. (1980). *Texto y contexto*. Madrid: Cátedra.

van Dijk, T. (1983). *La ciencia del texto. Un enfoque interdisciplinario*. Buenos Aires: Paidós.

van Dijk, T. (1995). *Prensa, racismo y poder*. México D.F.: Universidad Iberoamericana.

van Dijk, T. (1997). *Discourse studies: A multidisciplinary introduction*. London: Sage.

van Dijk, T. (1999). *Ideología. Una aproximación multidisciplinaria*. Barcelona: Gedisa.

van Dijk, T. (2003). *Racismo y discurso de las élites*. Barcelona: Gedisa.

van Dijk, T. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

van Overschelde, J.P. (2002). The influence of word frequency on recency effects in directed free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1160-1171.

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

Venegas, R. (2003). *La estructura argumentativa en textos producidos por estudiantes de educación superior*. Ponencia presentada en el II Congreso Internacional Cátedra UNESCO Lectura y Escritura. Pontificia Universidad Católica de Valparaíso, Viña del Mar, Chile.

Ventura, M., Hu, X., Graesser, A., Louwerse, M. & Olney, A. (2004). The context dependant sentence abstraction model. Artículo presentado en 26rd The Annual Meeting of the Cognitive Science Society.

Viehweger, D. (1976). *Semantische Merkmale und Textstruktur*. En F. Daneš & D. Viehweger (Eds.), *Probleme der Textgrammatik* (pp. 195-206). Berlin: Akademie-Verlag.

- Warta, V. (1996). Embedded case reports: A genre-analysis issue in teaching English for medical purposes. Tesis de Magíster no publicada, Aston University, Aston, Estados Unidos.
- Wiemer-Hasting, P. (2004). The FAQ for using LSA at Umemphis [en línea]. Disponible en: <http://www.msci.memphis.edu/~wiemerhp/trg/lisa-followup.html>
- Wiemer-Hastings, P. (2000). Adding syntactic information to LSA. Actas de The Twenty-second Annual Conference of the Cognitive Science Society (pp. 989-993). Mahwah, N.J.: Erlbaum.
- Wiemer-Hastings, P. & Zipitria, I. (2001). Rules for syntax, vectors for semantics. Actas de The Twenty-third Annual Conference of the Cognitive Science Society (pp. 1112-1117). Mahwah, N.J.: Erlbaum.
- Wierzbicka, A. (1992). Semantic, culture and cognition: Universal human concepts culture-specific configurations. New York: Oxford University Press.
- Wierzbicka, A. (1997). Understanding cultures through their key words: English, Russian, Polish, German, and Japanese. New York: Oxford University Press.
- Winston, M., Chaffin, R. & Hermann, D. (1987). A taxonomy of part-whole relations. *Cognitive Science*, 11, 417-444.
- Wittgenstein, L. (1988). Investigaciones filosóficas. México D.F.: Universidad Nacional Autónoma de México.
- Wodak, R. & Meyer. M. (2003). Métodos de análisis crítico del discurso. Barcelona: Gedisa.
- Wolfe, M.B., Schreiner, M.E., Rehder, B., Laham, D., Foltz, P.W., Kintsch, W. & Landauer, T.K. (1998). Learning from text: Matching readers and text by latent semantic analysis. *Discourse Processes*, 25, 309-336.
- Wundt, W. (1983). Über psychologischen Methoden. *Philosophischen Studien*, 1, 1-38.
- Yáñez, J. & Berndt, H. (2000). Nuevo sistema integrado cromatografía líquida de alta resolución/nebulizador hidráulico de alta presión y temperatura (HPLC/HT-HHPN) para la separación automática de la matriz en la determinación de metales en salmueras con espectrometría de absorción atómica (FAAS). *Boletín de la Sociedad Chilena de Química*, 45(4), 535-549.

Zazo, A., Figuerola, C., Alonso, J.L. & Gómez, R. (2002). Recuperación de información utilizando el modelo vectorial. Reporte Técnico DPTOIA-IT-2002-006, Departamento de Informática y Automática, Universidad de Salamanca [en línea]. Disponible en: <http://tejo.usal.es/inftec/2002/DPTOIA-IT-2002-006.pdf>.

ANEXO

1

“Codificación y dirección electrónica de las revistas usadas en la conformación del corpus”

CÓDIGO	REVISTA	DIRECCIÓN ELECTRÓNICA
	CIENCIAS EXACTAS	
BSCQ	BOLETÍN DE LA SOCIEDAD CHILENA DE QUIMICA	http://www.scielo.cl/scielo.php?pid=0366-1644&script=sci_serial
RCQ	REVISTA COLOMBIANA DE QUIMICA	http://www.icfes.gov.co/revistas/recolqui/
RPQ	REVISTA PERUANA DE QUÍMICA E INGENIERÍA QUÍMICA	http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Reglamento.htm
AAL	ACTAS DE LA ACADEMIA LUVENTICUS	http://www.luventicus.org/Actas/
ITERC	INTERCIENCIA	http://www.interciencia.org/
ACV	ACTA CIENTÍFICA VENEZOLANA	http://acta.ivic.ve/
	CIENCIAS BIOLÓGICAS	
GC	GAYANA CONCEPCIÓN	http://www.scielo.cl/scielo.php?pid=0717-6538&script=sci_serial

RCHA	REVISTA CHILENA DE ANATOMIA	http://www.scielo.cl/scielo.php?pid=0716-9868&script=sci_serial
RCI	REVISTA CHILENA DE INFECTOLOGIA	http://www.scielo.cl/scielo.php?pid=0716-1018&script=sci_serial
RCHN	REVISTA DE HISTORIA NATURAL	http://www.scielo.cl/scielo.php?pid=0716-078X&script=sci_serial
AMV	ARCHIVOS DE MEDICINA VETERINARIA	http://www.scielo.cl/scielo.php?pid=0301-732X&script=sci_serial
	CIENCIAS SOCIALES	
AMB	ÁMBITOS. REVISTA INTERNACIONAL DE COMUNICACIÓN	http://www.ull.es/publicaciones/latina/ambitos/ambitos.htm
CHU	CHUNGARÁ.	http://www.scielo.cl/scielo.php?pid=0717-7356&script=sci_serial
AD	ANALES DE DOCUMENTACIÓN	http://www.um.es/fccd/anales/
NS	NUEVA SOCIEDAD	http://www.nuevasoc.org.ve/home/
INE	INVESTIGACIONES ECONÓMICAS	http://www.funep.es/invecon/sp/sie.asp

ANEXO

2

“Codificación de los artículos y referencias bibliográficas del corpus ARTICO”

CIENCIAS EXACTAS

ID	Código	Referencia Bibliográfica
1	BSCQ1-1999 44(3)	Garzón, G., Galarza, E., Hoyos, O. & Fackler, J. (1999). Adición oxidativa de sustratos tipo adición oxidativa de sustratos tipo PHZ-ZPH (z = s, se, te) al complejo heterotrinuclear AU ₂ PT[CH ₂ P(S)PH ₂] 4. <i>Boletín de la Sociedad Chilena de Química</i> , 44(3), 255-261.
2	BSCQ2-1999 44(3)	Acosta, G.A., Rodríguez, Y., Plutin, A., Márquez, H., Dago, A., Pomes, R. & Esteva, A. (1999). Estructura del bromohidrato de 2,4-diamino-5- feniltiazol (amifenazol) en disolución y en estado sólido <i>Boletín de la Sociedad Chilena de Química</i> , 44 (3), 263-269.
3	BSCQ3-1999 44(3)	Ponzi, E., Masini, O., Comelli, N., Grzona, L., Carruscull, A., & Ponzi, M. (1999). Isomerización de α -PINENO. Influencia de la concentración de hierro en óxido de circonio sulfatado. <i>Boletín de la Sociedad Chilena de Química</i> , 44(3), 271-277.
4	BSCQ4-1999 44(3)	Taboada, M., Graber, T. & Bastias, E. (1999). Un nuevo método de obtención de datos cinéticos para cristalización batch desde soluciones acuosas. <i>Boletín de la Sociedad Chilena de Química</i> , 44(3), 279-287.
5	BSCQ5-1999 44(3)	Amador-Hernández, J., López-Martínez, L., López-de-Alba, P.L., Caldera, A., Estela, J.M. & Cerda, V. (1999). Determinación simultánea de seis hidrocarburos policíclicos aromáticos en medio micelar por regresión de mínimos cuadrados parciales (pls-1) utilizando espectros de fluorescencia de ángulo variable lineal. <i>Boletín de la Sociedad Chilena de Química</i> , 44(3), 299-313.
6	BSCQ6-1999 44(3)	Alvez, G., García, R., Arriagada, R. & Cid, R (1999). Arcillas pilareadas II. Hidrodesulfuración de tiofeno sobre catalizadores Ni, W soportados sobre arcillas pilareadas. <i>Boletín de la Sociedad Chilena de Química</i> , 44(3), 337-344.

7	BSCQ7-1999 44(3)	Báez, M. & Zincker, J. (1999). Parámetros de calidad analítica de un método de determinación multiresiduos de plaguicidas por HPLC-DAD. <i>Boletín de la Sociedad Chilena de Química</i> , 44(3), 357-366.
8	BSCQ8-1999 44(3)	Bartulin, J., Ramos, M. & Zunza, H. (1999). Propiedades mesomórficas de 1-(4-n-alcoxifenil)-3-(5-n-alquil-2-tienil)-1,3-propanodionas y de sus derivados isoxazólicos y pirazólicos. <i>Boletín de la Sociedad Chilena de Química</i> , 44(3), 375-384.
9	BSCQ9-1999 44(4)	Neira, J., Bruhn, C., Steger, E. & Reyes, N. (1999). Nebulización hidráulica de alta presión como medio de transporte e introducción de muestra en la determinación directa de trazas de cobre, hierro y manganeso. Aplicación en licores de blanqueo de celulosa. <i>Boletín de la Sociedad Chilena de Química</i> , 44(4), 427-434.
10	BSCQ10-1999 44(4)	García, R., Cid, R. & Arriagada, R. (1999). Retención de Cr(III) y Hg(II) en zeolitas. Influencia de la naturaleza de la zeolita y de variables de proceso. <i>Boletín de la Sociedad Chilena de Química</i> , 44(4), 435-442.
11	BSCQ11-1999 44(4)	López-de-Alba, P., López-Martínez, L. & Amador-Hernández, J. (1999). Determinación espectrofotométrica de cobre en formulaciones médicas, salvado de trigo y aguas potables. <i>Boletín de la Sociedad Chilena de Química</i> , 44(4), 469-477.
12	BSCQ12-1999 44(4)	Mendoza, S., Jaramillo, V. & Vega, M. (1999). Determinación de metamizol-sódico en supositorios por cromatografía en capa fina de alta eficiencia (HPTLC). <i>Boletín de la Sociedad Chilena de Química</i> , 44(4), 479-485.
13	BSCQ13-1999 44(4)	Vásquez, S. & Flint, Y C. (1999). Estudio de los procesos de transferencia de energía entre iones Pr ³⁺ dopados en Cs ₂ NaYCl ₆ . <i>Boletín de la Sociedad Chilena de Química</i> , 44(4), 487-496.
14	BSCQ14-1999 44(4)	Guevara, J. L., Rojas, R., Narvaez, A. & Quijada, R. (1999). Estudio del comportamiento catalítico de la homo y copolimerización de 1-octadeceno con

		catalizadores metallocenos racémicos rac-Et(Ind) ₂ ZrCl ₂ y rac-Me ₂ Si(Ind) ₂ ZrCl ₂ . <i>Boletín de la Sociedad Chilena de Química</i> , 44(4), 497-504.
15	BSCQ15-1999 44(4)	Catalán, R., Neira, A., Melo, R. & Rivas, B. (1999). Copolimerización por injerto de acetato de vinilo en fibra celulósica secundaria. <i>Boletín de la Sociedad Chilena de Química</i> , 44(4), 505-515.
16	BSCQ1-2000 45(1)	Cordero De Troconis, M., Pedrique De Aulacio, M. & Colman De S., T. (2000). Modelado molecular y relaciones estructura- actividad antibacteriana de quinolinas análogas al ácido nalidixico. <i>Boletín de la Sociedad Chilena de Química</i> , 45(1), 5-13.
17	BSCQ2-2000 45(1)	Montes, S., Valero, E. & Schmidt, R. (2000). Estudio de los mecanismos de flotabilidad del cuarzo en presencia de sales de alquilamonio. I. Efecto del pH. <i>Boletín de la Sociedad Chilena de Química</i> , 45(1), 31-39.
18	BSCQ3-2000 45(1)	Antilen, M., Escudey, M. & Förster, J. (2000). Efecto del catión homoionizante sobre la competencia de adsorción de fosfato-ftalato-cloruro en un andisol chileno. <i>Boletín de la Sociedad Chilena de Química</i> , 45(1), 41-48.
19	BSCQ4-2000 45(1)	Machuca, M., De Gregori, I. & Pinochet, H. (2000). Determinación amperométrica indirecta de Se(IV) con preconcentración en línea por análisis de inyección en flujo. <i>Boletín de la Sociedad Chilena de Química</i> , 45(1), 99-107.
20	BSCQ5-2000 45(1)	Becerra, A., Dimitrijewits, M., Arciprete, C. & Castro-Luna, A. (2000). Desarrollo de un catalizador estable de níquel sobre alúmina para el reformado de metano con dióxido de carbono. <i>Boletín de la Sociedad Chilena de Química</i> , 45(1), 147-154.

21	BSCQ1-2000 45(2)	Díaz, F., Sánchez, C., Del Valle, M., Bernède, J. & Tregouët, Y. (2000). Síntesis, caracterización y propiedades eléctricas de poli (p-fenilsulfonil-anilina). <i>Boletín de la Sociedad Chilena de Química</i> , 45(2), 181-189.
22	BSCQ2-2000 45(2)	Guzmán, M., Ortega, P., & Vera, L. (2000). Método polinomial general para calcular parámetros de corrosión de datos obtenidos por técnica potencioestática o intencioestática. <i>Boletín de la Sociedad Chilena de Química</i> , 45(2), 191-197.
23	BSCQ3-2000 45(2)	Bailac, P., Dellacasa, A., Bernasconi, H., Firpo, N. & Ponzi, M. (2000). Composición del aceite esencial y actividad antimicrobiana de eupatorium patens. <i>Boletín de la Sociedad Chilena de Química</i> , 45(2), 207-211.
24	BSCQ4-2000 45(2)	Yazdani-Pedram, M., Calderón, K. & Quijada, R. (2000). Funcionalización de polipropileno en solución mediante injerto de monometil itaconato. <i>Boletín de la Sociedad Chilena de Química</i> , 45(2), 269-282.
25	BSCQ5-2000 45(2)	Valencia, E., Valenzuela, E., Barros, E., Hernández, M., Lazo, C., Gutiérrez, C., Gonzalez-Coloma, A., González, A. & Bermejo, J. (2000). Estudio fitoquímico y actividad antialimentaria de senna stipulaceae. <i>Boletín de la Sociedad Chilena de Química</i> , 45(2), 297-301.
26	BSCQ6-2000 45(2)	Bunster, M., Contreras-Martel, C., Bruna, C. & Martínez- Oyanede, J. (2000). R-Ficoeritrina de gracilaria chilensis: estabilidad e interacción entre subunidades. <i>Boletín de la Sociedad Chilena de Química</i> , 45(2), 303-309.
27	BSCQ1-2000 45(3)	Alvarado, Y., Cubillán, N., Soscun, H., Soscun, H., Osorio, E. & Vallejo, R. (2000). Estudio teórico de la polarizabilidad molecular en triafulveno y compuestos análogos. <i>Boletín de la Sociedad Chilena de Química</i> , 45(3), 339-346.
28	BSCQ2-2000 45(3)	Costamagna, J., Lillo, L.E., Matsuhira, B. & Villagrán, M. (2000) Poli-D-galactosamina como ligante en la formación de complejos con Cu(II) y Ni(II). <i>Boletín de la Sociedad Chilena de Química</i> , 45(3), 377-383.

29	BSCQ3-2000 45(3)	Catalán, R., Neira, A., Rivas, B. & Melo, R. (2000). Polimerización de injerto de fibra celulósica secundaria con metacrilato de metilo y metacrilato de etilo. <i>Boletín de la Sociedad Chilena de Química</i> , 45(3), 391-402.
30	BSCQ4-2000 45(3)	Lisperguer, J., Ballerini, A., Nuñez, M. & Palavecino, P. (2000). Análisis térmico de adhesivos fenólicos modificados con lignina. <i>Boletín de la Sociedad Chilena de Química</i> , 45(3), 403-408.
31	BSCQ5-2000 45(3)	Orellana, F., Ahumada, E., Suárez, C., Cote, G. & Lizama, H. (2000) Estudio termodinámico de parámetros involucrados en la formación de los precipitados de Arsénico (V) con Bario(II). <i>Boletín de la Sociedad Chilena de Química</i> , 45(3), 415-422.
32	BSCQ6-2000 45(3)	Ortiz, J., Barbato, S. & Gautier, J.L. (2000). Reducción del oxígeno sobre calcosita, pirita y galena en medio alcalino. <i>Boletín de la Sociedad Chilena de Química</i> , 45(3), 441-448.
33	BSCQ1-2000 45(4)	Yáñez, J. & Berndt, H. (2000). Nuevo sistema integrado cromatografía líquida de alta resolución / nebulizador hidráulico de alta presión y temperatura (HPLC/HT-HHPN) para la separación automática de la matriz en la determinación de metales en salmueras con espectrometría de absorción atómica (FAAS). <i>Boletín de la Sociedad Chilena de Química</i> , 45(4), 535-549.
34	BSCQ2-2000 45(4)	Maurelia, R., Zamora, R., Guevara, M., Rojas, M. & Rojas, L. (2000). Caracterización y separación de una estilbita originaria de la región de Atacama, Chile. <i>Boletín de la Sociedad Chilena de Química</i> , 45(4), 569-575.
35	BSCQ3-2000 45(4)	Ríos-Escudero, A., Costamagna, J. & Villagrán, M. (2000). Síntesis y caracterización de nuevos complejos de Ni(II) con bases de schiff derivadas de n-Octilamina, n-Dodecilamina y n-Octadecilamina. <i>Boletín de la Sociedad Chilena de Química</i> , 45(4), 593-599.
36	BSCQ4-2000 45(4)	Morales, G., Sierra, P., Bórquez, J. & Loyola, L. (2000). Rigidusido, un nuevo glicoditerpenoide de haplopappus rigidus. <i>Boletín de la Sociedad Chilena de</i>

		<i>Química</i> , 45(4), 611-614.
37	BSCQ5-2000 45(4)	Mennickent, S., Vega, M., Godoy, C.G. & Yates, T. (2000). Desarrollo de un método por cromatografía en capa fina instrumental para análisis cuantitativo de ácido acetilsalicílico. <i>Boletín de la Sociedad Chilena de Química</i> , 45(4), 615-620.
38	BSCQ6-2000 45(4)	Burgos, F., Arancibia, V., Manríquez, J.M. & Chávez, I. (2000). Síntesis y caracterización [Cp*Ru(C8H7)], un intermediario útil para la construcción de oligómeros homo y hetero polinucleares de rutenio. <i>Boletín de la Sociedad Chilena de Química</i> , 45(4), 621-628.
39	BSCQ1-2001 46(1)	Toral, M. I., Lara, N., Gómez, J. & Richter, P. (2001). Determinación de hierro en fase sólida por espectrofotometría derivada de segundo orden. <i>Boletín de la Sociedad Chilena de Química</i> , 46(1), 51-60.
40	BSCQ2-2001 46(1)	Dimitrijewits, M. I., Arciprete, C., Becerra, A. & Castro-Luna, A. (2001). Reformado de metano con Co ₂ sobre catalizadores de níquel: efectos de la técnica de preparación. <i>Boletín de la Sociedad Chilena de Química</i> , 46(1), 65-75.
41	BSCQ1-2001 46(2)	Capobianco, M. & Cordero de Troconis, M.I. (2001). Diseño racional de compuestos antibacterianos derivados del 5-nitrofurano usando modelado molecular. <i>Boletín de la Sociedad Chilena de Química</i> , 46(2), 113-119.
42	BSCQ2-2001 46(2)	Valenzuela, F., Pizarro, F., Basualto, C., Sapag, J. & Tapia, C. (2001). Extracción de Mo (VI) con aminas terciarias en un reactor de membranas líquidas de soporte sólido. <i>Boletín de la Sociedad Chilena de Química</i> , 46(2), 187-195.
43	BSCQ3-2001 46(2)	Barral, R., Pozo, K., Urrutia, R., Cisternas, M., Pacheco, P. & Focardi, S. (2001). Plaguicidas organoclorados persistentes en sedimentos de tres lagos costeros y un lago andino de Chile central. <i>Boletín de la Sociedad Chilena de</i>

		<i>Química</i> , 46(2), 149-159.
44	BSCQ4-2001 46(2)	Contreras, P., Faini, F. & Rojas, M.C. (2001). Obtención y separación de giberelinas lactónicas producidas por el hongo gibberella fujikuroi para usarlas como estándares cromatográficos. <i>Boletín de la Sociedad Chilena de Química</i> , 46(2), 197-202.
45	BSCQ5-2001 46(2)	Basáez, L., Peric, I., Aguirre, C. & Vanysek, P. (2001). Estudio electroquímico del antibiótico amoxicilina a través de interfases líquido-líquido. <i>Boletín de la Sociedad Chilena de Química</i> , 46(2), 203-208.
46	BSCQ1-2001 46(3)	Pizarro, J., Rubio, M.A. & Lira, G. (2001). Optimización de una técnica de muestreo in situ: análisis de Cu y Fe en sistemas acuáticos. <i>Boletín de la Sociedad Chilena de Química</i> , 46(3), 281-285.
47	BSCQ2-2001 46(3)	Belmar, J., Jiménez, C. & Rodríguez, J. (2001). Una síntesis eficiente del n,n',n"-tris- (2,3 dihidroxibenzoil)-1,1,1-tris-(aminometil)ano, un análogo de la enterobactina. <i>Boletín de la Sociedad Chilena de Química</i> , 46(3), 301-307.
48	BSCQ3-2001 46(3)	Rubio, M. A., Lissi, E., Riveros, V. & Páez, M. (2001). Remoción de contaminantes por lluvias y rocíos en la región Metropolitana. <i>Boletín de la Sociedad Chilena de Química</i> , 46(3), 353-361.
49	BSCQ4-2001 46(3)	Álvez, G., García, R., Cid, R., Escalona, N. & Gil-Llambias, F.J. (2001). Hidrodesulfuración de tiofeno sobre catalizadores Ni-W y Ni-Re. Efecto del soporte. <i>Boletín de la Sociedad Chilena de Química</i> , 46(3), 363-372.
50	BSCQ1-2001 46(4)	Baeza, P., Ojeda, J., Escalona, N., Alvarez, G., García, R., Cid, R. & Gil-Llambía, F.J. (2001). Estudio de la sinergia del sistema Ni-Re sobre la actividad catalítica en la reacción de hidrodesulfuración de un gasoil. <i>Boletín de la Sociedad Chilena de Química</i> , 46(4), 415-418.
51	BSCQ2-2001 46(4)	Bittner, M., Alarcón, J., Aqueveque, P., Becerra, J., Hernández, V., Hoeneisen, M. & Silva, M. (2001). Estudio químico de especies de la familia

		euphorbiaceae en Chile. <i>Boletín de la Sociedad Chilena de Química</i> , 46(4), 419-431.
52	BSCQ3-2001 46(4)	Pastene, E., Wilkomirsky, T., Bocaz, G., Havel, J., Peric, I., Vega, M., González, M. & Alderete, J. (2001). Uso de espectroscopía de RMN y Maldi-Tof Ms en la elucidación estructural de flavonoides antioxidantes provenientes de la planta medicinal chilena <i>Cheilanthes glauca</i> (Cav.) Mett. <i>Boletín de la Sociedad Chilena de Química</i> , 46(4), 449-457.
53	BSCQ4-2001 46(4)	Matsuhiro, B. & Rubio, M.J. (2001). Modificación química de pectina de limón. Chemical modification of lemon pectin. <i>Boletín de la Sociedad Chilena de Química</i> , 46(4), 481-486.
54	BSCQ5-2001 46(4)	Herrera, C., Tello, M., Ibáñez, I., Velásquez, L., Mancini, M., & Álvarez, S. (2001). Digestión proteica con microondas y su aplicación en preparación de muestras para análisis de harina de pescado. <i>Boletín de la Sociedad Chilena de Química</i> , 46(4), 487-494.
55	BSCQ1-2002 47(1)	Cid, R., Bendezú, S. & López A. (2002). Actividad catalítica de sulfuros de Ni y W soportados sobre zeolita ultraestable. <i>Boletín de la Sociedad Chilena de Química</i> , 47(1), 13-18.
56	BSCQ2-2002 47(1)	Modak, B., Arrieta, A., Torres, R. & Urzúa, A. (2002). Actividad antibacteriana de flavonoides aislados del exudado resinoso de <i>heliotropium sinuatum</i> : efecto del tipo de estructura. <i>Boletín de la Sociedad Chilena de Química</i> , 47(1), 19-23.
57	BSCQ3-2002 47(1)	Lisperguer, J. & Droguett, C. (2002). Caracterización del curado de resinas urea formaldehído por calorimetría diferencial de barrido (dsc). <i>Boletín de la Sociedad Chilena de Química</i> , 47(1), 33-38.
58	BSCQ4-2002 47(1)	Catalán, R., Neira, A. & Rivas, B. (2002). Estudio de la copolimerización por injerto de acrilonitrilo en fibra celulósica secundaria. <i>Boletín de la Sociedad</i>

		<i>Chilena de Química, 47(1), 39-47.</i>
59	BSCQ5-2002 47(1)	Cornwell, E. (2002). Optimización del exponente implícito en el algoritmo correspondiente al índice de conectividad de valencia tridimensional corregido $\lambda(q)$. Aplicación a hidrocarburos saturados. <i>Boletín de la Sociedad Chilena de Química, 47(1), 53-60.</i>
60	BSCQ6-2002 47(1)	Figuroa, C., Vega, M., Saelzer, R., Ríos, G. & Rodríguez, H. (2002). Análisis de la fracción lipídica de la yema de huevos por amd-hptlc. <i>Boletín de la Sociedad Chilena de Química, 47(1), 61-66.</i>
61	BSCQ1-2002 47(2)	Guevara, J.L., Quijada, R., Saavedra, P., Palza, H. & Galland, G. (2002). Estudio de la polimerización de propeno utilizando diferentes catalizadores metallocenos. <i>Boletín de la Sociedad Chilena de Química, 47(2), 81-90.</i>
62	BSCQ2-2002 47(2)	Palma, G., Alvear, M. & Salazar, I. (2002). Utilización de desechos celulósicos en la preparación de formulaciones de liberación controlada de los herbicidas simazina y trifluralina. <i>Boletín de la Sociedad Chilena de Química, 47(2), 175-180.</i>
63	BSCQ3-2002 47(2)	Collao, S., Arce, E. & Andía, A. (2002). Mineralogía, química e inclusiones fluidas en los depósitos de nitratos de María Elena, IIa. región, Chile. <i>Boletín de la Sociedad Chilena de Química, 47(2), 181-190.</i>
64	BSCQ1-2002 47(3)	Toral, M.I., Lara, N., Tapia, A., Torres, C. & Richter, Pablo. (2002). Estudio espectral y determinación simultánea de sulfametoxazol y trimetoprim por espectrofotometría derivada digital. <i>Boletín de la Sociedad Chilena de Química, 47(3), 241-251.</i>
65	BSCQ2-2002 47(3)	Torres, R., Modak, B., Urzúa, A., Delle Monache, F., Damonte, E. & Pujol, C. (2002). Propiedades antivirales de compuestos naturales y semi-sintéticos de la resina de <i>heliotropium filifolium</i> . <i>Boletín de la Sociedad Chilena de Química, 47(3), 259-263.</i>

66	BSCQ1-2002 47(4)	Bolaños, A., Cuenú, F. & Vargas, R. (2002). Caracterización espectroscópica y análisis térmico de los complejos trifenilfosfina-rutenio(II) conteniendo el ligando 2,6-dimetilpirazina y otros n-heterocíclicos. <i>Boletín de la Sociedad Chilena de Química</i> , 47(3), 299-310.
67	BSCQ2-2002 47(4)	Sienra, M., Préndez, M. & Romero, R. (2002). Metodología para la extracción, fraccionamiento y cuantificación de hidrocarburos aromáticos policíclicos presentes en material particulado urbano. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 311-324.
68	BSCQ3-2002 47(4)	Espinoza, L. & Cortés, M. (2002) Síntesis y actividad biológica de dos nuevos análogos de brasinoesteroides funcionalizados en el anillo C. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 335-347.
69	BSCQ4-2002 47(4)	Escudey, M., Förster, J., Galindo, G. & Pizarro, C. (2002). Uso de la aproximación de Rothmund-Kornfeld para la descripción del intercambio catiónico binario en aluminosilicatos sintéticos de carga variable. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 349-359.
70	BSCQ5-2002 47(4)	Belmar, J., Zúñiga, C., Jiménez, C. & Saavedra, J. (2002). Síntesis de N,N'-Bis-3,5-DI-TER-Butilsalicilamidas. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 371-377.
71	BSCQ6-2002 47(4)	Moya, S., Azócar-Guzmán, I., Guerrero, J. & Aguirre, P. (2002). Síntesis y caracterización de complejos mono y bimetálico de Renio(I) y Rutenio (II) con 4,4'-bipiridina como ligando puente. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 393-401.
72	BSCQ7-2002 47(4)	Rudolph, A., Franco, C., Becerra, J., Barros, A. & Ahumada, R. (2002). Evaluación de materia orgánica e hidrocarburos aromáticos policíclicos en sedimentos superficiales, bahía Concepcion, Chile. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 403-410.

73	BSCQ8-2002 47(4)	Montes, S., Valero, E. & Montes, G. (2002). Estudio de los mecanismos de flotabilidad del cuarzo en presencia de sales de alquilamonio: II. Efecto del almidón. . <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 449-456.
74	BSCQ9-2002 47(4)	Urrutia, R., Yevenes, M. & Barra, R. (2002). Determinación de los niveles basales de metales traza, en sedimentos de tres lagos andinos de Chile: lagos Chungará, Laja y Castor. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 457-467.
75	BSCQ10-2002 47(4)	Tapia, J., Freer, J., Mansilla, H., Villaseñor, J., Bruhn, C. & Basualto, S. (2002). Estudio de reducción fotocatalizada de cromo hexavalente. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 469-476.
76	BSCQ11-2002 47(4)	Gutiérrez, M., Becerra, J. & Barra, R. (2002). Tribromofenol empleado en aserraderos: métodos de análisis, características físico-químicas y presencia en componentes ambientales. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 485-493.
77	BSCQ12-2002 47(4)	Nieto, I. & Valencia, M. (2002). Esteroles, ácidos grasos e hidrocarburos de los cuerpos fructíferos de ganoderma australe. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 511-516. Disponible en: http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0366-16442002000400023&lng=es&nrm=iso&tlng=es
78	BSCQ13-2002 47(4)	Espinoza, L. & Cortes, M. (2002). Síntesis y actividad biológica de análogos de brasinoesteroides. <i>Boletín de la Sociedad Chilena de Química</i> , 47(4), 511-516. Disponible en: http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0366-16442002000400024&lng=es&nrm=iso&tlng=es
79	RCQ1-1997 26(1)	Pinzón, J. (1997). Superficie específica de una bentonita mediante la adsorción de azul de metileno. <i>Revista Colombiana de Química</i> , 26(1), 1-14.
80	RCQ2-1997 26(1)	Kouznetzov, V. & Palma A. (1997). Nuevos aminoalcoholes y amidas de la

		serie γ -piperidínica. <i>Revista Colombiana de Química</i> , 26(1), 15-22.
81	RCQ3-1997 26(1)	Lozano, A., Guzmán, G., Ferrucho, A & Rojas, L. (1997). Isotermas de adsorción de Potasio en un oxisol de Llanos Orientales de Colombia. <i>Revista Colombiana de Química</i> , 26(1), 23-30.
82	RCQ4-1997 26(1)	Gartner, D. & Restrepo, P. (1997). Determinación de vitamina a por cromatografía líquida de alta eficiencia (clae) en bienestarina cruda. <i>Revista Colombiana de Química</i> , 26(1), 31-42.
83	RCQ5-1997 26(1)	Sáez, A., Blair, S. & Sáez, J. (1997). Aporfinoides en corteza de guatterialehmanii y evaluación de su actividad antimalárica invitro. <i>Revista Colombiana de Química</i> , 26(1), 43-50.
84	RCQ6-1997 26(1)	Luque, J. E. (1997). Comparación de las lectinas extraídas de la semilla y la raíz de arveja (pisum sativum). <i>Revista Colombiana de Química</i> , 26(1), 50-55.
85	RCQ7-1997 26(2)	Tellez, A., Álvarez, G. & Roa, A. (1997). Diseño y puesta en marcha de un "sistema semicontinuo en dos etapas: hidrólisis-fermentación" para la producción de etanol a partir de almidón de papa usando simultáneamente aspergillus niger saccharomyces cerevisiae. <i>Revista Colombiana de Química</i> , 26(2),1-10.
86	RCQ8-1997 26(2)	Melgarejo, M., Ballesteros M.I. & Bendeck, M. (1997). Evaluación de algunos parámetros fisicoquímicos y nutricionales en humus de lombriz y composts derivados de diferentes sustratos. <i>Revista Colombiana de Química</i> , 26(2), 11-20.
87	RCQ9-1997 26(2)	Alarcón, E., Lozano, A. & Chaparro, H. (1997). Caracterización fenotípica de aislamientos rizobianos de Acacia (acacia sp.) y Retamo (teline monpessulana). <i>Revista Colombiana de Química</i> , 26(2), 21-34.
88	RCQ10-1997 26(2)	Echavarría, A. & Saldarriaga, C. (1997). Nuevo método para la síntesis de

		fosfatos de zinc microporosos. <i>Revista Colombiana de Química</i> , 26(2), 51- 55.
89	RCQ1-1998 27(1)	Barragán, D., Giraldo, L., Agreda J., & Moreno, J. (1998). Estructura general de un software para captura y procesamiento de datos. Aplicaciones a potenciometría y microcalorimetría de alta precisión. <i>Revista Colombiana de Química</i> , 27(1), 1- 12.
90	RCQ2-1998 27(1)	Delgado, W., Avella, E., & de Díaz, A. (1998). Alcaloides bencilisoquinolínicos del tallo de piper augustum rudge. <i>Revista Colombiana de Química</i> , 27(1), 13- 22.
91	RCQ3-1998 27(1)	Estupiñán, A., Sarmiento, D. & Belalcázar de Galvis, A. (1998). Remoción de Cobre y Níquel por intercambio catiónico con una Zeolita natural. <i>Revista Colombiana de Química</i> , 27(1), 31- 40.
92	RCQ4-1999 27(1)	Pacheco, L., López, C., Uribe, C. & Peláez, C. (1998). Limitaciones del valor de ΔG para la diferenciación del ambiente químico molecular. <i>Revista Colombiana de Química</i> , 27(1), 41-50.
93	RCQ5-1998 27(1)	Quintero, C., Segura, E. & Delgado, E. (1998). Caracterización química y mineralógica de las cromitas de la zona de Bello y San Pedro (departamento de Antioquia). <i>Revista Colombiana de Química</i> , 27(1), 51- 60.
94	RCQ6-1998 27(1)	Gómez, M.A. & Vargas, W. (1998). Desarrollo de un catalizador sólido de carácter ácido. <i>Revista Colombiana de Química</i> , 27(1), 61- 76.
95	RCQ7-1998 27(1)	Sáenz, J., Granados, H., Escobar, G., Cardona, W., Atehortua, L., Callejas, R., Cortes, D. & Gonzalez, C. (1998). Piperlonguminina y estigmasterol, compuestos de raíces y tallos de Piper auritum, actividad insecticida de extractos. <i>Revista Colombiana de Química</i> , 27(1), 77- 82.
96	RCQ8-1998 27(1)	Echavarría, A., Moreno, M., Ramírez, C., Tamayo, C. & Saldarriaga, C. (1998). Uso de bentonita en aguas residuales de curtiembres para la remoción

		de Cr(III). <i>Revista Colombiana de Química</i> , 27(1), 83- 88.
97	RCQ9-1998 27(1)	Gil, E. & Saldarriaga, C. (1998). Uso de ceniza volante en la remoción de Cromo (III) de los efluentes líquidos de curtiembres. <i>Revista Colombiana de Química</i> , 27(1), 89- 95.
98	RCQ1-1999 28(1)	Calvo, J., Barrera, N., García, J., Guzman, F., Espejo, F. & Patarroyo, M. (1999). Síntesis de oxitocina en fase sólida usando derivados de terbutoxicarbonilo y fluorenilmtoxycarbonilo. <i>Revista Colombiana de Química</i> , 28(1), 19- 26.
99	RCQ2-1999 28(1)	Calvo, J., Rivera, Z., Carvajal, C., Bernal, A., Chaves, F., Lozano, J., Trujillo, M., Guzmán, F. & Patarroyo, M. (1999). Síntesis de péptidos dendriméricos usando la estrategia de doble dimero. <i>Revista Colombiana de Química</i> , 28(1), 27- 34.
100	RCQ3-1999 28(1)	García, E., Bustamante, F. & Montes de Correa, C. (1999). Reducción de óxidos de nitrógeno con catalizadores zeolíticos intercambiados con cobalto. <i>Revista Colombiana de Química</i> , 28(1), 35- 44.
101	RCQ4-1999 28(1)	Quiroz, N., Stashenko, E., Páez, E. & Martínez, J. (1999). Zeolitas Nay intercambiadas con metales de transición (Fe ²⁺ , Co ²⁺ , Mo ²⁺ y Mn ²⁺) como catalizadores para la oxidación de limoneno. <i>Revista Colombiana de Química</i> , 28(1), 45- 54.
102	RCQ5-1999 28(1)	Quijano, C., Suárez, M. & Duque, C. (1999). Constituyentes volátiles de dos variedades de guayaba (psidium guajava l.) : Palmira Ica-1 y Glum Sali. <i>Revista Colombiana de Química</i> , 28(1), 55- 64.
103	RCQ6-1999 28(1)	Murillo, E. & Viña, A. (1999). Determinación de constituyentes volátiles de la albahaca (ocimum spp) mediante dos métodos de extracción. <i>Revista Colombiana de Química</i> , 28(1), 65- 74.

104	RCQ7-1999 28(1)	Farías, D., Ballesteros, M.I. & Bendeck, M. (1999). Variación de parámetros fisicoquímicos durante un proceso de compostaje. <i>Revista Colombiana de Química</i> , 28(1), 75- 86.
105	RCQ8-1999 28(1)	Mendoza, D., Peña, J. & Franco, A. (1999). Efecto del glifosato y paraquat sobre el proceso de nitrificación en un suelo del corregimiento de Río Frío (Magdalena, Colombia). <i>Revista Colombiana de Química</i> , 28(1), 87- 96.
106	RCQ9-1999 28(1)	Yu. Sakharov, I., Ardila, G., Sakharova, I., Rojas, A. & Yu. Pletjuschkina, O. (1999). Peroxidasa de plantas tropicales. <i>Revista Colombiana de Química</i> , 28(1), 97- 106.
107	RCQ10-1999 28(1)	Lee, C. & Navarro, Y. (1999). Rhizobium leguminosarum como organismo biocontrolador de la interacción hospedero-patógeno: Clavel (dianthus caryophyllus)- fusarium oxysporum f.sp.dianthi <i>Revista Colombiana de Química</i> , 28(1), 107- 115.
108	RPQ1-2000 3(1)	Erazo, R. (2000). Diseño del relleno sanitario «La Vizcacha» del distrito de Puente Piedra, Lima-Perú. <i>Revista Peruana de Química e Ingeniería Química</i> , 3(1), [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Vol3_Nº1/diseño_relleno.htm
109	RPQ2-2000 3(1)	Pizarro, R. & Salas, G. (2000). Diseño y construcción de un sistema de flotación por aire disuelto, de efluentes líquidos, a nivel de laboratorio. <i>Revista Peruana de Química e Ingeniería Química</i> , 3(1), [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Vol3_Nº1/diseño.htm
110	RPQ3-2000 3(1)	Erazo, R. & Cárdenas, J.L. (2000). Planta de tratamiento de aguas residuales en la Universidad Nacional Mayor de San Marcos, Lima-Perú. <i>Revista Peruana de Química e Ingeniería Química</i> , 3(1), [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Vol3_Nº1/pla

		nta.htm
111	RPQ4-2000 3(1)	Garrido, A. & Otiniano, M. (2000). Síntesis, caracterización y actividad catalítica de los catalizadores Fe-Mo-Sb. <i>Revista Peruana de Química e Ingeniería Química</i> , 3(1), [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Vol3_Nº1/sintesis.htm
112	RPQ5-2000 3(1)	Gamarra, G. & Woolcott, J.C. (2000). Optimización de las condiciones fermentativas para la producción y extracción de b--galactosidasa de <i>Kluyveromyces marxianus</i> [En línea]. <i>Revista Peruana de Química e Ingeniería Química</i> , 3(1). Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Vol3_Nº1/optimizacion.htm
113	RPQ6-2000 3(1)	Angulo, J. & Sotelo, A. (2000). Determinación de la constante de estabilidad del complejo Bis(1,1-dietil-3-benzoiltioureato) Plomo(II). <i>Revista Peruana de Química e Ingeniería Química</i> , 3(1), [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Vol3_Nº1/determinacion.htm
114	AAL1-2002 1	Luetich, J.J. (2002). La ecuación fundamental de los sistemas multifásicos en equilibrio químico en el marco de la aproximación física. <i>Actas de la Academia Luventicus</i> , 1, 1- 22.
115	AAL2-2002 1	Luetich, J.J. (2002). Combinaciones lineales, multiplicadores de Lagrange y transformadas de Legendre. Una mirada a tres procedimientos para escribir el potencial de Gibbs de una mezcla reactiva. <i>Actas de la Academia Luventicus</i> , 1, 23-36.
116	AAL3-2002 1	Luetich, J.J. (2002). Tres teoremas fundamentales de la termodinámica química. <i>Actas de la Academia Luventicus</i> , 1, 37-50.

117	AAL4-2002 1	Luetich, J.J. (2002). La reversibilidad microscópica y la información contenida en el vector composición. <i>Actas de la Academia Luventicus</i> , 1, 51-68.
118	AAL1-2003 1	Luetich, J.J. (2003). La transformada de Legendre completa del potencial de Gibbs del sistema (ácido sulfhídrico+agua+acetato de sodio) respecto de las variables de composición de los elementos. <i>Actas de la Academia Luventicus</i> , 1, 1-13.
119	AAL2-2003 3	Luetich, J.J. (2003). La termodinámica y los mecanismos de las reacciones químicas. <i>Actas de la Academia Luventicus</i> , 3, 1-4.
120	ITERC1-2001 26(8)	Romero, R., Adame, R. & Sen, M. (2001). Estudio de los parámetros que afectan la transferencia de calor conjugada en intercambiador de calor de tubos y placas-aleta. <i>Revista Interciencia</i> , 26(8), 321- 326.
121	ITERC1-2003 28(1)	Sanhueza, E., Alfonso, Y. & Santana, M. (2003). Compuestos orgánicos volátiles en la atmósfera de La Gran Sabana. II: HCHO, HCOOH y C[H.sub.3]COOH en lluvias. <i>Revista Interciencia</i> , 28(1), 29- 35.
122	ITERC2-2003 28(3)	Ribón, M., Salgado, S., Palma, D. & Lagunas, L. del C. (2003). Propiedades químicas y físicas de un vertisol cultivado con caña de azúcar. <i>Revista Interciencia</i> , 28(3), 154-159.
123	ITERC3-2003 28(8)	Alzate, L., Pérez, A., Nevárez, V., Rinderknecht, N. & Poggi, H. (2003). Comparación de dos sistemas anaerobios acoplados para la biometanización de la fracción orgánica de residuos sólidos urbanos. <i>Revista Interciencia</i> , 28(8), 436- 442.
124	ITERC4-2003 28(8)	Del Valle, J. (2003). Cantidad, calidad y nutrientes reciclados por la hojarasca fina en bosques pantanosos del Pacífico sur colombiano. <i>Revista Interciencia</i> , 28(8), 443-449.
125	ITERC5-2003 28(3)	Del Valle, J. (2003). Descomposición de la hojarasca fina en bosques

		pantanosos del Pacífico colombiano. <i>Revista Interciencia</i> , 28(3), 148- 153
126	AVC1-1999 50(1)	Carbognani, L., Méndez, A. & Sánchez, V. (1999). Caracterización de crudos venezolanos. Orígenes, razones y necesidades. <i>Actas Venezolanas de Ciencias</i> , 50(1), 11- 17
127	AVC2-1999 50(1)	Pereira, P., Machín, I., Salerno, G., Cotte, E., Higuerey, I., Andriollo, A., Cordova, J., Zacarias, L., Marzín, R. & Rivas, G. (1999). La investigación y desarrollo en mejoramiento de crudos extrapesados en pdvsa-intevp: Aquaconversion(r). <i>Actas Venezolanas de Ciencias</i> , 50(1), 48- 53.
128	AVC3-1999 50(1)	Rivas, H. & Gutiérrez, X. (1999). Los surfactantes: comportamiento y algunas de sus aplicaciones en la industria petrolera. <i>Actas Venezolanas de Ciencias</i> , 50(1), 54-65.
129	AVC4-1999 50(1)	Martínez, M.A., Núñez, A., López, R., Morales, F. & Núñez, O. (1999). Degradación de fenoles usando TiO ₂ y luz UV, seguimiento de la reacción En un amplio rango de pH. <i>Actas Venezolanas de Ciencias</i> , 50(1), 81- 86.
130	AVC5-1999 50	Cabrera, G. & Márquez, C. (1999). Síntesis de 4-aminoquinolinas promovida por ácidos de Lewis. <i>Actas Venezolanas de Ciencias</i> , 50, 173-176.
131	AVC1-2000 51	Millán, E., Ramírez, M. & Perdomo, G. (2000). Síntesis de poliésteres-uretano a partir de 1,4-ciclohexanodiol y metileno-bis-4-fenilisocianato. <i>Actas Venezolanas de Ciencias</i> , 51, 150-159.
132	AVC2-2000 51	Quintero, T., Pernía, E., Guzmán, W. & Márquez, C. (2000). Síntesis de 3-carboetoxi-2-fenil-4-quinolonas. Potenciales antimicrobianos. <i>Actas Venezolanas de Ciencias</i> , 51, 137-142.

CIENCIAS BIOLÓGICAS

ID	Código	Referencia Bibliográfica
1	AMV1-2000 32(1)	Sierra, N., Ramírez, R. & Mota, D. (2000). Aislamiento del virus de PRRS en México: Estudio clínico, serológico y virológico. <i>Archivos de Medicina Veterinaria</i> , 32(1), 1-9.
2	AMV2-2000 32(1)	Islas, A., Quezada, M., Bernal, A., Mora, G., López-Rivero, J.L., Merino, V., Dossow, P., Rojas, V. & Marín, L. (2000). Características histoquímicas, morfométricas y metabólicas del músculo Gluteus medius de equinos entrenados para competencias de salto. <i>Archivos de Medicina Veterinaria</i> , 32(1), 11-19.
3	AMV3-2000 32(1)	Sucre, L., Finol, H., & Pérez, R. (2000). Cambios ultraestructurales del músculo Gluteus medius en caballos mestizos de tiro no entrenados y sometidos a trabajo de labranza. <i>Archivos de Medicina Veterinaria</i> , 32(1), 21-32.
4	AMV4-2000 32(1)	Aller, J.F., Alberio, R.H., Palma, G. (2000). Gestación con embriones producidos in vitro a partir de ovocitos recuperados de vacas ovariectomizadas. <i>Archivos de Medicina Veterinaria</i> , 32(1), 33-39.
5	AMV5-2000 32(1)	Silva, M., Gatica, R. & Correa, J. (2000). Inducción de mellizos mediante la transferencia de un segundo embrión ipsilateral o contralateral al cuerpo lúteo en vacas cubiertas. <i>Archivos de Medicina Veterinaria</i> , 32(1), 41-48.
6	AMV6-2000 32(1)	López-Ortega, A., Márquez, Y. & Mendoza, C. (2000). Variación de la concentración sanguínea del colesterol total y de las lipoproteínas en conejos hembras mantenidas a baja temperatura. <i>Archivos de Medicina Veterinaria</i> , 32(1), 49-54.
7	AMV7-2000 32(1)	Oblitas, F., Contreras, P., Böhmwald, H. & Wittwer, F. (2000). Efecto de la suplementación con selenio sobre la actividad sanguínea de glutatión peroxidasa (GSH-Px) y ganancia de peso en bovinos selenio deficientes.

		mantenidos a pastoreo. <i>Archivos de Medicina Veterinaria</i> , 32(1), 55-62.
8	AMV8-2000 32(1)	Tadich, N. & Hernández, M. (2000). Prevalencia de lesiones podales en ovinos de 25 explotaciones familiares de la provincia de Valdivia, Chile. <i>Archivos de Medicina Veterinaria</i> , 32(1), 63-74.
9	AMV9-2000 32(1)	Zamora, J., Reinhardt, G., Polette, M., Macías, P. & González, I. (2000). Escherichia coli aislada de cerditos diarreicos. Presunción de cepas productoras del Factor Citotóxico Necrosante (CNF). <i>Archivos de Medicina Veterinaria</i> , 32(1), 75-81.
10	AMV10-2000 32(2)	Recabarren, S., Lobos, A., Schneider, C., Cox, J. & Parilo, J. (2000). Sensibilidad tisular a la insulina antes, durante y después de un ayuno en ovejas prepúberes. <i>Archivos de Medicina Veterinaria</i> , 32(2), 139-146.
11	AMV11-2000 32(2)	Recabarren, S., Lobos, A., Robinson, C., Orellana, P. & Parilo, J. (2000). Efecto de melatonina sobre la secreción pulsátil de hormona luteinizante y de hormona del crecimiento en borregas con restricción alimenticia. <i>Archivos de Medicina Veterinaria</i> , 32(2), 147-156.
12	AMV12-2000 32(2)	Gallo, C., Pérez, S., Sanhueza, C. & Gasic, J. (2000). Efectos del tiempo de transporte de novillos previo al faenamiento sobre el comportamiento, las pérdidas de peso y algunas características de la canal. <i>Archivos de Medicina Veterinaria</i> , 32(2), 157-170.
13	AMV13-2000 32(2)	Tadich, N., Gallo, C. & Alvarado, M. (2000). Efectos de 36 horas de transporte terrestre con y sin descanso sobre algunas variables sanguíneas indicadoras de estrés en bovinos. <i>Archivos de Medicina Veterinaria</i> , 32(2), 171-183.
14	AMV14-2000 32(2)	Otero, P., Jacomet, L., Pisera, D., Rebuelto, M. & Hallu, R. (2000). Estudio preliminar de los efectos cardio-respiratorios del fentanilo en caninos anestesiados con halotano. <i>Archivos de Medicina Veterinaria</i> , 32(2), 185-192.

15	AMV15-2000 32(2)	Baroni, E., Ahumada, F., Lüders, C., Saavedra, I. & Zambucetti, R. (2000). Farmacocinética de propofol en potrillos de 10 a 15 días de edad. <i>Archivos de Medicina Veterinaria</i> , 32(2), 193-200.
16	AMV16-2000 32(2)	López, A., Morales, M., Cabrera, R., Urra, X. (2000). Ingestión y digestibilidad aparente de forrajes por la llama (<i>Lama glama</i>): I.- Heno de alfalfa (<i>Medicago sativa</i>) y paja de trigo (<i>Triticum Aestivum</i>) en diferentes proporciones. <i>Archivos de Medicina Veterinaria</i> , 32(2), 201-208.
17	AMV17-2000 32(2)	Patitucci, A., M. Phil, R.S.A., Pérez, M., Israel, K. & Rozas, M. (2000). Prevalencia de anticuerpos séricos contra <i>Neospora caninum</i> en dos rebaños lecheros de la IX Región de Chile. <i>Archivos de Medicina Veterinaria</i> , 32(2), 209-214.
18	AMV18-2000 32(2)	Cubillos, V., Suarez, D., Enriquez, R., Alberdi, A. & Monras, M. (2000). Caracterización clínica y aspectos lesionales macro y microscópicos de la leucemia linfoblástica en salmones y truchas de cultivo. <i>Archivos de Medicina Veterinaria</i> , 32(2), 215-227.
19	AMV19-2000 32(2)	Lightowler, C., Pidal, G., Mercado, M. & Cattaneo, M. (2000). Evaluación ecocardiográfica de la función sistólica en el caballo. Parte 1. Valores de referencia para el porcentaje de acortamiento fraccional y fracción de eyección. <i>Archivos de Medicina Veterinaria</i> , 32(2), 229-234.
20	AMV1-2001 33(1)	Gesche, E., Madrid, E. & Águila, C. (2001). Efecto del pH, cepa bacteriana y tipo de muestra, en la detección microbiológica, de ácido oxolínico y oxitetraciclina en peces. <i>Archivos de Medicina Veterinaria</i> , 33(1), 21-29.
21	AMV2-2001 33(1)	Díaz, D., Picco, E., Encinas, T. Rubio, M. & Litterio, N. (2001). Residuos tisulares de nicotinato de norfloxacin administrado por vía oral en cerdos. <i>Archivos de Medicina Veterinaria</i> , 33(1), 37-42.
22	AMV3-2001 33(1)	Gallo, C., Espinoza, M. & Gasic, J. (2001). Efectos del transporte por camión durante 36 horas con y sin período de descanso sobre el peso vivo y algunos

		aspectos de calidad de carne en bovinos. <i>Archivos de Medicina Veterinaria</i> , 33(1), 43-53.
23	AMV4-2001 33(1)	Islas, A., Quezada, M., Mora, G., López-Rivero, J.L., Merino, V., Rojas, H., Carrillo, C., Rojas, M. & Zegpig, A. (2001). Caracterización de las fibras del músculo gluteus medius en equinos por métodos histoquímicos e inmunohistoquímicos. <i>Archivos de Medicina Veterinaria</i> , 33(1), 55-62.
24	AMV5-2001 33(1)	Moine, R., Rivera, M., Vivas, A., Ferraris, G., Galán, A. & Natali, J. (2001). Morfometría y determinación de calcio y fósforo en la parte media de la diáfisis del metacarpiano III en yeguas mestiza con criollo. <i>Archivos de Medicina Veterinaria</i> , 33(1), 63-68.
25	AMV6-2001 33(1)	Rubilar, L., Donoso, S., Díaz, L., Godoy, C., Muñoz, L. & Pérez, R. (2001). Eficacia antihelmíntica de tres endectocidas administrados por vía oral en caballos. <i>Archivos de Medicina Veterinaria</i> , 33(1), 69-75.
26	AMV7-2001 33(1)	Pérez, R., Cabezas, I., Godoy, C., Rubilar, L., Díaz, L., Muñoz, L., Arboix, M. & Alvinerie, M. (2001). Disposición plasmática y fecal de moxidectina administrada por vía oral en caballos. <i>Archivos de Medicina Veterinaria</i> , 33(1), 77-88.
27	AMV8-2001 33(1)	Sepúlveda, N., Risopatrón, J., Oberg, J. & Neumann, A. (2001). Suplementación pre y post parto en ovejas. Efecto sobre la pubertad y actividad reproductiva de sus hijas. <i>Archivos de Medicina Veterinaria</i> , 33(1), 89-96.
28	AMV9-2001 33(2)	Pulido, R., Balocchi, O. & Fernandez, J. (2001). Efecto del nivel de producción de leche sobre el comportamiento ingestivo en vacas lecheras en pastoreo primaveral. <i>Archivos de Medicina Veterinaria</i> , 33(2), 137-144.
29	AMV10-2001 33(2)	López, A., Morales, M., Cabrera, R. & Arias, M. (2001). Ingestión y digestibilidad aparente de forrajes por la llama (<i>Lama glama</i>). II. heno de trébol rosado (<i>Trifolium pratense</i>), heno de ballica (<i>Lolium multiflorum</i>), paja

		de poroto (<i>Phaseolus vulgaris</i>) y paja de avena (<i>Avena sativa</i>). <i>Archivos de Medicina Veterinaria</i> , 33(2), 145-152.
30	AMV11-2001 33(2)	Pokniak, J., Cornejo, S., Bravo, I. & Battaglia, J. (2001). Pigmentación de trucha arco iris (<i>oncorhynchus mykiss</i>) tipo mar, alimentadas con dos niveles de astaxantina en dietas de crecimiento-engorda. <i>Archivos de Medicina Veterinaria</i> , 32(2), 155-164.
31	AMV12-2001 33(2)	Celedón, M., Sandoval, A., Droguett, J., Calfio, R., Ascencio, L. Pizarro, J. & Navarro, C. (2001). Pesquisa de anticuerpos seroneutralizantes para pestivirus y herpesvirus en ovinos, caprinos y camélidos sudamericanos de Chile. <i>Archivos de Medicina Veterinaria</i> , 33(2), 165-172.
32	AMV13-2001 33(2)	Reinhardt, G., Carrasco, L., Tadich, N. & Riedemann, S. (2001). Comparación entre dos técnicas de diagnóstico para diarrea viral bovina (DVB) en 50 predios de la X región, Chile. Seroneutralization y enzimoimmunoensayo indirecto (ELISA-I). <i>Archivos de Medicina Veterinaria</i> , 33(2), 173-183.
33	AMV14-2001 33(2)	Ronchi, J., Estela, E., Leunda, M., & Odeón, C. (2001). Infección experimental con el virus de la diarrea viral bovina (VDVB) genotipo 2 en terneros con anticuerpos neutralizantes al VDVB genotipo 1. <i>Archivos de Medicina Veterinaria</i> , 33(2), 185-192.
34	AMV15-2001 33(2)	López, I., Cubillos, V., Ernst, S. & Rosenfeld, C. (2001). Frecuencia y distribución, según edad y estacionalidad, de patologías hepáticas en salmonídeos, provenientes de la Xª a XIIª Región de Chile. Período 1988-1998. <i>Archivos de Medicina Veterinaria</i> , 33(2), 193-202.
35	AMV16-2001 33(2)	Cubillos, V., López, I. & Alberdi, A. (2001). Aspectos lesionales microscópicos en hígados de salmonídeos, provenientes de centros de cultivo de la Xª a XIIª Región de Chile (1988-1998). <i>Archivos de Medicina Veterinaria</i> , 33(2), 203-216.

36	AMV17-2001 33(2)	Cavestany, D., Galina, C. & Viñoles, C. (2001). Efecto de las características del reinicio de la actividad ovárica posparto en la eficiencia reproductiva de vacas Holstein en pastoreo. <i>Archivos de Medicina Veterinaria</i> , 33(2), 217-226.
37	AMV1-2002 34(1)	Kelly, L., Postiglioni, A., De Andrés, D., Gagliardi, R., Biagetti, R. & Franco J. (2002). Variabilidad genética de los caballos criollos del Uruguay. <i>Archivos de Medicina Veterinaria</i> , 34(1), 13-23.
38	AMV2-2002 34(1)	Thibaut, J., Rivera, T. & Ahumada, F. (2002). Anestesia endovenosa en perros mediante el uso de propofol en dosis única, premedicado con acepromazina-atropina y xilazina-atropina. <i>Archivos de Medicina Veterinaria</i> , 34(1), 25-35.
39	AMV3-2002 34(1)	Sievers, G., Jara, M., Cárdenas, C. & Núñez, J. (2002). Estudio anual de la eliminación de huevos y ooquistes de parásitos gastrointestinales y larvas de nemátodos pulmonares en ovinos de una estancia en Magallanes, Chile. <i>Archivos de Medicina Veterinaria</i> , 34(1), 37-47.
40	AMV4-2002 34(1)	Wittwer, F., Araneda, P., Ceballos, A., Contreras, P., Andaur, M. & Böhmwald, H. (2002). Actividad de glutatión peroxidasa (GSH-Px) en sangre de bovinos a pastoreo de la IX Región, Chile y su relación con la concentración de selenio en el forraje. <i>Archivos de Medicina Veterinaria</i> , 34(1), 49-57.
41	AMV5-2002 34(1)	Llorente, P., Leoni, L. & Martínez Vivot, M. (2002). Leptospirosis en camélidos sudamericanos. Estudio de prevalencia serológica en distintas regiones de la Argentina. <i>Archivos de Medicina Veterinaria</i> , 34(1), 59-68.
42	AMV6-2002 34(1)	Giraldo de León, G., Orrego Uribe, A. & Betancurth, A. (2002). Los roedores como reservorios de Leptospiras en planteles porcinos de la zona central cafetera de Colombia. <i>Archivos de Medicina Veterinaria</i> , 34(1), 69-78.
43	AMV7-2002 34(1)	Giraldo de León, G., Orrego Uribe, A., Santacruz, M. & Yepes, E. (2002). Leptospirosis. Las aguas de la explotación porcina como vehículo de la Leptospira, en la zona central cafetera de Colombia. <i>Archivos de Medicina</i>

		<i>Veterinaria</i> , 34(1), 79-87.
44	AMV8-2002 34(1)	Lottersberger, J. Pauli, R. & Vanasco, N. (2002). Desarrollo y validación de un enzimoimmunoensayo para el diagnóstico de Leptospirosis bovina. <i>Archivos de Medicina Veterinaria</i> , 34(1), 89-95.
45	AMV9-2002 34(2)	Zúñiga, C., Vargas, R. & Vergara, U. (2002). Evolución de la infección con <i>Trypanosoma cruzi</i> en cepas susceptibles y resistentes de ratones. <i>Archivos de Medicina Veterinaria</i> , 34(2), 183-188.
46	AMV10-2002 34(2)	Anrique, R. & Viveros, M. P. (2002). Efecto del ensilado sobre la composición química y degradabilidad ruminal de la pomasa de manzana. <i>Archivos de Medicina Veterinaria</i> , 34(2), 189-197.
47	AMV11-2002 34(2)	Martínez, E., Pulido, R. & Latrille, L. (2002). Efecto de la paja de trigo tratada con alcali sobre el consumo de alimento y comportamiento ingestivo de vacas lecheras. <i>Archivos de Medicina Veterinaria</i> , 34(2), 199-212.
48	AMV12-2002 34(2)	Ramírez, M., Ernst, S., Elvinger, F., Rivera, A. & Rosenfeld, C. (2002). Respuesta serológica y tiempo de saneamiento en rebaños bovinos con brucelosis vacunados con Cepa 19 o Cepa RB-51; Xª Región, Chile. <i>Archivos de Medicina Veterinaria</i> , 34(2), 213-220.
49	AMV13-2002 34(2)	San Martín, B., Kruze, J., Morales, M., Agüero, H., Leon, B., Espinosa, S., Iragüen, D., Puga, J. & Borie, C. (2002). Resistencia bacteriana en cepas patógenas aisladas de mastitis en vacas lecheras de la V Región, Región Metropolitana y Xª Región, Chile. <i>Archivos de Medicina Veterinaria</i> , 34(2), 221-234.
50	AMV14-2002 34(2)	Díaz, G., Galina, C., Basurto, C. & Ochoa, G. (2002). Efecto de la progesterona natural con o sin la adición de benzoato de estradiol sobre la presentación de celo, ovulación y gestación en animales tipo <i>Bos indicus</i> en el trópico mexicano. <i>Archivos de Medicina Veterinaria</i> , 34(2), 283-286.

51	AMV15-2002 34(2)	Perfumo, C., Sanguinetti, H., Giorgio, N., Armocida, A., Machuca, M., Massone, A., Risso, M., Aguirre, J. & Idiart, J. (2002). Constrictura rectal en cerdos necropsiados en una granja de ciclo completo en confinamiento. Consideraciones sobre su prevalencia, hallazgos anatomopatológicos y etiopatogenia. <i>Archivos de Medicina Veterinaria</i> , 34(2), 245-252.
52	AMV16-2002 34(2)	Soto, J.P., Kruze, J. & Leiva, S. (2002). Comparación de tres métodos de diagnóstico de Paratuberculosis bovina en rebaños lecheros infectados. <i>Archivos de Medicina Veterinaria</i> , 34(2), 253-263.
53	AMV1-2003 35(1)	Anrique, R. & Dossow, C. (2003). Efectos de la pulpa de manzana ensilada en la ración de vacas lecheras sobre el consumo, la tasa de sustitución y la producción de leche. <i>Archivos de Medicina Veterinaria</i> , 35(1), 13-22.
54	AMV2-2003 35(1)	Lertora, J., Paredes, E., Reinhardt, G. & Alberdi, A. (2003). Inmunohistoquímica en biopsias de piel tratadas con proteína K y microondas para el diagnóstico en animales persistentemente infectados con el virus diarreico viral bovina. <i>Archivos de Medicina Veterinaria</i> , 35(1), 23-36.
55	AMV3-2003 35(1)	Redondo, E., Gómez, L., Kelling, C., Gázquez, A., Masot, A. J. (2003). Hibridación in situ del virus respiratorio syncytial bovino en pulmón de cordero a diferentes tiempos postinfección. <i>Archivos de Medicina Veterinaria</i> , 35(1), 37-49.
56	AMV4-2003 35(1)	Adarmes, H., Riveros, A., Galleguillos, M. & González, E. (2003). Contenido de glicosaminoglicanos del líquido sinovial de la articulación metacarpofalángica de caballos castrados y yeguas de diferentes edades. <i>Archivos de Medicina Veterinaria</i> , 35(1), 51-59.
57	AMV5-2003 35(1)	Gallardo, C., Paredes, E. & Pérez, J. (2003). Estudio histopatológico de hígado y riñón de caninos y su relación con las concentraciones de urea, creatinina, proteínas, enzimas (ALT y SAP) en sangre premortem y en humor acuoso a las 0 y 24 horas postmortem. <i>Archivos de Medicina Veterinaria</i> , 35(1), 61-74.

58	GC1-2000 64(2)	Chiappa, E., Rodríguez, S. & Bascuñán, R. (2000). Preferencia y manipulación floral de <i>Centris rodophthalma</i> Perez (hymenoptera: anthophoridae). <i>Gayana (Concepc.)</i> , 64(2), 133-138.
59	GC2-2000 64(2)	Iannacone, J., Alvariño, L., Caballero, C. & Sánchez, J. (2000). Cuatro ensayos ecotoxicológicos para evaluar Lindano y Clorpirifos. <i>Gayana (Concepc.)</i> , 64(2), 139-146.
60	GC3-2000 64(2)	Louis-Maldonado, M. & Badilla, R. (2000). Descripción de los estados preimaginales de <i>Hemieuxoa conchidia</i> (Butler). <i>Gayana (Concepc.)</i> , 64(2), 147-153.
61	GC4-2000 64(2)	Rodríguez, M., Olivares, T. & Angulo, A. (2000). Nuevo género y nuevas especies de Noctuidos altoandinos v. ternera Rodríguez & Angulo n. Gen. v/s géneros afines (lepidoptera: noctuidae: cucullinae). <i>Gayana (Concepc.)</i> , 64(2), 155-160.
62	GC5-2000 64(2)	Romero, F., Altieri, E., Quiñehual, C. & Cayuqueo, A. (2000). Actividad contráctil del músculo papilar cardíaco y conducto deferente de rata inducida por veneno de la araña <i>latrodectus mactans</i> de Chile. <i>Gayana (Concepc.)</i> , 64(2), 161-170.
63	GC6-2000 64(2)	Stuardo, J. & Vargas-Almonacid, P. (2000). Moluscos terrestres de Chile. Sinonimia y problemas relacionados: 1. Familias veronicellidae, pipillidae y achatinellidae (gastropoda: pulmonata). <i>Gayana (Concepc.)</i> , 64(2), 171-188.
64	GC1-2001 65(1)	Marino, P. & Spinelli, G. (2001). El subgénero <i>forcipomya</i> (<i>euprojoannisia</i>) en la Patagonia (diptera: ceratopogonidae). <i>Gayana (Concepc.)</i> , 65(1), 11-18.
65	GC2-2001 65(1)	Chiaraviglio, M. & Briguera, V. (2001). Participación de señales químicas en el reconocimiento y discriminación de sexos en <i>boa constrictor</i> occidentales (serpentes, boidae). <i>Gayana (Concepc.)</i> , 65(1), 5-10.

66	GC3-2001 65(1)	Quirán, E., Corró Molas, B., Caramuti, V. & Bernardos, J. (2001). Determinación preliminar de subcastas en obreras de <i>Acromyrmex lobicornis emery</i> 1887 (hymenoptera: formicidae). <i>Gayana (Concepc.)</i> , 65(1), 19-25.
67	GC4-2001 65(1)	Rodríguez, M. & Angulo, A. (2001). Tres nuevas especies del género <i>scriptania hampson</i> 1905, en Chile (lepidoptera: noctuidae: hadeninae). <i>Gayana (Concepc.)</i> , 65(1), 27-36.
68	GC5-2001 65(1)	Sagretti, L. & Bistoni, M. (2001). Alimentación de <i>Odontesthes bonariensis</i> (cuvier y valenciennes 1835) (atheriniformes, atherinidae) en la Laguna Salada de Mar Chiquita (Córdoba, Argentina). <i>Gayana (Concepc.)</i> , 65(1), 37-42.
69	GC6-2001 65(1)	Lépez, I., Furet, L. & Aracena, O. (2001). Población de emérita análoga (<i>stimpson</i> 1857) en playas Amarilla y Rinconada, Antofagasta: aspectos abióticos, bióticos y concentración de cobre. <i>Gayana (Concepc.)</i> , 65(1), 55-76.
70	GC1-2001 65(2)	Quirán, E. & Steibel, J. (2001). Relación entre el peso de <i>Acromyrmex lobicornis emery</i> 1887 (hymenoptera:formicidae) y el peso de la carga, en condiciones de laboratorio. <i>Gayana (Concepc.)</i> , 65(2), 113-118.
71	GC2-2001 65(2)	Retana-Salazar, A. & Soto-Rodríguez, G. (2001). Filogenia del género neotropical <i>zeugmatothrips</i> (thysanoptera: phlaeothripidae). <i>Gayana (Concepc.)</i> , 65(2), 119-128.
72	GC3-2001 65(2)	Mendoza, G. & Jerez, V. (2001). Coleópteros epigeos asociados a <i>Peumus boldus</i> mol. En la reserva nacional Isla Mocha. Chile. (insecta - coleoptera). <i>Gayana (Concepc.)</i> , 65(2), 129-136.
73	GC4-2001 65(2)	Muñoz, P. & Salamanca, M. (2001). Flujo de plomo particulado a los sedimentos marinos y su incorporación en <i>Paraprionospio pinnata</i> (polychaeta: spionidae) en Bahía Concepción (36° lat. S), Chile. <i>Gayana (Concepc.)</i> , 65(2), 155-166.

74	GC5-2001 65(2)	Milessi, A., Vögler, R. & Bazzino, G. (2001). Identificación de tres especies del género <i>squatina</i> (chondrichthyes, squatinidae) en La Zona Común de Pesca Argentino-Uruguaya (ZCPAU). <i>Gayana (Concepc.)</i> , 65(2), 167-172.
75	GC6-2001 65(2)	Muñoz, E., Mendoza, G. & Valdovinos, C. (2001). Evaluación rápida de la biodiversidad en cinco sistemas lénticos de Chile Central: macroinvertebrados Bentónicos. <i>Gayana (Concepc.)</i> , 65(2), 173-180.
76	GC7-2001 65(2)	Bazzino, G. & Quiñones, R. (2001). Alimentación del calamar <i>Illex argentinus</i> (cephalopoda, ommastrephidae) en la Zona Común de Pesca Argentino-Uruguaya (ZCPAU). <i>Gayana (Concepc.)</i> , 65(2), 173-180.
77	GC1-2002 66(1)	Berrios, P. (2002). Artrópodos asociados a suelo de renovales de <i>Nothofagus obliqua</i> (mirb.) Oersted (fagacea) en la zona costera de la VIII Región de Chile. <i>Gayana (Concepc.)</i> , 66(1), 1-6.
78	GC2-2002 66(1)	Moreno, R., Moreno, J., Ortiz, J., Victoriano, P. & Torres-Pérez, F. (2002). Herpetofauna del Parque Nacional Llanos de Challe (III Región, Chile). <i>Gayana (Concepc.)</i> , 66(1), 7-10.
79	GC3-2002 66(1)	Elgueta, M. & Cerda, M.†. (2002). <i>Acalodegma vidali</i> , nueva especie altoandina de prioninae (coleoptera: cerambycidae) de Chile. <i>Gayana (Concepc.)</i> , 66(1), 11-21.
80	GC4-2002 66(1)	Olivares, T., Rodríguez, M., Navarro, M. & Angulo, A. (2002). Descripción de las estructuras genitales y estados inmaduros de <i>Pseudocerura thoracica butleri</i> 1882 (lepidoptera: noctuidae: cuculliinae). <i>Gayana (Concepc.)</i> , 66(1), 23-30.
81	GC5-2002 66(1)	Lunaschi, L. (2002). Redescrición de <i>Saccocoelioides octavus szidat</i> 1970 y <i>S. bacilliformis szidat</i> , 1973 (digenea, haploporidae), parásitos de peces del género <i>astyanax baird & girard</i> 1854 (pisces: characidae). <i>Gayana (Concepc.)</i> , 66(1), 31-37.

82	GC6-2002 66(1)	Sobrazo, C., Jara-Seguel, P., Peredo, S. & Parada, E. (2002). Primer registro de <i>Musculium argentinum</i> (d'orbigny 1835) (bivalvia: sphaeriidae) en aguas continentales chilenas. <i>Gayana (Concepc.)</i> , 66(1), 39-43.
83	GC7-2002 66(1)	Alayo, M. & Iannacone, J. (2002). Ensayos ecotoxicológicos con petróleo crudo, diesel 2 y diesel 6 con dos subespecies de <i>Brachionus plicatilis</i> müller 1786 (rotifera: monogononta). <i>Gayana (Concepc.)</i> , 66(1), 45-58.
84	GC1-2002 66(2)	Castro, D. & Cicchino, A. (2002). Una especie nueva de phtheiropoios (phthiraptera, amblycera) parásita de <i>ctenomys</i> sp. (rodentia, octodontidae) de la provincia de Córdoba, Argentina. <i>Gayana (Concepc.)</i> , 66(2), 89-96.
85	GC2-2002 66(2)	Vargas, H., Parra, L., Vargas, H. & Bobadilla, D. (2002). Aspectos biológicos de <i>Eupithecia sibylla</i> butler 1882 (lepidoptera: geometridae). <i>Gayana (Concepc.)</i> , 66(2), 103-106.
86	GC3-2002 66(2)	Lara, G., Parada, E. & Peredo, S. (2002). Alimentación y conducta alimentaria de la almeja de agua dulce <i>Diplodon chilensis</i> (bivalvia: hyriidae). <i>Gayana (Concepc.)</i> , 66(2), 107-112.
87	GC4-2002 66(2)	Lara, G., Contreras, A. & Encina, F. (2002). La almeja de agua dulce <i>Diplodon chilensis</i> (bivalvia:hyriidae) potencial biofiltro para disminuir los niveles de coliformes en pozos. Experimento de laboratorio. <i>Gayana (Concepc.)</i> , 66(2), 113-118.
88	GC5-2002 66(2)	Palma, A., Figueroa, R., Ruiz, V., Araya, E. & Berríos, P. (2002). Composición de la dieta de <i>Oncorhynchus mykiss</i> (walbaum 1792) (pisces: salmonidae) en un sistema fluvial de baja intervención antrópica: Estero Nonguén, VIII Región, Chile. <i>Gayana (Concepc.)</i> , 66(2), 129-139.
89	GC1-2003 67(1)	Muñoz, L., Aguilera, M. & Casanueva, M. (2003). Prevalencia e intensidad de ectoparásitos asociados a <i>Tadarida brasiliensis</i> (geoffroy saint-hilaire, 1824) (chiroptera: molossidae) en (Concepc.). <i>Gayana (Concepc.)</i> , 67(1), 1-8.

90	GC2-2003 67(1)	Iannacone, J. (2003). Helmintos parásitos de <i>Atelopus bomolochus peters</i> 1973 (anura: bufonidae) de Piura, Perú. <i>Gayana (Concepc.)</i> , 67(1), 9-15.
91	GC3-2003 67(1)	Alarcón, M. (2003). Sifonapterofauna de tres especies de roedores de Concepción, VI Región, Chile. <i>Gayana (Concepc.)</i> , 67(1), 16-24.
92	GC4-2003 67(1)	Díaz-Páez, H. & Ortiz, J.C. (2003). Hábitos alimentarios de <i>Pleurodema thauli</i> (anura, leptodactylidae) en Concepción, Chile. <i>Gayana (Concepc.)</i> , 67(1), 25-32.
93	GC5-2003 67(1)	Grob, C., Quiñones, R. & Figueroa, D. (2003). Cuantificación del transporte de agua costa-océano a través de filamentos y remolinos ricos en Clorofila a, en la zona centro-sur de Chile (35.5-37.5 °S). <i>Gayana (Concepc.)</i> , 67(1), 55-67.
94	GC6-2003 67(1)	Hernández, K., González, R. & Quiñones, R. (2003). Actividad de la lactato dehidrogenasa (ldh) en <i>Emerita analoga</i> (stimpson 1857) crustacea, hippidae. <i>Gayana (Concepc.)</i> , 67(1), 68-74.
95	GC7-2003 67(1)	Sánchez, R., Galvis, G. & Victoriano, P. (2003). Relación entre características del tracto digestivo y los hábitos alimentarios de peces del Río Yucao, sistema del Río Meta (Colombia). <i>Gayana (Concepc.)</i> , 67(1), 75-86.
96	GC8-2003 67(1)	Sepúlveda, R., Moreno, R. & Carrasco, F. (2003). Diversidad de macroinvertebrados asociados a arrecifes de <i>Phragmatopoma moerchi</i> kinberg, 1867 (polychaeta: sabellariidae) en el intermareal rocoso de Cocholgue, Chile. <i>Gayana (Concepc.)</i> , 67(1), 45-54.
97	GC9-2003 67(1)	Silva, J., Torrejón, G., Bay-Schmith, E. & Larraín, A. (2003). Calibración del bioensayo de toxicidad aguda con <i>Daphnia pulex</i> (crustacea: cladocera) usando un tóxico de referencia. <i>Gayana (Concepc.)</i> , 67(1), 87-96.
98	RCHA1-2000 18(1)	Ghezzi, M., Lupidio, M., Castro, A., Gómez, S., Bilbao, G. & Landi, H. (2000). Desarrollo morfológico del estómago en terneros alimentados con dos

		sustitutos lácteos. <i>Revista Chilena de Anatomía</i> , 18(1), 19-26.
99	RCHA2-2000 18(1)	Pellegrino, F. (2000). Revisión anatómica del neopalio del perro. <i>Revista Chilena de Anatomía</i> , 18(1), 35-46.
100	RCHA3-2000 18(1)	Ghezzi, M., Alzola, R., Lupidio, M., Massone, A., Castro, A. & Rodríguez, J. (2000). Conducto hepatopancreático de la llama (<i>Lama glama</i>). <i>Revista Chilena de Anatomía</i> , 18(1), 27-34.
101	RCHA4-2000 18(1)	Galván, S., Gimeno, M., Nuviala, J., Gil, J., Laborda, J., Andreotti, C., Sbodio, O. & Pastor, R. (2000). Potencialidades y limitaciones del uso de recursos multimediales en la enseñanza de anatomía veterinaria. <i>Revista Chilena de Anatomía</i> , 18(1), 75-83.
102	RCHA5-2000 18(1)	Del Sol, M., Olave, E., Hettich, M. & Bustos, K. (2000). Arteria ulnar superficial. Importancia anatomo-clínica. <i>Revista Chilena de Anatomía</i> , 18(1), 85-88.
103	RCHA6-2000 18(1)	Oróstegui, C., Parraguez, V., Adaro, L., Peñailillo, P. & Cepeda, R. (2000). Cambios histológicos y morfométricos de las vesículas seminales, inducidos por variaciones estacionales, en la Chinchilla laniger (grey) en cautiverio. <i>Revista Chilena de Anatomía</i> , 18(1), 89-96.
104	RCHA7-2000 18(1)	Braga, M., Gabrielli, C., De Souza, A., Rodrigues, C. & Marino, J. C. (2000). Huesos suturales en el pterión. <i>Revista Chilena de Anatomía</i> , 18(1), 97-101.
105	RCHA8-2000 18(1)	Vásquez, B., Peña, P., Veuthey, C. & Sánchez, R. (2000). Localización inmunocitoquímica de la Serpina CBG-símil en el sistema reproductor de mamíferos hembra. <i>Revista Chilena de Anatomía</i> , 18(1), 109-116.
106	RCHA9-2000 18(1)	Henríquez, J., Sandoval, P. & Fuentes, R. (2000). Anatomía radiológica del hueso hioides. <i>Revista Chilena de Anatomía</i> , 18(1), 117-124.
107	RCHA1-2000	Rodríguez, A., Rojas, M., Montenegro, M. & Regadera, J. (2000). Expresión

	18(2)	de filamentos intermedios durante el desarrollo embrionario de cerdo (<i>Sus scrofa</i>) y bovino (<i>Bos taurus</i>). <i>Revista Chilena de Anatomía</i> , 18(2), 237-244.
108	RCHA2-2000 18(2)	Veiga, M., Egami, M., Ranzani-Paiva, M. & Rodrigues, E. (2000). Aspectos morfológicos y citoquímicos de las células sanguíneas de <i>Salminus maxillosus</i> Valenciennes, 1840 (Characiformes, characidae). <i>Revista Chilena de Anatomía</i> , 18(2), 245-250.
109	RCHA3-2000 18(2)	Mansilla, M., Veuthey, C., Vasconcellos, A. & Peña, P. (2000). Teratoma dermoide quístico córneo-conjuntivo en un perro ovejero alemán (<i>Canis familiaris</i>): Caso clínico. <i>Revista Chilena de Anatomía</i> , 18(2), 273-277.
110	RCHA4-2000 18(2)	Del Sol, M., Olave, E. & Gabrielli, C. (2000). Frecuencia e importancia anátomo-clínica del músculo accesorio del músculo cuadrado plantar. <i>Revista Chilena de Anatomía</i> , 18(2), 279-282.
111	RCHA5-2000 18(2)	Pellegrino, F. (2000). Evidencias anatómicas del área de cobertura de electrodos especiales temporales en electroencefalografía canina. <i>Revista Chilena de Anatomía</i> , 18(2), 287-295.
112	RCHA6-2000 18(2)	Rodrigues, C., Alves, A., Olave, E. & Gabrielli, C. (2000). Consideraciones anatómicas sobre la anastomosis arterial recto-sigmoídea. <i>Revista Chilena de Anatomía</i> , 18(2), 297-300.
113	RCHA7-2000 18(2)	Olave, E., Gabrielli, C., Braga, M., Del Sol, M. & De Souza, A. (2000). Ramo comunicante entre los nervios musculocutáneo y mediano en el hombre. <i>Revista Chilena de Anatomía</i> , 18(2), 301-304.
114	RCHA1-2001 19(1)	Merlo, A., Albanese, A., Gómez, E., Miño, J., Mascitti, T., Ingratta, A. & Albanese, E. (2001). Las relaciones de dominancia en el lóbulo límbico. <i>Revista Chilena de Anatomía</i> , 19(1), 5-10.
115	RCHA2-2001	Dallard, B., Romano, G., Lorente, J., Lamas, H. & Ortega, H. (2001). Estudio lectinhistoquímico y morfométrico de la glándula mamaria de llama (<i>Lama</i>

	19(1)	glama) durante el periodo de lactancia. <i>Revista Chilena de Anatomía</i> , 19(1), 11-16.
116	RCHA3-2001 19(1)	Alonso, C., Rafasquino, M., Anguis, F., Piove, M. & Idiart, J. (2001). Características macro y microscópicas del corazón y grandes vasos del coipo (<i>Myocastor coypus</i> , Molina) de diferentes edades. <i>Revista Chilena de Anatomía</i> , 19(1), 29-37.
117	RCHA4-2001 19(1)	Moine, R., Rivera, M., Ferraris, G., Natali, J., Galán, A., Gigena, M., Barotto, O. & Vivas, A. (2001). Descripción histológica de la substancia cortical dorsal en la parte media de la diáfisis del metacarpiano III en yegua mestiza. <i>Revista Chilena de Anatomía</i> , 19(1), 39-43.
118	RCHA5-2001 19(1)	Gabrielli, C., Olave, E., Mandiola, E. & Braga, M. (2001). Arterias metatarsianas plantares en el hombre. <i>Revista Chilena de Anatomía</i> , 19(1), 51-56.
119	RCHA6-2001 19(1)	Perriard, D. & Losardo, R. (2001). La anatomía en una propuesta curricular integrada. <i>Revista Chilena de Anatomía</i> , 19(1), 57-60.
120	RCHA1-2001 19(2)	Viegas, K., Orsi, A., Matheus, S., Francia-Farje, L. Orsi, D. & Mello, J. (2001). Características estructurales de la aorta de conejo (<i>Oryctolagus cuniculus</i>). <i>Revista Chilena de Anatomía</i> , 19(2), 131-137.
121	RCHA2-2001 19(2)	Barrovecchio, J., Pérez, B., Bella de Paz, L., Busmail, L. & Ruggero, E. (2001). Evaluación de los exámenes de anatomía normal en la Universidad Nacional de Rosario, Argentina. <i>Revista Chilena de Anatomía</i> , 19(2), 139-144.
122	RCHA3-2001 19(2)	Braga, A., Catirse, A., Lopes, R., Semprini, M., Sala, M. & Watanabe, I. (2001). Efecto tóxico del ión cadmio sobre la glándula submandibular de rata adulta. <i>Revista Chilena de Anatomía</i> , 19(2), 183-190.

123	RCHA4-2001 19(2)	Guiraldes, H., Oddó, H., Mena, B., Velasco, N. & Paulos, J. (2001). Enseñanza de la anatomía humana: experiencias y desafíos en una escuela de medicina. <i>Revista Chilena de Anatomía</i> , 19(2), 205-212.
124	RCHA5-2001 19(2)	Rojas, M., Vignon, X., Montenegro, M., Del Sol, M. Bustos-Obregón, E. & Fléchon, J. (2001). Caracterización, aislamiento y cultivo de células germinales primordiales de conejo. <i>Revista Chilena de Anatomía</i> , 19(2), 213-220.
125	RCHA6-2001 19(2)	Vásquez, B. & Del Sol, M. (2001). Estudio morfológico de la glándula bulbouretral de conejo (<i>Oryctolagus cuniculus</i>). <i>Revista Chilena de Anatomía</i> , 19(2), 221-228.
126	RCHA1-2001 19(3)	Vilela, A., Vieira da Costa, J., Lopes, R. & Sala, M. (2001). Acción de la hipervitaminosis A en el germen dentario de fetos de rata. Estudio histológico y morfométrico. <i>Revista Chilena de Anatomía</i> , 19(3), 231-238.
127	RCHA2-2001 19(3)	Amorim, M., De Amorim Júnior, A., De Oliveira Villarouco, F., Bernadez, D. & Duarte Costa Filho, J. (2001). Capacidad del estómago de conejos sin raza definida (SRD) (<i>Oryctolagus cuniculus</i>). <i>Revista Chilena de Anatomía</i> , 19(3), 259-262.
128	RCHA3-2001 19(3)	Silva, H. (2001). Ángulos del planp sagital de la columna lumbosacra en una muestra de adolescentes de la ciudad de Temuco, Chile. <i>Revista Chilena de Anatomía</i> , 19(3), 271-277.
129	RCHA4-2001 19(3)	Vásquez, B. (2001). Presencia de CBG en el estroma ovárico de mamíferos. <i>Revista Chilena de Anatomía</i> , 19(3), 279-284.
130	RCHA5-2001 19(3)	Castro, A., Ghezzi, M., Alzota, R., Lupidio, M. & Rodríguez, J. (2001). Morfología del hígado de llama (<i>Lama glama</i>). <i>Revista Chilena de Anatomía</i> , 19(3), 291-296.
131	RCHA6-2001	Adaro, L., Mendoza, J., Cepeda, R. & Oróstegui, C. (2001). Estudio anatómico-radiográfico de las vesículas seminales de la chinchilla (<i>Chinchilla laniger</i>).

	19(3)	Grey) en cautiverio. <i>Revista Chilena de Anatomía</i> , 19(3), 297-300.
132	RCHA7-2001 19(3)	Montenegro, M. & Rojas, M. (2001). Transformación epitelio-mesenquimática durante el desarrollo embrionario. <i>Revista Chilena de Anatomía</i> , 19(3), 301-310.
133	RCHA8-2001 19(3)	Gabrielli, C., Braga, M. Teixeira, Olave, E., Froehner Jr., I. & Kunze, I. (2001). Aspectos anatómicos y biométricos de la arteria dorsal del pie en humanos. <i>Revista Chilena de Anatomía</i> , 19(3), 325-330.
134	RCHA1-2002 20(1)	Espinoza, O., Bustos-Obregón, E. & Suja, J. (2002). Efecto del parathion sobre el índice de apoptosis en hepatocitos de ratones CF1. <i>Revista Chilena de Anatomía</i> , 20(1), 29-36.
135	RCHA2-2002 20(1)	Collipal, E. (2002). Conceptualización a través de redes semánticas naturales de los módulos de autoaprendizaje en anatomía humana. <i>Revista Chilena de Anatomía</i> , 20(1), 63-67.
136	RCHA1-2002 20(2)	Merlo, A., Albanese, A., Miño, J., Gómez, E., Saubidet, A., Mascitti, T., Ingrata, A. & Albanese, E. (2002). Cuerpo caloso en el plano mediano de imágenes de resonancia magnética: perímetro, superficie y forma en relación a edad y sexo. <i>Revista Chilena de Anatomía</i> , 20(2), 131-138.
137	RCHA2-2002 20(2)	Felipe, A., Cabodevila, J., & Callejas, S. (2002). Morfología de la unión útero-tubárica en <i>Myocastor Coypus</i> (Coipo). <i>Revista Chilena de Anatomía</i> , 20(2), 139-144.
138	RCHA3-2002 20(2)	Inzunza, O. & Bravo, H. (2002). Animación computacional de fotografías, un real aporte al aprendizaje práctico de Anatomía Humana. <i>Revista Chilena de Anatomía</i> , 20(2), 151-157.
139	RCHA4-2002 20(2)	Olivares, R., Soto-Moyano, R., Hernández, A. & Aboitiz, F. (2002). La malnutrición prenatal proteica afecta el desarrollo del cuerpo caloso anterior

		<i>Revista Chilena de Anatomía, 20(2), 159-163.</i>
140	RCHA5-2002 20(2)	Briones, F., Calderón, M., Muñoz, J., Venegas, F. & Araya, N. (2002). El anticuerpo monoclonal Ki-67 como elemento de valor diagnóstico y pronóstico en neoplasias mamarias caninas. <i>Revista Chilena de Anatomía, 20(2), 165-168.</i>
141	RCHA6-2002 20(2)	Vásquez, B. & Del Sol, M. (2002). Complejo prostático en el conejo (<i>Oryctolagus cuniculus</i>). <i>Revista Chilena de Anatomía, 20(2), 175-180.</i>
142	RCHA7-2002 20(2)	Amorim, M., Amorim Júnior, A., Silva Júnior, V., Villarouco, F. & Henrique V. (2002). Longitud total del intestino de conejos sin raza definida (<i>Oryctolagus cuniculus</i>). <i>Revista Chilena de Anatomía, 20(2), 181-183.</i>
143	RCHA8-2002 20(2)	Ferraris M., Carranza, M., Actis, A. & Simbrón, A. (2002). Cambios estructurales del complejo articular temporomandibular (CATM) humano en distintas edades gestacionales. <i>Revista Chilena de Anatomía, 20(2), 185-191.</i>
144	RCHA9-2002 20(2)	Del Sol, M. & Olave, E. (2002). Origen axilar de la arteria profunda del brazo. <i>Revista Chilena de Anatomía, 20(2), 201-204.</i>
145	RCHA10-2002 20(2)	Silva, H., Moraes, S. & Tashiro, T. (2002). Análisis morfométrico de los elementos de la cadena estilohioídea. <i>Revista Chilena de Anatomía, 20(2), 205-210.</i>
146	RCHA11-2002 20(2)	Cabalin, S., Collipal, L. & Doussoulin, S. (2002). Evaluación de módulos de autoaprendizaje en Anatomía Humana. <i>Revista Chilena de Anatomía, 20(2), 211-214.</i>
147	RCHA12-2002 20(2)	Montes, G., Cavallieri, A., Favaretto, P. & Del Sol, M. (2002). Una experiencia con la evaluación individual de la productividad académica en la Facultad de Medicina de la Universidad de San Pablo, Brasil. <i>Revista Chilena de Anatomía, 20(2), 215-230.</i>
148	RCHA13-2002	Olave, E., Gabrielli, C., Braga, M. & Del Sol, M. (2002). Aspectos biométricos

	20(2)	de los ramos motores del nervio musculocutáneo para el músculo braquial. <i>Revista Chilena de Anatomía, 20(2), 231-236.</i>
149	RCHA1-2002 20(3)	Torrez, J. & Del Sol, M. (2002). Músculos flexores accesorios del IV y V dedos del pie. Variación anatómica. <i>Revista Chilena de Anatomía, 20(3), 251-253.</i>
150	RCHA2-2002 20(3)	Babinski, M., Chagas, M., Costa, W. & Pereira, M. (2002). Morfología y fracción del área del lumen glandular de la zona de transición en la próstata humana. <i>Revista Chilena de Anatomía, 20(3), 255-262.</i>
151	RCHA3-2002 20(3)	Parra, R. & Del Sol, M. (2002). Estudios histológico e histoquímico del epidídimo de conejo (<i>Oryctolagus cuniculus</i>). <i>Revista Chilena de Anatomía, 20(3), 269-274.</i>
152	RCHA4-2002 20(3)	Brasil, F., Babinski, M., Sgrott, E. A. & Luz, H. (2002). El conocimiento de la anatomía topográfica de los estudiantes en la práctica del examen físico. Análisis cuantitativo. <i>Revista Chilena de Anatomía, 20(3), 275-280.</i>
153	RCHA5-2002 20(3)	Galán, A., Rivera, M., Moine, R., Ferraris, G., Gigena, M. & Natali, J. (2002). Propiedades morfométricas del metacarpiano III de potrillos mestizos. <i>Revista Chilena de Anatomía, 20(3), 285-290.</i>
154	RCHA6-2002 20(3)	Mandiola, E., Olave, E. & Alarcón, E. (2002). Relaciones anatómicas entre la arteria cerebelar anterior inferior y el nervio trigémino. <i>Revista Chilena de Anatomía, 20(3), 291-294.</i>
155	RCHA1-2003 21(1)	Schencke, C. & del Sol, M. (2003). Morfología del conducto deferente en el conejo (<i>Oryctolagus cuniculus</i>). <i>International Journal of Morphology, 21(1), 15-21.</i>
156	RCHA2-2003 21(1)	Machado, F., Babinski, M., Brasil, F., Favorito, L., Abidu-Figueiredo, M. & Costa, M. (2003). Variaciones anatómicas entre el nervio isquiático y el músculo piriforme durante el período fetal humano. <i>International Journal of</i>

		<i>Morphology</i> , 21(1), 29-35.
157	RCHA3-2003 21(1)	Lacerda, S., Sala, M., Lopes, R., Semprini, M. & Watanabe, I. (2003). Efectos de la administración prenatal de cafeína sobre la glándula submandibular de fetos de rata. <i>International Journal of Morphology</i> , 21(2), 37-42.
158	RCHA4-2003 21(1)	García-Hernández, F. (2003). Evaluación del aprendizaje práctico de la Anatomía Humana para Odontología en la Universidad de Antofagasta, Chile. <i>International Journal of Morphology</i> , 21(2), 43-47.
159	RCHA1-2003 21(2)	Merlo, A., Albanese, E., Gómez, E., Miño, J., Ingratta, A., Mascitti, T. & Albanese, A. (2003). Compensación: un nuevo concepto relacionado con la lateralidad cerebral. Su aplicación a regiones del lóbulo frontal, incluida el área motora del lenguaje (de broca). <i>International Journal of Morphology</i> , 21(2), 113-116.
160	RCHA2-2003 21(2)	Inzunza, O., D'Acuña, E. & Bravo, H. (2003). Evaluación práctica de anatomía. Rendimiento de los alumnos de primer año de medicina ante distintas formas de preguntar. <i>International Journal of Morphology</i> , 21(2), 131-136,
161	RCHA3-2003 21(2)	Babinski, M., Sgrott, E., Luz, H., Brasil, F., Chagas, M., & Abidu-Figueiredo, M. (2003). La relación de los estudiantes con el cadáver en el estudio práctico de anatomía: la reacción e influencia en el aprendizaje. <i>International Journal of Morphology</i> , 21(2), 137-142.
162	RCHA4-2003 21(2)	Henríquez, J., Fuentes, R., Sandoval, P. & Muñoz, A. (2003). Análisis de la estabilidad ortostática cráneocervical en adultos jóvenes mapuches. <i>International Journal of Morphology</i> , 21(2), 149-153.
163	RCHA5-2003 21(2)	Olave, E. & Braga, M. (2003). Inervación del músculo semitendinoso. Biometría de sus ramos y localización de sus puntos motores. <i>International Journal of Morphology</i> , 21(2), 161-165.

164	RCI1-2000 17(3)	Pavletic, C. (2000). Hantavirus: Su distribución geográfica entre los roedores silvestres de Chile. <i>Revista Chilena de Infectología</i> , 17(3), 186-196.
165	RCI2-2000 17(3)	Spotorno, A., Palma, R.E. & Valladares, J.P. (2000). Biología de roedores reservorios de hantavirus en Chile. <i>Revista Chilena de Infectología</i> , 17(3), 197-210.
166	RCI3-2000 17(3)	Galeno, H., Villagra, E., Fernández, J., Ramírez, E. & Mora, J. (2000). Técnicas diagnósticas de infección humana por hantavirus. <i>Revista Chilena de Infectología</i> , 17(3), 211-215.
167	RCI4-2000 17(3)	Galeno, H., Villagra, E., Fernández, J., Ramírez, E. & Mora, J. (2000). Infección por hantavirus en humanos: Experiencia del Laboratorio de Referencia para Enfermedades Infecciosas Emergentes. <i>Revista Chilena de Infectología</i> , 17(3), 216-219.
168	RCI5-2000 17(3)	Sotomayor, V. & Aguilera, X. (2000). Epidemiología de la infección humana por hantavirus en Chile. <i>Revista Chilena de Infectología</i> , 17(3), 220-232.
169	RCI6-2000 17(3)	Navarrete, M., Saldias, F., Mancilla, R., Zaror, L. & Ferres, M. (2000). Evidencia clínico epidemiológica de la existencia de síndrome pulmonar por hantavirus en Valdivia-Chile desde 1993. <i>Revista Chilena de Infectología</i> , 17(3), 233-240.
170	RCI7-2000 17(3)	Castillo, C., Naranjo, J. & Ossa, G. (2000). Síndrome cardiopulmonar por hantavirus en 21 adultos en la IX Región de Chile. <i>Revista Chilena de Infectología</i> , 17(3), 241-247.
171	RCI8-2000 17(3)	Soza, G., Lorca, P., Puebla, S., Wenzel, M., Navarrete, M., Villagra, E., Mora, J., Levis, S. & Avilés, G. (2000). Síndrome pulmonar por hantavirus en población infantil. Chile: Regiones IX y X. 1998-2000. <i>Revista Chilena de Infectología</i> , 17(3), 248-257.

172	RCI9-2000 17(3)	Tapia, M., Mansilla, C. & Vera, J. (2000). Síndrome pulmonar por hantavirus: Experiencia clínica en diagnóstico y tratamiento. Hospital Coyhaique-Chile. <i>Revista Chilena de Infectología</i> , 17(3), 258-269.
173	RCI10-2000 17(3)	Pavletic, C. (2000). Hantavirus: Medidas de prevención y manejo ambiental. <i>Revista Chilena de Infectología</i> , 17(3), 270-275.
174	RCI11-2000 17(4)	Salazar, A., Pret, C., Chavez, A., Garcia, P., Millan, Z., Goycoolea, M., Parada, J., Urra, L., Ahumada, E., Yoma, T., Duque, C., Herman, O. & Quiroga, T. (2000). Evaluación de métodos diagnósticos para sífilis congénita. <i>Revista Chilena de Infectología</i> , 17(4), 289-296.
175	RCI12-2000 17(4)	Chávez, A., Álvarez, A. & Wu, E. & Comité Nacional de SIDA Pediátrico y Sociedad Chilena de Pediatría (2000). Transmisión vertical de la infección por virus de inmunodeficiencia humana: Impacto de la aplicación del protocolo ACTG 076 en Chile. <i>Revista Chilena de Infectología</i> , 17(4), 297-301.
176	RCI11-2001 18(1)	Mella, S., Zemelman, C., Bello, H., Domínguez, M., González, G. & Zemelman, R. (2001). Antimicrobianos: Propiedades microbiológicas, clasificación y relación estructura-actividad de cefalosporinas e importancia de las cefalosporinas de cuarta generación. <i>Revista Chilena de Infectología</i> , 18(1), 7-19.
177	RCI-2001 18(2)	Fica, A., Alexandre, M., Prat, S., Fernández, A., Fernández, J. & Heitmann, I. (2001). Cambios epidemiológicos de las salmonelosis en Chile. Desde <i>Salmonella typhi</i> a <i>Salmonella enteritidis</i> . <i>Revista Chilena de Infectología</i> , 18(2), 85-93.
178	RCI11-2001 18(3)	Soler, M., Vergara, R., Silva, M., Collao, X. & Navarrete, E. (2001). Detección de enterovirus mediante transcripción reversa y reacción de polimerasa en cadena en líquido cefalorraquídeo de niños con meningitis aséptica. <i>Revista Chilena de Infectología</i> , 18(3), 175-181.

179	RCI2-2001 18(3)	Castillo, L., Espinoza, R., Castro, J., Jonquera, R., Hayden, A., Núñez, N., Ramos, A. & Acuña, G. (2001). Evaluación in vitro de linezolid frente a <i>Staphylococcus aureus</i> meticilina resistentes aislados de infecciones osteoarticulares. <i>Revista Chilena de Infectología</i> , 18(3), 182-186.
180	RCI3-2001 18(3)	Guzmán, A., Abarzúa, F., Belmar, C. & García, P. (2001). Resultados de la aplicación del protocolo basado en screening para la búsqueda de <i>Streptococcus agalactiae</i> en el tercer trimestre del embarazo. Posible impacto sobre la sepsis neonatal precoz por este agente. <i>Revista Chilena de Infectología</i> , 18(3), 187-192.
181	RCI4-2001 18(3)	Ulloa, M., Silva, V., Piñones, E., Porte, L., Fica, A. & Pinto, E. (2001). Caracterización molecular de cepas de <i>Streptococcus pyogenes</i> aisladas de cuadros invasores basada en el polimorfismo del regulón vir. <i>Revista Chilena de Infectología</i> , 18(3), 193-202.
182	RCI1-2001 18(4)	González, R., Coria, P. & Martínez, M. (2001). Colonización de recién nacidos prematuros y de sus madres por <i>Ureaplasma urealyticum</i> . <i>Revista Chilena de Infectología</i> , 18(4), 250-254.
183	RCI2-2001 18(4)	Ulloa, M., Porte, L., Carmi, A, Varela, C. & Fica, A. (2001). Comparación de reacción de polimerasa en cadena, látex y antibiograma para detección de <i>Staphylococcus aureus</i> meticilina resistente. <i>Revista Chilena de Infectología</i> , 18(4), 255-260.
184	RCI1-2002 19(1)	Muñoz, G. & Venegas, M. (2002). Estudio de reproducibilidad del examen de cuantificación de VIH-1 por la técnica Nuclisens HIV-1 QT® en muestras de plasma de pacientes infectados con VIH-1. <i>Revista Chilena de Infectología</i> , 19(1), 25-31.
185	RCI2-2002 19(1)	Mella, S., Sepúlveda, M., Acosta, P., Bello, H., Domínguez, M., González, G., Zemelman, R. & Cofré, O. (2002). Aislamiento de <i>Enterococcus faecium</i> resistente a vancomicina con genotipo van B en el Hospital Clínico Regional

		de (Concepc.). <i>Revista Chilena de Infectología</i> , 19(1), 32-36.
186	RCI1-2002 19(2)	Klaassen, J., García, P., Maíz, A. & Campano, M. (2002). Mecanismos de contaminación de las fórmulas para nutrición enteral. <i>Revista Chilena de Infectología</i> , 19(2), 69-73.
187	RCI1-2002 19(3)	García, P., De la Cerda, G., Calvo, M., Godoy, R., Covarrubias, C., Potin, M. & Quiroga, T. (2002). Inmunogenicidad de una vacuna recombinante anti hepatitis B en personal de salud. <i>Revista Chilena de Infectología</i> , 19(3), 133-139.
188	RCI1-2002 19(4)	Fica, A. (2002). Prevención y tratamiento de la neumonía adquirida en la comunidad en pacientes adultos: Un enfoque para la atención primaria. <i>Revista Chilena de Infectología</i> , 19(4), 207-219.
189	RCI1-2003 20(1)	Navarrete, M., Wilson, M., Oth, L., Hoffman, I. & Zaror, L. (2003). Frecuencia de Mycoplasma pneumoniae y Chlamydia pneumoniae en pacientes con distress respiratorio y serología negativa para hantavirus. <i>Revista Chilena de Infectología</i> , 20(1), 7-10.
190	RCI2-2003 20(1)	Casellas, J., Visser, M., Coco, B., Mac Dougall, N., Cohen, h., Soutric, J., Arenoso, H. & Grupo Colaborativo Latinoamericano. (2003). Actividad in vitro de niveles séricos y urinarios de amoxicilina y amoxicilina-sulbactam sobre 820 cepas de Escherichia coli aisladas de infecciones urinarias bajas extrahospitalarias. Estudio sudamericano. <i>Revista Chilena de Infectología</i> , 20(1), 11-18.
191	RCI3-2003 20(1)	Diomedi, A. (2003). La guerra biológica en la conquista del nuevo mundo. Una revisión histórica y sistemática de la literatura. <i>Revista Chilena de Infectología</i> , 20(1), 19-25.
192	RCI4-2003 20(1)	Corvalán, A., Aguayo, F., Lévicán, J. & Corvalán, I. (2003). Biología molecular en Infectología. Parte II: Diagnóstico molecular de agentes

		infecciosos. <i>Revista Chilena de Infectología</i> , 20(1), 26-38.
193	RCI1-2003 20(3)	Delpiano, L., Guillén, B. & Casado, C. (2003). Comportamiento clínico-epidemiológico de la influenza en niños hospitalizados. <i>Revista Chilena de Infectología</i> , 20(3), 159-165.
194	RCI2-2003 20(3)	Retamal, P., Martínez, A. & Ábalos, P. (2003). Secuencias de inserción IS6110 e IS1081 en cepas de Mycobacterium bovis provenientes de bovinos beneficiados en la Región Metropolitana. <i>Revista Chilena de Infectología</i> , 20(3), 166-170.
195	RCI3-2003 20(3)	Gaete, L., Schatloff, O., Anziani, F., Serrano, C., Ceballos, V., Bello, P., Ortiz, M., Saavedra, I. & Yatabe, Y. (2003). Bioequivalencia entre dos formulaciones de claritromicina comprimidos de liberación modificada existentes en el mercado chileno. <i>Revista Chilena de Infectología</i> , 20(3), 171-177.
196	RCI4-2003 20(3)	Gaete, L., Schatloff, O., Bello, P., Serrano, C., Ceballos, V., Anziani, F., Solís, J., Venegas, P. & Saavedra, I. (2003). Bioequivalencia entre dos formulaciones de claritromicina en suspensión pediátrica existentes en el mercado chileno. <i>Revista Chilena de Infectología</i> , 20(3), 178-183.
197	RCHN1-2000 73(1)	Pardiñas, U., Moreira, G., García-Esponda, C. & De Santis, L. (2000). Deterioro ambiental y micromamíferos durante el Holoceno en el nordeste de la estepa patagónica (Argentina). <i>Revista Chilena de Historia Natural</i> , 73(1), 9-21.
198	RCHN2-2000 73(1)	Hernández, C., Galleguillos, R. & Oyarzún, C. (2000). Diferenciación genética de <i>Merluccius gayi gayi</i> y <i>Merluccius gayi peruanus</i> (Pisces, Merlucciidae) y antecedentes paleogeográficos de su área de distribución. <i>Revista Chilena de Historia Natural</i> , 73(1), 23-29.
199	RCHN3-2000 73(1)	De Los Rios, P. & Zúñiga, O. (2000). Comparación biométrica del lóbulo frontal en poblaciones americanas de <i>Artemia</i> (Anostraca, Artemiidae). <i>Revista Chilena de Historia Natural</i> , 73(1), 31-36.

		<i>Chilena de Historia Natural</i> , 73(1), 31-38.
200	RCHN4-2000 73(1)	Luna, F. Antinuchi, D & Busch, C. (2000). Ritmos de actividad locomotora y uso de las cuevas en condiciones seminaturales en <i>Ctenomys talarum</i> (Rodentia, Octodontidae). <i>Revista Chilena de Historia Natural</i> , 73(1), 39-46.
201	RCHN5-2000 73(1)	Martínez, G. (2000). Conducta alimentaria de <i>Daphnia ambigua</i> Scourfield 1947, <i>Moina micrura</i> Kurz 1874 y <i>Ceriodaphnia dubia</i> Richard 1895 (Cladocera) frente a un gradiente de concentración de alimento. <i>Revista Chilena de Historia Natural</i> , 73(1), 47-54.
202	RCHN6-2000 73(1)	Pinto, M., Lizana, C., Pastenes, C., Riquelme, A. & Berti, M. (2000). Efecto de la radiación ultravioleta-B sobre el crecimiento y la fotosíntesis de siete variedades de trigo (<i>Triticum aestivum</i> L.). <i>Revista Chilena de Historia Natural</i> , 73(1), 55-66.
203	RCHN7-2000 73(1)	Jerez, V. (2000). Diversidad y patrones de distribución geográfica de insectos coleópteros en ecosistemas desérticos de la región de Antofagasta, Chile. <i>Revista Chilena de Historia Natural</i> , 73(1), 79-92.
204	RCHN8-2000 73(1)	Lancellotti, D. & Vásquez, J. (2000). Zoogeografía de macroinvertebrados bentónicos de la costa de Chile: contribución para la conservación marina. <i>Revista Chilena de Historia Natural</i> , 73(1), 99-129.
205	RCHN9-2000 73(1)	Fariña, J., Aldana, M., Ogalde, F. & Ojeda, P. (2000). Ecología trófica de <i>Girella laevifrons</i> (Pisces: Kyphosidae) en zonas intermareales rocosas del norte de Chile afectadas y no afectadas por contaminantes derivados de la minería del cobre. <i>Revista Chilena de Historia Natural</i> , 73(1), 139-149.
206	RCHN10-2000 73(1)	Cisternas, M., Torres, L., Urrutia, R., Araneda, A & Parra, O. (2000). Comparación ambiental, mediante registros sedimentarios, entre las condiciones prehispánicas y actuales de un sistema lacustre. <i>Revista Chilena de Historia Natural</i> , 73(1), 151-162.

207	RCHN11-2000 73(1)	Figueroa, J. & Castro, S. (2000). Efecto de herbívoros y patógenos en la sobrevivencia y crecimiento de plántulas en un fragmento del bosque templado húmedo de Chiloé, Chile. <i>Revista Chilena de Historia Natural</i> , 73(1), 163-173.
208	RCHN1-2000 73(2)	Astete-Espinoza, L. & Cáceres, C. (2000). Efecto del parasitismo del isópodo bopírido <i>Ionella agassizi</i> (Isopoda: Epicaridea) (Bornnier, 1900) sobre la fisiología nutricional del nape. <i>Revista Chilena de Historia Natural</i> , 73(2), 243-252.
209	RCHN2-2000 73(2)	Cáceres, C. & Ojeda, P. (2000). Patrones de forrajeo en dos especies de peces intermareales herbívoros de las costas de Chile: Efecto de la abundancia y composición química del alimento. <i>Revista Chilena de Historia Natural</i> , 73(2), 253-260.
210	RCHN3-2000 73(2)	Martínez, G., Molina, X., Olmedo, M. & Montecino, V. (2000). Susceptibilidad a la radiación ultravioleta_B del dinoflagelado <i>Alexandrium catenella</i> Kofoid Balech y de la diatomea <i>Phaeodactylum tricornutum</i> Bohlin. <i>Revista Chilena de Historia Natural</i> , 73(2), 323-330.
211	RCHN4-2000 73(2)	Cortés, A., Rosenmann, M. & Bozinovic, F. (2000). Relación costo_beneficio en la termorregulación de <i>Chinchilla lanigera</i> . <i>Revista Chilena de Historia Natural</i> , 73(2), 351-357.
212	RCHN1-2000 73(4)	Saiz, F., Yates, L., Núñez, C., Daza, M., Varas, M. & Vivar, C. (2000). Biodiversidad del complejo de artrópodos asociados al follaje de la vegetación del norte de Chile, II región. <i>Revista Chilena de Historia Natural</i> , 73(4), 671-692.
213	RCHN2-2000 73(4)	Neira, E. & Lara, A. (2000). Desarrollo de cronologías de ancho de anillos para alerce (<i>Fitzroya cupressoides</i>) en Contao y Mirador, Chile. <i>Revista Chilena de Historia Natural</i> , 73(4), 693-703.
214	RCHN3-2000	Muñoz-Pedrerros, A., Moncada-Herrera, J. & Larrain, A. (2000). Variación de la percepción del recurso paisaje en el sur de Chile. <i>Revista Chilena de</i>

	73(4)	<i>Historia Natural</i> , 73(4), 729-738.
215	RCHN4-2000 73(4)	Quijada, P. & Cáceres, C. (2000). Patrones de abundancia, composición trófica y distribución espacial del ensamble de peces intermareales de la zona centro-sur de Chile. <i>Revista Chilena de Historia Natural</i> , 73(4), 739-747.
216	RCHN5-2000 73(4)	Borie, F., Rubio, R., Morales, A. & Castillo, C. (2000). Relación entre densidad de hifas de hongos micorrizógenos arbusculares y producción de glomalina con las características físicas y químicas de suelos bajo cero labranza. <i>Revista Chilena de Historia Natural</i> , 73(4), 749-756.
217	RCHN6-2000 73(4)	Véliz, D. & Vásquez, J. (2000). La Familia Trochidae (Mollusca: Gastropoda) en el norte de Chile: consideraciones ecológicas y taxonómicas. <i>Revista Chilena de Historia Natural</i> , 73(4), 757-769.
218	RCHN7-2000 73(4)	Martínez, G. & Montecino, V. (2000). Competencia en Cladocera: implicancias de la sobreposición en el uso de los recursos tróficos. <i>Revista Chilena de Historia Natural</i> , 73(4), 787-795.
219	RCHN1-2001 74(1)	Luxoro, M. & Vinet, R. (2001). Bases fisiológicas para una interacción entre las células cromafines y las endoteliales de la glándula adrenal. <i>Revista Chilena de Historia Natural</i> , 74(1), 39-45.
220	RCHN2-2001 74(1)	Godoy-Herrera, R. (2001). La conducta de larvas de <i>Drosophila</i> (Diptera: Drosophilidae): su etología, desarrollo, genética y evolución. <i>Revista Chilena de Historia Natural</i> , 74(1), 55-64.
221	RCHN3-2001 74(1)	Frías, D. (2001). Diferencias genéticas y morfológicas de los estados inmaduros de dos razas de <i>Rhagoletis conversa</i> (Bréthes) (Diptera: Tephritidae) asociadas a plantas <i>Solanum</i> : distribución geográfica y posible origen en simpatria de una nueva especie. <i>Revista Chilena de Historia Natural</i> , 74(1), 73-90.

222	RCHN4-2001 74(1)	Spotorno, A., Walker, L., Flores, S., Yevenes, M., Marín, J. & Zuleta, C. (2001). Evolución de los filotinos (Rodentia, Muridae) en los Andes del Sur. <i>Revista Chilena de Historia Natural</i> , 74(1), 151-166.
223	RCHN5-2001 74(1)	Rothhammer, F. & Manríquez, G. (2001). Danko Brncic y la coadaptación. <i>Revista Chilena de Historia Natural</i> , 74(1), 177-181.
224	RCHN6-2001 74(1)	Toro, C. (2001). Brenand: "resonancia y universalidad de las formas". <i>Revista Chilena de Historia Natural</i> , 74(1), 183-194.
225	RCHN1-2001 74(2)	Bertrán, C., Arenas, J. & Parra, O. (2001). Macrofauna del curso inferior y estuario del río Biobío (Chile): cambios asociados a variabilidad estacional del caudal hídrico. <i>Revista Chilena de Historia Natural</i> , 74(2), 331-340.
226	RCHN2-2001 74(2)	Silva, H., Acevedo, E. & Silva, P. (2001). Anatomía del tejido fotosintético de diez taxa de Opuntia establecidos en el secano árido mediterráneo de Chile. <i>Revista Chilena de Historia Natural</i> , 74(2), 341-351.
227	RCHN3-2001 74(2)	Giraldo-Cañas, D. (2001). Relaciones fitogeográficas de las sierras y afloramientos rocosos de la Guayana colombiana: Un estudio preliminar. <i>Revista Chilena de Historia Natural</i> , 74(2), 353-364.
228	RCHN4-2001 74(2)	Muñoz-Schick, M., Pinto, R., Mesa, A. & Moreira-Muñoz, A. (2001). "Oasis de neblina" en los cerros costeros del sur de Iquique, región de Tarapacá, Chile, durante el evento El Niño 1997-1998. <i>Revista Chilena de Historia Natural</i> , 74(2), 389-405.
229	RCHN5-2001 74(2)	Formas, J.R., Núñez, J. & Brieva, L. (2001). Osteología, taxonomía y relaciones filogenéticas de las ranas del género Telmatobufo (Leptodactylidae). <i>Revista Chilena de Historia Natural</i> , 74(2), 365-387.
230	RCHN6-2001 74(2)	Cruces, F., Urrutia, R., Araneda, A., Torres, L., Cisternas, M. & Vyverman, W. (2001). Evolución trófica de Laguna Grande de San Pedro (VIII Región, Chile) durante el último siglo, mediante el análisis de registros sedimentarios. <i>Revista Chilena de Historia Natural</i> , 74(2), 407-418.

		<i>Chilena de Historia Natural</i> , 74(2), 407-418.
231	RCHN7-2001 74(2)	Burru, L., Trivi de Mandri, M., Palacio, P. & Lombardo, M. (2001). Relaciones polen-vegetación de algunos taxa de la estepa patagónica (Argentina). <i>Revista Chilena de Historia Natural</i> , 74(2), 419-427.
232	RCHN8-2001 74(2)	Ide, S. & Lanfranco, D. (2001). Longevidad de <i>Orgilus obscurator</i> Ness (Hymenoptera: Braconidae) en presencia de diferentes fuentes de alimento. <i>Revista Chilena de Historia Natural</i> , 74(2), 469-475.
233	RCHN1-2001 74(3)	Daneri, G. (2001). Estudio de flujos verticales de carbono y nitrógeno en ambientes acuáticos controlados en la bahía de Knebel, Dinamarca, utilizando isótopos estables de nitrógeno y carbono (^{15}N y ^{13}C). <i>Revista Chilena de Historia Natural</i> , 74(3), 515-522.
234	RCHN2-2001 74(3)	Peñaloza, A., Caviaras, L., Arroyo, M. & Torres, C. (2001). Efecto nodriza intra-específico de <i>Kageneckia angustifolia</i> D. Don (Rosaceae) sobre la germinación de semillas y sobrevivencia de plántulas en el bosque esclerófilo montano de Chile central. <i>Revista Chilena de Historia Natural</i> , 74(3), 539-548.
235	RCHN3-2001 74(3)	Roig-Juñent, S. & Domínguez, M. (2001). Diversidad de la familia Carabidae (Coleoptera) en Chile. <i>Revista Chilena de Historia Natural</i> , 74(3), 549-571.
236	RCHN4-2001 74(3)	Camus, P. (2001). Biogeografía marina de Chile continental. <i>Revista Chilena de Historia Natural</i> , 74(3), 587-617.
237	RCHN5-2001 74(3)	Avilés, I., Maddison, W., Salazar, P., Estévez, G., Tufiño, P. & Cañas, G. (2001). Arañas sociales de la Amazonía ecuatoriana, con notas sobre seis especies sociales no descritas previamente. <i>Revista Chilena de Historia Natural</i> , 74(3), 619-638.
238	RCHN6-2001 74(3)	Quatrini, R., Albino, A. & Barg, M. (2001). Variación morfológica y dieta en dos poblaciones de <i>Liolaemus elongatus</i> Koslowsky, 1896 (Iguania: Tropiduridae) del noroeste patagónico. <i>Revista Chilena de Historia Natural</i> ,

		74(3), 639-651.
239	RCHN7-2001 74(3)	Torres-Contreras, H. (2001). Antecedentes biológicos de hormigas presentes en Chile publicados en revistas científicas nacionales y extranjeras durante el Siglo XX. <i>Revista Chilena de Historia Natural</i> , 74(3), 653-668.
240	RCHN8-2001 74(3)	Avendaño, R., Riquelme, C., Escribano, R. & Reyes, N. (2001). Sobrevivencia y crecimiento de post-larvas de <i>Argopecten purpuratus</i> (Lamarck, 1819) en Bahía Inglesa, Chile: Efectos del origen, distribución en la bahía y bacterioflora larval. <i>Revista Chilena de Historia Natural</i> , 74(3), 669-679.
241	RCHN9-2001 74(3)	Pavéz, E. (2001). Biología reproductiva del águila <i>Geranoaetus melanoleucus</i> (Aves: Accipitridae) en Chile central. <i>Revista Chilena de Historia Natural</i> , 74(3), 687-697.
242	RCHN10-2001 74(3)	Lusk, C., Donoso, C., Jiménez, M., Moya, C., Oyarce, G., Reinoso, R., Saldaña, A., Villegas, P. & Matus, F. (2001). Descomposición de hojarasca de <i>Pinus radiata</i> y tres especies arbóreas nativas. <i>Revista Chilena de Historia Natural</i> , 74(3), 705-710.
243	RCHN11-2001 74(3)	Moraga, M., Aspillaga, E., Santoro, C., Standen, V., Carvallo, P. & Rothhammer, F. (2001). Análisis de ADN mitocondrial en momias del norte de Chile avala hipótesis de origen amazónico de poblaciones andinas. <i>Revista Chilena de Historia Natural</i> , 74(4), 719-726.
244	RCHN1-2001 74(4)	Valenzuela, E., Leiva, S., & Godoy, R. (2001). Variación estacional y potencial enzimático de microhongos asociados con la descomposición de hojarasca de <i>Nothofagus pumilio</i> . <i>Revista Chilena de Historia Natural</i> , 74(4), 737-749.
245	RCHN2-2001 74(4)	Muñoz, G. (2001). Diferencias en las historias de vida de dos especies de isópodos bopíridos (Isopoda: Epicaridea) que parasitan al nape <i>Neotrypaea uncinata</i> (Milne-Edwards, 1837) (Decapoda: Thalassinidea). <i>Revista Chilena</i>

		<i>de Historia Natural</i> , 74(4), 755-761.
246	RCHN3-2001 74(4)	González, J., Pastenes, C. & Horton, P. (2001). Efecto de la temperatura, el estrés hídrico y luminoso sobre la heterogeneidad del fotosistema II en cuatro variedades de poroto (<i>Phaseolus vulgaris</i> L.). <i>Revista Chilena de Historia Natural</i> , 74(4), 779-791.
247	RCHN4-2001 74(4)	Villagrán, C. (2001). Un modelo de la historia de la vegetación de la Cordillera de La Costa de Chile central-sur: la hipótesis glacial de Darwin. <i>Revista Chilena de Historia Natural</i> , 74(4), 793-803.
248	RCHN5-2001 74(4)	Faúndez-Báez, P., Morales, C. & Arcos, D. (2001). Variabilidad espacial y temporal en la hidrografía invernal del sistema de bahías frente a la VIII región (Chile centro-sur). <i>Revista Chilena de Historia Natural</i> , 74(4), 793-803.
249	RCHN6-2001 74(4)	Garcías, F., Mendoza, R. & George-Nascimento, M. (2001). Variación entre años de las infracomunidades de parásitos metazoos de la corvina <i>Cilus gilberti</i> (Pisces: Sciaenidae) en Chile. <i>Revista Chilena de Historia Natural</i> , 74(4), 833-840.
250	RCHN7-2001 74(4)	Cerda, M. & Castilla, J. (2001). Diversidad y biomasa de macro-invertebrados en matrices intermareales del tunicado <i>Pyura praeputialis</i> (Heller, 1878) en la Bahía de Antofagasta, Chile. <i>Revista Chilena de Historia Natural</i> , 74(4), 841-853.
251	RCHN8-2001 74(4)	Rivadeneira, M. & Oliva, E. (2001). Patrones asociados a la conducta de desplazamiento local en <i>Phymactis clematis</i> Drayton (Anthozoa: Actiniidae). <i>Revista Chilena de Historia Natural</i> , 74(4), 855-863.
252	RCHN9-2001 74(4)	Morales, M., Villalba, R., Grau, H., Villagra, P., Boninsegna, J., Ripalta, A. & Paolini, L. (2001). Potencialidad de <i>Prosopis ferox</i> Griseb (Leguminosae subfamilia: Mimosoideae) para estudios dendrocronológicos en desiertos subtropicales de alta montaña. <i>Revista Chilena de Historia Natural</i> , 74(4),

			865-872.
253	RCHN10-2001 74(4)		Guerra-García, J. & Thiel, M. (2001). La fauna de caprélidos (Crustacea: Amphipoda: Caprellidea) de la costa de Coquimbo, centro-norte de Chile, con una clave taxonómica para la identificación de las especies. <i>Revista Chilena de Historia Natural</i> , 74(4), 873-883.
254	RCHN1- 75(1)	2002	Zenuto, R., Vassallo, A. & Busch, C. (2002). Comportamiento social y reproductivo del roedor subterráneo solitario <i>Ctenomys talarum</i> (Rodentia: Ctenomyidae) en condiciones de semicautiverio. <i>Revista Chilena de Historia Natural</i> , 75(1), 165-177.
255	RCHN2- 75(1)	2002	Braccini, J. & Chiaramonte, G. (2002). Biología de la raya <i>Psammobatis extenta</i> (Garman, 1913) (Batoidea: Rajidae). <i>Revista Chilena de Historia Natural</i> , 75(1), 179-188.
256	RCHN3- 75(1)	2002	Oyarzún, C., Godoy, R. & Leiva, S. (2002). Deposición atmosférica de nitrógeno en un transecto valle longitudinal-cordillera de Los Andes, centro-sur de Chile. <i>Revista Chilena de Historia Natural</i> , 75(1), 233-243.
257	RCHN4- 75(1)	2002	Cuevas, C. & Formas, J.R. (2002). <i>Telmatobius philippii</i> , una nueva especie de rana acuática de Ollagüe, norte de Chile (Leptodactylidae). <i>Revista Chilena de Historia Natural</i> , 75(1), 245-258.
258	RCHN1- 75(2)	2002	Canals, M., Atala, C., Olivares, R., Novoa, F. & Rosenmann, M. (2002). La asimetría y el grado de optimización del árbol bronquial en <i>Rattus norvegicus</i> y <i>Oryctolagus cuniculus</i> . <i>Revista Chilena de Historia Natural</i> , 75(2), 271-282.
259	RCHN2- 75(2)	2002	Castro, D. & Cicchino, A. (2002). Las especies del género <i>Gyropus</i> Nitzsch, 1818 (Phthiraptera: Gyropidae) parásitas de Octodontidae (Mammalia: Rodentia). <i>Revista Chilena de Historia Natural</i> , 75(2), 293-298.
260	RCHN3-	2002	Meyhnard, A., Palma, R. & Rivera-Milla, E. (2002). Filogeografía de las Llacas chilenas del género <i>Thylamys</i> (Marsupialia, Didelphidae) en base a

	75(2)		secuencias del gen mitocondrial citocromo b. <i>Revista Chilena de Historia Natural</i> , 75(2), 299-306.
261	RCHN4- 75(2)	2002	Márquez, S., Funes, G., Cabido, M. & Pucheta, E. (2002). Efectos del pastoreo sobre el banco de semillas germinable y la vegetación establecida en pastizales de montaña del centro de Argentina. <i>Revista Chilena de Historia Natural</i> , 75(2), 327-337.
262	RCHN5- 75(2)	2002	Cerrillo, M., Dana, E., Castro, H., Rodríguez-Tamayo, M. & Mota, J. (2002). Selección de áreas prioritarias para la conservación de flora gipsícola en el sureste de la Península Ibérica. <i>Revista Chilena de Historia Natural</i> , 75(2), 395-408.
263	RCHN6- 75(2)	2002	Neme, G., Moreira, G., Atencio, A. & De Santis, L. (2002). El registro de microvertebrados del sitio arqueológico Arroyo Malo 3 (Provincia de Mendoza, Argentina). <i>Revista Chilena de Historia Natural</i> , 75(2), 409-421.
264	RCHN7- 75(2)	2002	Díaz, I., Sarmiento, C., Ulloa, L., Moreira, R., Navia, R., Véliz, E. & Peña, C. (2002). Vertebrados terrestres de la Reserva Nacional Río Clarillo, Chile central: Representatividad y conservación. <i>Revista Chilena de Historia Natural</i> , 75(2), 433-448.
265	RCHN1- 75(3)	2002	Damascos, M. & Rapoport, E. (2002). Diferencias en la flora herbácea y arbustiva entre claros y áreas bajo dosel en un bosque de <i>Nothofagus pumilio</i> en Argentina. <i>Revista Chilena de Historia Natural</i> , 75(3), 465-472.
266	RCHN2- 75(3)	2002	Valladares, J., Etheridge, R., Schulte, J., Manriquez, G. & Spotorno, A. (2002). Nueva especie de lagartija del norte de Chile, <i>Liolaemus molinai</i> (Reptilia: Liolaeminae). <i>Revista Chilena de Historia Natural</i> , 75(3), 473-489.
267	RCHN3- 75(3)	2002	Romero, M. & Valdebenito, E. (2002). Larvas veliger de gastrópodos Prosobranchia provenientes de Punta de Lobos, Cuarta Región, Chile. <i>Revista Chilena de Historia Natural</i> , 75(3), 491-514.

268	RCHN4- 75(3)	2002	Astorga, M., Guíñez, R., Ortiz, J. & Castilla, J. (2002). Variación fenotípica y genética en el tunicado <i>Pyura praeputialis</i> (Heller, 1878) en el área norte de la Bahía de Antofagasta, Chile. <i>Revista Chilena de Historia Natural</i> , 75(3), 515-526.
269	RCHN5- 75(3)	2002	Betancourt, J. & Saavedra, B. (2002). Paleomadrigueras de roedores, un nuevo método para el estudio del Cuaternario en zonas áridas de Sudamérica. <i>Revista Chilena de Historia Natural</i> , 75(3), 527-546.
270	RCHN6- 75(3)	2002	Delgado, L. & Serey, I. (2002). Distribución del cobre en ecosistemas forestales de tipo mediterráneo. <i>Revista Chilena de Historia Natural</i> , 75(3), 557-565.
271	RCHN7- 75(3)	2002	Cortés, A., Rau, J., Miranda, E. & Jiménez, J. (2002). Hábitos alimenticios de <i>Lagidium viscacia</i> y <i>Abrocoma cinerea</i> : Roedores sintópicos en ambientes altoandinos del norte de Chile. <i>Revista Chilena de Historia Natural</i> , 75(3), 583-593.
272	RCHN8- 75(3)	2002	Pistone, E., Carezzano, F. & Bee-De-Speroni, N. (2002). Tamaño relativo encefálico e índices cerebrales en <i>Vanellus c. chilensis</i> (Aves: Charadriidae). <i>Revista Chilena de Historia Natural</i> , 75(3), 595-602.
273	RCHN9- 75(3)	2002	Ginocchio, R. & Narváez, J. (2002). Importancia de la forma química y de la matriz del sustrato en la toxicidad por cobre en <i>Noticastrum sericeum</i> (Less.) Less. ex Phil. <i>Revista Chilena de Historia Natural</i> , 75(3), 603-612.
274	RCHN10- 75(3)	2002	Nai-Bregaglio, M., Pucheta, E. & Cabido, M. (2002). El efecto del pastoreo sobre la diversidad florística y estructural en pastizales de montaña del centro de Argentina. <i>Revista Chilena de Historia Natural</i> , 75(3), 613-623.
275	RCHN1- 75(4)	2002	Díaz, F. & George-Nascimento, M. (2002). Estabilidad temporal de las infracomunidades de parásitos en la borrachilla <i>Scartichthys viridis</i> (Valenciennes, 1836) (Pisces : Blenniidae) en la costa central de Chile. <i>Revista Chilena de Historia Natural</i> , 75(4), 625-635.

			<i>Chilena de Historia Natural</i> , 75(4), 641-649.
276	RCHN2- 75(4)	2002	Maldonado, C., Pujado, E. & Squeo, F. (2002). El efecto de la disponibilidad de agua durante el crecimiento de <i>Lycopersicon chilense</i> sobre la capacidad de sus semillas para germinar a distintas temperaturas y concentraciones de manitol y NaCl. <i>Revista Chilena de Historia Natural</i> , 75(4), 651-660.
277	RCHN3- 75(4)	2002	Muñoz, G., Valdebenito, V. & George-Nascimento, M. (2002). La dieta y la fauna de parásitos metazoos del torito <i>Bovichthys chilensis</i> Regan 1914 (Pisces: Bovichthyidae) en la costa de Chile centro-sur: variaciones geográficas y ontogenéticas. <i>Revista Chilena de Historia Natural</i> , 75(4), 661-671.
278	RCHN4- 75(4)	2002	Muñoz-Pedrerros, A. & Larraín, A. (2002). Impacto de la actividad silvoagropecuaria sobre la calidad del paisaje en un transecto del sur de Chile. <i>Revista Chilena de Historia Natural</i> , 75(4), 673-689.
279	RCHN5- 75(4)	2002	Parada, E. & Peredo, S. (2002). Estado actual de la taxonomía de bivalvos dulceacuícolas chilenos: Progresos y conflictos. <i>Revista Chilena de Historia Natural</i> , 75(4), 691-701.
280	RCHN6- 75(4)	2002	Renison, D., Cingolani, A. & Suarez, R. (2002). Efectos del fuego sobre un bosquecillo de <i>Polylepis australis</i> (Rosaceae) en las montañas de Córdoba, Argentina. <i>Revista Chilena de Historia Natural</i> , 75(4), 719-727.
281	RCHN7- 75(4)	2002	Torrejón, F. & Cisternas, M. (2002). Alteraciones del paisaje ecológico araucano por la asimilación mapuche de la agroganadería hispano-mediterránea (siglos XVI y XVII). <i>Revista Chilena de Historia Natural</i> , 75(4), 729-736.
282	RCHN8- 75(4)	2002	Torres, R., Squeo, F., Jonquera, C., Aguirre, E. & Ehleringer, J. (2002). Evaluación de la capacidad estacional de utilizar eventos de precipitación en tres especies de arbustos nativos de Chile con distintos sistemas radiculares. <i>Revista Chilena de Historia Natural</i> , 75(4), 737-749.

283	RCHN9- 75(4)	2002	Camus, P. & Barahona, R. (2002). Insectos del intermareal de (Concepción) Chile: Perspectivas para la investigación ecológica. <i>Revista Chilena de Historia Natural</i> , 75(4), 793-803.
284	RCHN10- 75(4)	2002	Badano, E., Molina-Montenegro, M., Quiroz, C. & Caviares, L. (2002). Efectos de la planta en cojín <i>Oreopolus glacialis</i> (Rubiaceae) sobre la riqueza y diversidad de especies en una comunidad alto-andina de Chile central. <i>Revista Chilena de Historia Natural</i> , 75(4), 757-765.
285	RCHN11- 75(4)	2002	Gerding, V. & Thiers, O. (2002). Caracterización de suelos bajo bosques de <i>Nothofagus betuloides</i> (Mirb) Blume, en Tierra del Fuego, Chile. <i>Revista Chilena de Historia Natural</i> , 75(4), 819-833.
286	RCHN1- 76(1)	2003	Habit, E., Victoriano, P. & Rodríguez-Ruiz, A. (2003). Variaciones espacio-temporales del ensamble de peces de un sistema fluvial de bajo orden del centro-sur de Chile. <i>Revista Chilena de Historia Natural</i> , 76(1), 3-14.
287	RCHN2- 76(1)	2003	Dorador, C., Pardo, R. & Vila, I. (2003). Variaciones temporales de parámetros físicos, químicos y biológicos de un lago de altura: El caso del lago Chungará. <i>Revista Chilena de Historia Natural</i> , 76(1), 15-22.
288	RCHN3- 76(1)	2003	Maestre, F. (2003). Variaciones en el patrón espacial a pequeña escala de los componentes de la costra biológica en un ecosistema mediterráneo semiárido. <i>Revista Chilena de Historia Natural</i> , 76(1), 35-46.
289	RCHN4- 76(1)	2003	Victoriano, P., Torres-Pérez, F., Ortiz, J., Parra, L., Northland, I. & Capetillo, J. (2003). Variación aloenzimática y parentesco evolutivo en especies de <i>Microlophus</i> del grupo "peruvianus" (Squamata: Tropiduridae). <i>Revista Chilena de Historia Natural</i> , 76(1), 65-78.
290	RCHN5- 76(1)	2003	Norambuena, C. & Paredes, M. (2003). Variabilidad y estructura genética en dos poblaciones de <i>Vicugna vicugna</i> (Camelidae) del norte de Chile. <i>Revista Chilena de Historia Natural</i> , 76(1), 99-104.

291	RCHN6- 76(1)	2003	Kunst, C., Bravo, S., Moscovich, F., Herrera, J., Godoy, J. & Vélez, S. (2003). Fecha de aplicación de fuego y diversidad de herbáceas en una sabana de Elionorus muticus (Spreng) O. Kuntze. <i>Revista Chilena de Historia Natural</i> , 76(1), 105-115.
292	RCHN7- 76(1)	2003	Parra, L. & Pascual-Toca, M. (2003). Revisión taxonómica de los géneros Oratha Walker 1863 y Hasodima Butler 1882 (Lepidoptera: Geometridae). <i>Revista Chilena de Historia Natural</i> , 76(1), 117-128.
293	RCHN1- 76(2)	2003	Ortiz, M., Silva, H., Silva, P. & Acevedo, E. (2003). Estudio de parámetros hídricos foliares en trigo (<i>Triticum aestivum</i> L.) y su uso en selección de genotipos resistentes a sequía. <i>Revista Chilena de Historia Natural</i> 76(2), 219-233.
294	RCHN2- 76(2)	2003	Martínez, G., Contreras, M. & Vila, I. (2003). Modelos conceptuales de abundancia de fitoplancton asociados a la heterogeneidad espacial en el Embalse Rapel (Chile central). <i>Revista Chilena de Historia Natural</i> 76(2), 255-266.
295	RCHN3- 76(2)	2003	Figueroa, R., Valdovinos, C., Araya, E. & Parra, O. (2003). Macroinvertebrados bentónicos como indicadores de calidad de agua de ríos del sur de Chile. <i>Revista Chilena de Historia Natural</i> 76(2), 275-285.
296	RCHN4- 76(2)	2003	Hengst, M. & Artigas, J. (2003). Revisión del género <i>Philonerax</i> Bromley, 1932 (Diptera: Asilidae: Asilinae) y dos nuevas especies. <i>Revista Chilena de Historia Natural</i> 76(2), 287-296.

CIENCIAS SOCIALES

ID	Código	Referencia Bibliográfica
----	--------	--------------------------

1	AMB1-2000 5	López, X. (2000). La comunicación del futuro se escribe con L de local. <i>Ámbitos</i> , 5, 107-117.
2	AMB2-2000 5	Román, M. (2000). Aspectos metodológicos de la historia de la comunicación. <i>Ámbitos</i> , 5, 119-128.
3	AMB3-2000 5	Sandoval, M. (2000). La Red: El nuevo mercado de la vida privada. <i>Ámbitos</i> , 5, 143-155.
4	AMB4-2000 5	Fernández, M. (2000). Luis Buñuel: Trascendiendo el tópico. <i>Ámbitos</i> , 5, 157-168.
5	AMB5-2000 5	Espinosa, P. (2000). Los géneros periodísticos informativos en la actualidad Internacional. <i>Ámbitos</i> , 5, 169-190.
6	AMB6-2000 5	Vela, J. (2000). Los estudios de la audiencia: Una visión crítica desde la economía política. <i>Ámbitos</i> , 5, 191-210.
7	AMB7-2000 5	Sierra, F. (2000). Horizontes y problemas del audiovisual andaluz. <i>Ámbitos</i> , 5, 213-224.
8	AMB8-2000 5	Bernal, A. (2000). Joly: Un grupo de comunicación andaluz dentro de la estructura de la prensa en España. <i>Ámbitos</i> , 5, 225-240.
9	AMB9-2000 5	Marín, J. (2000). La crónica deportiva: José A. Sánchez Araujo. <i>Ámbitos</i> , 5, 241-257.
10	AMB10-2000 5	Reig, R., García, M., Moreno, L. & Luque, J. (2000). Investigación sobre la comunicación en Sevilla. Once años de la Facultad de Ciencias de la Información (1989-2000). <i>Ámbitos</i> , 5, 259-311.
11	AMB1-2001 6	Emanuelli, P. (2001). Dominante cultural y productos televisivos: Géneros que homogeneizan preferencias. <i>Ámbitos</i> , 6, 7-20.

12	AMB2-2001 6	Esteinou, J. (2001). Internet y la transformación del Estado. <i>Ámbitos</i> , 6, 21-34.
13	AMB3-2001 6	Fandiño, X. (2001). De Analogía a Digitalia. <i>Ámbitos</i> , 6, 35-50.
14	AMB4-2001 6	Edo, C. (2001). Las incertidumbres del periodismo en Internet. <i>Ámbitos</i> , 6, 51-62.
15	AMB5-2001 6	García de Torres, E. (2001). El medio-portal y los canales de actualidad. <i>Ámbitos</i> , 6, 63-80.
16	AMB6-2001 6	Labio, A. (2001). La profesión periodística ante los intereses de la estructura informativa. Una aproximación al tema. <i>Ámbitos</i> , 6, 81-94.
17	AMB7-2001 6	López, A. (2001). La “historia de vida” periodística, un género poco usual en la prensa española. <i>Ámbitos</i> , 6, 95-106.
18	AMB8-2001 6	Moreno, P. (2001). Géneros para la persuasión en prensa: Los artículos de opinión del diario El País. <i>Ámbitos</i> , 6, 107-121.
19	AMB9-2001 6	Santos, M.T. (2001). Los periódicos gratuitos con contenidos especializados en la Comunidad Autónoma Vasca. <i>Ámbitos</i> , 6, 123-142.
20	AMB10-2001 6	Troncoso, J. (2001). Enfatemática del antiespañolismo en los textos de historia en países europeos y americanos. <i>Ámbitos</i> , 6, 143-169.
21	AMB11-2001 6	Barrero, A. (2001). Juicios paralelos y Constitución: Su relación con el Periodismo. <i>Ámbitos</i> , 6, 171-189.
22	AMB12-2001 6	Pineda, A. (2001). El modelo de propaganda de Noam Chomsky: Medios mainstream y control del pensamiento. <i>Ámbitos</i> , 6, 191-210.
23	AMB13-2001 6	Otero, M. (2001). A propósito de la telefonía móvil: Una reflexión desde la perspectiva de la psicología individual y social. <i>Ámbitos</i> , 6, 211-235.

24	AMB14-2001 6	Egea, C. (2001). La carrera por la comunicación local (1998-2000) “Los grandes” se atreven con “lo pequeño”. <i>Ámbitos</i> , 6, 237-260.
25	AMB15-2001 6	Marín, J. (2001). La cobertura televisiva del fútbol en el sur de Brasil y de España: RBS – Canal Sur. <i>Ámbitos</i> , 6, 263-278.
26	AMB16-2001 6	Elías, C. (2001). Periodismo especializado en medio ambiente: El caso Doñana como paradigma de manipulación informativa. <i>Ámbitos</i> , 6, 279-303.
27	AMB17-2001 6	González, J. (2001). Las relaciones históricas como precedentes de periodismo granadino. <i>Ámbitos</i> , 6, 305-320.
28	AMB1-2002 7-8	López, X. (2002). Nuevos perfiles de los periodistas en la sociedad de la información. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti01.htm
29	AMB2-2002 7-8	León, G. (2002). Teorías e investigación de la comunicación en América Latina: Situación actual. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti02.htm
30	AMB3-2002 7-8	Cisneros, J. (2002). El concepto de la comunicación: El cristal con que se mira. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti03/arti3.htm
31	AMB4-2002 7-8	De Pablos, J.M. (2002). Mantengamos la utopía: La Red es nuestra (1). <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti04.htm
32	AMB5-2002 7-8	Pestano, J.M. (2002). Las tecnologías emergentes como soportes de conocimiento: El libro digital (1). <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti05.htm

33	AMB6-2002 7-8	Fernández, M.A. (2002). Peculiaridades del editorial ante acontecimientos que marcan la historia: El caso de los editoriales del 12 de septiembre de 2001. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti06/arti6.htm
34	AMB7-2002 7-8	García, M.J. (2002). Mecanismos básicos de la propaganda de guerra en los medios informativos. El ejemplo de Kosovo. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/7-8/arti07.htm
35	AMB8-2002 7-8	Fernández, S. (2002). La globalización de la comunicación. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/7-8/arti08.htm
36	AMB9-2002 7-8	Sánchez, M. (2002). Importancia de la gestión de la comunicación en periodos de crisis empresariales. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti09/arti9.htm
37	AMB10-2002 7-8	Del Valle, C. (2002). La estructura argumentativa de un tipo particular de discurso jurídico: El caso de cuatro demandas de reivindicación territorial mapuche (IX Región – Chile) y la construcción de la identidad reivindicativa. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti10/arti10.htm
38	AMB11-2002 7-8	Ramos, F. (2002). Razones históricas de la imagen del ejército ante la sociedad española (desde la guerra de Cuba a nuestros días). <i>Ámbitos</i> , 7-8, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/7-8/arti11.htm
39	AMB12-2002 7-8	Díaz, J. (2002). Gacetas españolas de los Países Bajos en el siglo XVII: La Gazeta de Amsterdam y Noticias Principales y Verdaderas. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/7-8/arti12/arti12.htm

40	AMB13-2002 7-8	Gila, J. & Rodríguez, F. (2002). El periodista científico y el psicólogo de la comunicación: Su papel en la transmisión de mensajes científicos a través de los medios de comunicación de masas. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/7-8/arti13.htm
41	AMB14-2002 7-8	Perujo, F. (2002). Discapacidad y medios de comunicación: Entre la información y el estereotipo. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti14.htm
42	AMB15-2002 7-8	Ruiz, M. (2002). Prensa política en la Transición. Andalucía según El Socialista (1977-1982). <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti15.htm
43	AMB16-2002 7-8	Pérez, C. (2002). La actualidad informativa del “corazón” desde la especialización periodística: El periodismo rosa de Contraportada (Canal Sur Televisión). <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti16.htm
44	AMB17-2002 7-8	Guerrero, R. (2002). La prensa en la Segunda República: Breve aproximación como contexto vital de don Diego Martínez Barrio. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/7-8/arti17.htm
45	AMB18-2002 7-8	Sarmiento, R. (2002). Notas sobre psicoanálisis, literatura y comunicación. Algunas consideraciones a propósito del amor en la novela “Cuaderno de viaje” de Salvador Compán (1). <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti18.htm
46	AMB19-2002 7-8	Gómez, A. (2002). Aproximación al estudio de la prensa local de barrio en Sevilla. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti19/arti19.htm

47	AMB20-2002 7-8	Márquez, M. (2002). Miguel A. Yáñez Polo, fotosurrealista metafísico. <i>Ámbitos</i> , 7-8, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/7-8/arti20/arti20.htm
48	AMB1-2003 9-10	Núñez, S. (2003). Minorías nacionales y medios de comunicación: Una visión de Europa. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art01.htm
49	AMB2-2003 9-10	Del Valle, R. (2003). De Telefónica Media a Admira Media. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art02.htm
50	AMB3-2003 9-10	López, X. (2003). La Comunicación en Galicia 2000-2003: Realidades y Sueños. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art03.htm
51	AMB4-2003 9-10	Del Valle, C. (2003). Interculturalidad y justicia en Chile. De la escritura como violencia simbólica a la oralidad como economía. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art04.htm
52	AMB5-2003 9-10	Moreno, C. (2003). La investigación universitaria en periodismo científico. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art06.htm
53	AMB6-2003 9-10	Fernández, R. (2003). En torno al debate sobre la definición del periodismo ambiental. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art07.htm
54	AMB7-2003 9-10	Cortés, J. (2003). La educación para los medios en la oferta educativa profesional en el norte de México. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art08.htm

55	AMB8-2003 9-10	Elías, C. (2003). Los suplementos especializados como guetos de noticias en la prensa generalista: El caso de los científicos y sanitarios. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art09.htm
56	AMB9-2003 9-10	Ronda, J. (2003). El periodismo judicial en España. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/9/art10.htm
57	AMB10-2003 9-10	López, A. (2003). El análisis: ¿Un género periodístico? <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art11.htm
58	AMB11-2003 9-10	Moreno, P. (2003). Géneros para la persuasión en prensa: Los editoriales de diario El País. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en:
59	AMB12-2003 9-10	Alberdi, A., Armentia, J.I., Caminos, J.M. & Marín, F. (2003). La remodelación de El Periódico de Catalunya: Hacia el modelo de prensa de servicios. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art14.htm
60	AMB13-2003 9-10	Edo, C. (2003). Los medios digitales todavía no pueden con el papel. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art15.htm
61	AMB14-2003 9-10	Márquez, M. (2003). Restauración digital de la fotografía: Un concepto erróneo. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art16.htm
62	AMB15-2003 9-10	Laguna, A. (2003). Política y televisión: Las perversiones de la democracia. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art17.htm

63	AMB16-2003 9-10	Cebrián, M. (2003). Contenidos cinematográficos en televisión. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art18.htm
64	AMB17-2003 9-10	Rúas, X. (2003). Televisión pública para todos los públicos. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art19.htm
65	AMB18-2003 9-10	Santos, M. T. (2003). Las emisoras municipales en el País Vasco: Origen y desarrollo. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art20.htm
66	AMB19-2003 9-10	Sabés, F. (2003). La radio aragonesa y el fracaso del Plan de Adjudicación de Frecuencias de 1998. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art21.htm
67	AMB20-2003 9-10	Amar, V. (2003). Poesía y cine: Un ensayo sobre la palabra y la imagen. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art22.htm
68	AMB21-2003 9-10	Fernández, C. (2003). Procedimientos comunes en la comunicación política y en la comunicación empresarial. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art23.htm
69	AMB22-2003 9-10	Sánchez, M.L. (2003). Estructura del plan de crisis. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en: http://www.ull.es/publicaciones/latina/ambitos/9/art24.htm
70	AMB23-2003 9-10	Holgado, M. (2003). El papel de los medios de comunicación en la campaña electoral. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art25.htm

71	AMB24-2003 9-10	Rodríguez, R. (2003). Características socio-demográficas y laborales de los periodistas españoles e índice de satisfacción profesional. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art26.htm
72	AMB25-2003 9-10	Ruiz, M. (2003). Censura y consignas en la prensa franquista. Algunos ejemplos de dirigismo informativo. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art27.htm
73	AMB26-2003 9-10	Pérez, R. (2003). La construcción social de la emisión ideológica. El caso de la literatura espiritual en la España del Renacimiento. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art28.htm
74	AMB27-2003 9-10	Almiron, N. (2003). Sobre el progreso en una era de revolución científico-tecnológico-digital. <i>Ámbitos</i> , 9-10, [en línea]. Disponible en http://www.ull.es/publicaciones/latina/ambitos/9/art29.htm
75	AD1-2000 3	Arévalo, J. (2000). Características del comportamiento del mercado de trabajo en biblioteconomía, archivística y documentación. Bienio 98/99. <i>Anales de Documentación</i> , 3, 9-24.
76	AD2-2000 3	Anglada, L. (2000). ¿Mejor, peor o solo distinto? <i>Anales de Documentación</i> , 3, 25-39.
77	AD3-2000 3	Fernández, P. (2000). Contexto pragmático, géneros y comprensión lectora de resúmenes científicos en inglés. <i>Anales de Documentación</i> , 3, 41-53.
78	AD4-2000 3	Font, E. (2000). Gestión de la información en la utilización del proceso analítico jerárquico para la toma de decisiones de nuevos productos. <i>Anales de Documentación</i> , 3, 55-66.

79	AD5-2000 3	Licea, J. (2000). La formación de comunidades científicas en bibliotecología y ciencia de la información. <i>Anales de Documentación</i> , 3, 81-91.
80	AD6-2000 3	Merlo, J. (2000). El servicio bibliotecario de referencia. <i>Anales de documentación</i> , 3, 93-126.
81	AD7-2000 3	Rey, C. (2000). La satisfacción del usuario: Un concepto en alza. <i>Anales de documentación</i> , 3, 139-153.
82	AD8-2000 3	Sánchez, B. (2000). La información como recurso en el desarrollo de las organizaciones de las administraciones públicas. <i>Anales de documentación</i> , 3, 155-165.
83	AD9-2000 3	Vivas, A. (2000). Documentación del archivo universitario de salamanca. Análisis descriptivo de sus series documentales. <i>Anales de Documentación</i> , 3, 167-207.
84	AD1-2001 4	Bereijo, A. & Fuentes, J. (2001). Los soportes fílmicos, magnéticos y ópticos desde la perspectiva de la conservación de materiales. <i>Anales de Documentación</i> , 4, 7-37.
85	AD2-2001 4	Caldera, J. & Zapico, F. (2001). Principales fuentes de información audiovisual en las televisiones estatales. <i>Anales de Documentación</i> , 4, 39-50.
86	AD3-2001 4	Delgado, E. (2001). ¿Por qué enseñar métodos de investigación en las facultades de biblioteconomía y documentación? <i>Anales de Documentación</i> , 4, 51-71.
87	AD4-2001 4	Feria, L. & Rosario, M. (2001). La contribución de la Universidad de Colima al desarrollo de la tecnología óptica en México: El Cenedic y el Cepromed. <i>Anales de Documentación</i> , 4, 73-94.
88	AD5-2001 4	García, A. (2001). La compañía de mercaderes de libros de la corte a mediados del Siglo XVIII. <i>Anales de Documentación</i> , 4, 95-126.

89	AD6-2001 4	Garrido, M. (2001). Procedimientos automáticos de creación y transformación de los registros bibliográficos. <i>Anales de Documentación</i> , 4, 127-137.
90	AD7-2001 4	Licea, J., Valles, J. & Izáosla, J. (2001). Agenda científica y extracientífica en VIH/Sida. <i>Anales de Documentación</i> , 4, 139-149.
91	AD8-2001 4	Moreira, J. (2001). Licenciatura en documentación: Experiencias, desarrollo y problemática. El caso de la Universidad Carlos III de Madrid. <i>Anales de Documentación</i> , 4, 151-168.
92	AD9-2001 4	Morillo, M., Camarero, M. & Azorín, V. Proyectos interdisciplinares de documentación gráfica y multimedia. La función del documentalista. <i>Anales de Documentación</i> , 4, 169-178.
93	AD10-2001 4	Paniagua, E. (2001). La creatividad y las tecnologías de la información y las comunicaciones. <i>Anales de Documentación</i> , 4, 179-191.
94	AD11-2001 4	Pérez, M. (2001). Prácticas de lectura en prisión: Estudio de actitudes y comportamiento de los reclusos en el Centro Penitenciario de Badajoz. <i>Anales de Documentación</i> , 4, 193-213.
95	AD12-2001 4	Sánchez, M. & Saorín, T. (2001). Las comunidades virtuales y los portales como escenarios de gestión documental y difusión de información. <i>Anales de Documentación</i> , 4, 215-227.
96	AD13-2001 4	Schmolling, R. (2001). Las bibliotecas virtuales en Alemania: Conexiones entre el saber, la información y la técnica. <i>Anales de Documentación</i> , 4, 229-242.
97	AD1-2002 5	Asensi, V. & Parra, A. (2002). El método científico y la nueva filosofía de la ciencia. <i>Anales de Documentación</i> , 5, 9-19.

98	AD2-2002 5	Barnes, P. (2002). La biblioteca escolar como recurso central del aprendizaje. Propuesta de formación en el centro educativo. <i>Anales de Documentación</i> , 5, 21-49.
99	AD3-2002 5	Baró, M., Mañá, T. & Comalat, M. (2002). Las bibliotecas de los centros públicos de educación secundaria en la ciudad de Barcelona. <i>Anales de Documentación</i> , 5, 51-79.
100	AD4-2002 5	Fernández, C. (2002). La labor educadora de Cisneros y la primera biblioteca del Renacimiento en España. <i>Anales de Documentación</i> , 5, 81-97.
101	AD5-2002 5	Gil, I. (2002). Consistencia en la indización de documentos entre indizadores noveles. <i>Anales de Documentación</i> , 5, 99-111.
102	AD6-2002 5	Glazier, J. (2002). Propuesta de un modelo de las relaciones de los paradigmas en una disciplina emergente. <i>Anales de Documentación</i> , 5, 521.
103	AD7-2002 5	Gómez, J. (2002). Los problemas de las bibliotecas escolares de la región de Murcia en un contexto de crisis del sistema educativo. <i>Anales de Documentación</i> , 5, 125-156.
104	AD8-2002 5	González, J. & Rodríguez, J. (2002). La tecnología de flujo de trabajo en el contexto de la biblioteca digital. <i>Anales de Documentación</i> , 5, 157-175.
105	AD9-2002 5	Guardiola, P. & Hernández, M. (2002). La imagen social de las bibliotecas en la prensa digital y escrita. <i>Anales de Documentación</i> , 5, 177-196.
106	AD10-2002 5	Gutiérrez, M. (2002). El aprendizaje de la ciencia y de la información científica en la educación superior. <i>Anales de Documentación</i> , 5, 197-212.
107	AD11-2002 5	Licea, J., Arenas, M. & Valles, J. (2002). La píldora anticonceptiva en América Latina y El Caribe. Estudio bibliométrico. <i>Anales de Documentación</i> , 5, 213-222.

108	AD12-2002 5	López, I. (2002). Luis Méndez Albarrán y la CDU en España. <i>Anales de Documentación</i> , 5, 223-234.
109	AD13-2002 5	Marcos, M. (2002). Citar documentos electrónicos. Revisión de propuestas y planteamiento de pautas generales. <i>Anales de Documentación</i> , 5, 235-243.
110	AD14-2002 5	Moreira, J. (2002). Aplicaciones al análisis automático del contenido provenientes de la teoría matemática de la información. <i>Anales de Documentación</i> , 5, 273-286.
111	AD15-2002 5	Rodríguez, C. (2002). La biblioteca pública: Un servicio ligado a un proyecto de ciudad. <i>Anales de Documentación</i> , 5, 303-308.
112	AD16-2002 5	Salaberría, R. (2002). Las bibliotecas populares en la correspondencia de Juan Vicéns a Lulu Jourdain y Hernando Viñes (1933-1936). <i>Anales de Documentación</i> , 5, 309-332.
113	AD17-2002 5	Suaiden, E. (2002). El impacto social de las bibliotecas públicas. <i>Anales de Documentación</i> , 5, 333-344.
114	AD18-2002 5	Viñao, A. (2002). La enseñanza de la lectura y la escritura: Análisis socio-histórico. <i>Anales de Documentación</i> , 5, 345-359.
115	AD1-2003 6	Ayuso, M., & Martínez, V. (2003). La edición digital del boletín oficial de la región de Murcia: Parámetros e indicadores para su evaluación. <i>Anales de Documentación</i> , 6, 7-26.
116	AD2-2003 6	Bueno, D., Silva, J., & Olivera, G. (2003). Hongos ambientales en una biblioteca: Un año de estudio. <i>Anales de Documentación</i> , 6, 27-34.
117	AD3-2003 6	Fernández, I. (2003). Tabularivm: El archivo en época romana. <i>Anales de Documentación</i> , 6, 59-70.

118	AD4-2003 6	Fuentes, J. (2003). El concepto de biblioteca nacional a partir de los tres informes de la Unesco sobre las bibliotecas nacionales: Sylvestre (1987), Lin (1989) y Cornish (1991). <i>Anales de Documentación</i> , 6, 71-88.
119	AD5-2003 6	Gonzalez, E., Arenas, M. & Licea, J. (2003). Estudio bibliométrico de la actividad científica de los matemáticos mexicanos graduados en Estados Unidos en el periodo 1980-1998. <i>Anales de Documentación</i> , 6, 89-108.
120	AD6-2003 6	Gómez, C. (2003). Las nuevas formas de comunicación de la administración con el ciudadano. <i>Anales de Documentación</i> , 6, 109-119.
121	AD7-2003 6	Jaén, L. (2003). Los modelos teóricos de unificación de archivos. <i>Anales de Documentación</i> , 6, 121-135.
122	AD8-2003 6	Licea, J., Sandoval, M. & Arenas, M. La investigación agrícola en México. Un estudio bibliométrico con enfoque de género. <i>Anales de Documentación</i> , 6, 145-154.
123	AD9-2003 6	Martín, J.C., Blackwell, S. & Merlo, J. (2003). Las revistas electrónicas. Características, fuentes de información y medios de acceso. <i>Anales de Documentación</i> , 6, 155-186.
124	AD10-2003 6	Moya, G. (2003). La administración pública en línea: Estudio sobre los servicios públicos de los ayuntamientos de la región de Murcia basados en web. <i>Anales de Documentación</i> , 6, 187-202.
125	AD11-2003 6	Omella, E. (2003). Biblioteca pública y servicio de información a la comunidad: Propuestas para una mayor integración. <i>Anales de Documentación</i> , 6, 203-220.
126	AD12-2003 6	Pedraza, J.M. (2003). Algunas reflexiones sobre la tasación del libro antiguo como actividad documental. <i>Anales de Documentación</i> , 6, 221-239.

127	AD13-2003 6	Peset, M.F. (2003). Bibliotecas digitales en internet de libro raro, antiguo e incunables. <i>Anales de Documentación</i> , 6, 241-260.
128	AD14-2003 6	Prat, J. (2003). Las corrientes ideológicas. Pasado, presente y futuro de la cultura impresa. <i>Anales de Documentación</i> , 6, 261-265.
129	CHU1-2000 32(1)	Quevedo, S. (2000). Patrones de actividad a través de las patologías en la población arcaica de Punta Teatinos, norte semiárido chileno. <i>Chungará, Revista de Antropología Chilena</i> , 32(1), 7-9.
130	CHU2-2000 32(1)	Rodríguez, C. (2000). Estudio demográfico de la población Guanche de Tenerife. <i>Chungará, Revista de Antropología Chilena</i> 32(1), 27-32.
131	CHU3-2000 32(1)	Etxeberria, F., Romero, W. & Herraste, L. (2000). Cifosis angular de la columna vertebral: Identificación del mal de Pott en una momia Guanche prehispánica de Colombia. <i>Chungará, Revista de Antropología Chilena</i> 32(1), 41-48.
132	CHU4-2000 32(1)	Kauffmann, F. (2000). Primera expedición arqueológica a los mausoleos Chachapoya de la laguna de las momias (Dpto. de Amazonas, Perú). <i>Chungará, Revista de Antropología Chilena</i> , 32(1), 49-54.
133	CHU5-2000 32(1)	Costa, M., Cocilovo, J. & Quevedo, S. (2000). Patologías óseas, traumas y otros atributos en el grupo arcaico del Morro de Arica, norte de Chile. <i>Chungará, Revista de Antropología Chilena</i> , 32(1), 79-83.
134	CHU6-2000 32(1)	Rodríguez, C. (2000). Manifestaciones esqueléticas de las enfermedades parasitarias. <i>Chungará, Revista de Antropología Chilena</i> , 32(1), 117-121.
135	CHU1-2000 32(2)	Standen, V. & Arriaza, B. (2000). La treponematosi (yaws) en las poblaciones prehispánicas del desierto de Atacama (norte de Chile). <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 185-192.

136	CHU2-2000 32(2)	Malvido, E. (2000). La guerra contra las momias en nueva España. El Siglo XVIII, jesuitas, franciscanos, autoridades seculares e inquisición. <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 199-205.
137	CHU3-2000 32(2)	Mansilla, J. & Pijoan, M. (2000). Evidencia de treponematosi en la Cueva de la Candelaria, Coahuila, con énfasis en un bulto mortuorio infantil. <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 207-210.
138	CHU4-2000 32(2)	Mansilla, J. & Pijoan, M. (2000). La Cueva de la Candelaria: Bultos mortuorios y materiales. <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 211-215.
139	CHU5-2000 32(2)	Agüero, C. (2000). Las tradiciones de tierras altas y de valles occidentales en la textilería arqueológica del Valle de Azapa. <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 217-225.
140	CHU6-2000 32(2)	Rojas, A. & Hoces, S. (2000). Coexistencia y diversidad técnica, textural y formal en los textiles de un fardo perteneciente al sitio Coyo de San Pedro de Atacama. <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 227-233.
141	CHU7-2000 32(2)	Horta, H. (2000). Diademas de plumas en entierros de la costa del norte de Chile: ¿Evidencias de la vestimenta de una posible parcialidad pescadora? <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 235-243.
142	CHU8-2000 32(2)	Ulloa, L., Standen, V. & Gavilán, V. (2000). Estudio de una prenda textil asociada al Inca en la costa norte de Chile (camarones 9): Las "mantas" que envuelven los cuerpos. <i>Chungará, Revista de Antropología Chilena</i> , 32(2), 259-261.
143	CHU1-2001 33(1)	Cocilovo, J., Varela, H., Espouey, O. & Standen, V. (2001). El proceso microevolutivo de la población nativa antigua de Arica. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 13-20.

144	CHU2-2001 33(1)	Mena, F. & Reyes, O. (2001). Montículos y cuevas funerarias en Patagonia. Una visión desde Cueva Baño Nuevo-1, XI región. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 21-30.
145	CHU3-2001 33(1)	Correa, G. (2001). Patrones mortuorios en cazadores recolectores de pleistoceno y holoceno en Colombia. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 37-42.
146	CHU4-2001 33(1)	Santoro, C., Standen, V & Arriaza, B. (2001). ¿Patrón funerario arcaico o alteración postdeposicional? El enterratorio de Patapatane en los Andes centro-sur. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 43-49.
147	CHU5-2001 33(1)	Chacama, J. & Muñoz, I. (2001). Patrón funerario pre Chinchorro en un contexto de semi sedentarismo y complementariedad ecológica. El sitio Acha 2, extremo norte de Chile Ca. 9.500-10.000 años A.P. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 51-54.
148	CHU6-2001 33(1)	Costa, M. (2001). Modalidades de enterramientos humanos arcaicos en el norte de Chile. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 55-62.
149	CHU7-2001 33(1)	Bustos, V. & Vergara, N. (2001). Evolución de las prácticas mortuorias en el litoral de la octava región. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 73-78.
150	CHU8-2001 33(1)	Yacobaccio, H., Madero, C. & Reigadas, M. (2001). Inhumación de una cabeza aislada en la puna argentina. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 79-82.
151	CHU9-2001 33 (1)	Córdova, J. & Bernal, J. (2001). Fascinación por las momias, reforzamiento de la vida. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 91-93.

152	CHU10-2001 33(1)	Cases, B. & Rojas, A. (2001). Un planteamiento experimental de replicación Registro y conservación de una momia Chinchorro tardía. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 107-111.
153	CHU11-2001 33(1)	Belmonte, E., Ortega, M., Arévalo, P., Cassman, V. & Cartmell, L. (2001) Presencia de la hoja de coca en el ajuar funerario de tres cementerios de periodo Tiwanaku: az-140, az-6 y plm-31. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 125-135
154	CHU12-2001 33(1)	Pérez de Micou, C. (2001). Cestería y cordelería para los muertos. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 137-144
155	CHU13-2001 33(1)	Belmonte, E., Bastías, E., Gómez, M., Mujica, A. & Montenegro, G. Determinación taxonómica de fragmentos de madera de contexto funerario de la cultura Chinchorro. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 145-154.
156	CHU14-2001 33(1)	Muñoz, I. (2001). Uso de plantas en rituales funerarios del período formativo en Arica. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 155-160.
157	CHU15-2001 33(1)	Del Casal, M. (2001). Plantas para la eternidad. <i>Chungará, Revista de Antropología Chilena</i> , 33(1), 161-168.
158	CHU1-2001 33(2)	Aláez, A. (2001). Duelo andino: Sabiduría y elaboración de la muerte en los rituales mortuorios. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 173-178.
159	CHU2-2001 33(2)	Cachiguango, L. (2001). ¡Wantiay...! El ritual funerario andino de adultos en Otavalo, Ecuador. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 179-186.

160	CHU3-2001 33(2)	Cáceres, E. (2001). La muerte como sanción y compensación: Visión de equilibrio y reciprocidad en Cusco. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 187-200.
161	CHU4-2001 33(2)	Fernández, G. (2001). Almas y difuntos: Ritos mortuorios entre los Aymara lacustres del Titicaca. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 201-219.
162	CHU5-2001 33(2)	van Kessel, J. (2001). El ritual mortuorio de los Aymara de Tarapacá como vivencia y crianza de la vida. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 221-234.
163	CHU6-2001 33(2)	Madani, D. (2001). Alma Imaña. Rituales mortuorios andinos en las zonas rurales Aymara de puno circunlacustre (Perú). <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 235-244.
164	CHU7-2001 33(2)	Escalante, M. (2001). La mesa de Todos Santos en San Pedro de Atacama Chile. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 245-248.
165	CHU8-2001 33(2)	De Hoyos, M. (2001). Saliendo del cajón por el río Jordán: Costumbres funerarias del Valle del Cajón, Catamarca, Argentina. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 249-252.
166	CHU9-2001 33(2)	Ortega, M. (2001). Escatología andina: Metáforas del alma. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 253-258.
167	CHU10-2001 33(2)	Acosta, O. (2001). La muerte en el contexto Uru: El caso Chipaya. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 259-270.
168	CHU11-2001 33(2)	Bascopé, V. (2001). El sentido de la muerte en la cosmovisión andina: El caso de los valles andinos de Cochabamba. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 271-277.

169	CHU12-2001 33(2)	Ceruti, M. (2001). La capacocha del nevado de Chañi una aproximación preliminar desde la arqueología. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 279-282.
170	CHU13-2001 33(2)	Chávez, J. (2001). Investigaciones arqueológicas de alta montaña en el sur de Perú. <i>Chungará, Revista de Antropología Chilena</i> , 33(2), 283-288.
171	CHU1-2002 34(1)	Garay de Fumagalli, M. & Cremonte, M. (2002). Ocupaciones agropastoriles tempranas al sur de la quebrada de Huamahuaca (Jujuy, Argentina). <i>Chungará, Revista de Antropología Chilena</i> , 34(1), 35-52.
172	CHU2-2002 34(1)	Schiappacasse, V. & Niemeyer, H. (2002). Ceremonial Inca provincial: El asentamiento de Sagura (cuenca de Camarones). <i>Chungará, Revista de Antropología Chilena</i> , 34(1), 53-84.
173	CHU3-2002 34(1)	Gavilán, V. (2002). "Buscando vida": Hacia una teoría Aymara de la división del trabajo por género. <i>Chungará, Revista de Antropología Chilena</i> , 34(1), 101-117.
174	CHU4-2002 34(1)	Albeza, M., Acreche, N. & Caruso, G. (2002). Biodemografía en poblaciones de la puna (Chañarcito, Santa Rosa de los Pastos Grandes y Olacapato) Salta Argentina. <i>Chungará, Revista de Antropología Chilena</i> , 34(1), 119-126.
175	CHU5-2002 34(1)	Córdova, J., Ossandón, Y., Álvarez, N. & Bernal, J. (2002). De las sociedades Chinchorro al presente. Un enfoque de ciencia y tecnología en educación de museo. <i>Chungará, Revista de Antropología Chilena</i> , 34(1), 127-140.
176	CHU1-2002 34(2)	Méndez, C. (2002). Cazadores recolectores costeros y sus contextos de tarea. Una visión desde el asentamiento holocénico temprano de Punta Penitente (lv 014), los Vilos. <i>Chungará, Revista de Antropología Chilena</i> , 34(2), 153-166.

177	CHU2-2002 34(2)	Falabella, F., Sanhueza, L. & Fonseca, E. (2002). Las materias primas de la cerámica aconcagua salmón y sus implicancias para la interpretación de la organización de la producción alfarera. <i>Chungará, Revista de Antropología Chilena</i> , 34(2), 167-189.
178	CHU3-2002 34(2)	Romero, A. (2002). Cerámica doméstica del valle de Lluta: Cultura local y redes de interacción Inca. <i>Chungará, Revista de Antropología Chilena</i> , 34(2), 191-213.
179	CHU4-2002 34(2)	Ponce, E. (2002). Mazas prehispánicas de metal: Sur de Perú y extremo norte de Chile. <i>Chungará, Revista de Antropología Chilena</i> , 34(2), 215-223.
180	CHU5-2002 34(2)	Varela, V. (2002). Enseñanzas de alfareros toconceños: Tradición y tecnología en la cerámica. <i>Chungará, Revista de Antropología Chilena</i> , 34(2), 225-252.
181	CHU1-2003 35(1)	Llagostera, A. (2002). Patrones de momificación Chinchorro en las colecciones Uhle y Nielsen. <i>Chungará, Revista de Antropología Chilena</i> , 35(1), 5-22.
182	CHU2-2003 35(1)	Sanhueza, L., Vásquez, M. & Falabella, F. (2003). Las sociedades alfareras tempranas de la cuenca de Santiago. <i>Chungará, Revista de Antropología Chilena</i> , 35(1), 23-50.
183	CHU3-2003 35(1)	Rodríguez, M., Rúgolo de Agrasar, Z. & Aschero, C. (2003). El género <i>deyeuxia</i> (Poaceae, agrostidae) en sitios arqueológicos de la puna meridional Argentina, provincia de Catamarca. <i>Chungará, Revista de Antropología Chilena</i> , 35(1), 51-72.
184	CHU4-2003 35(1)	Henríquez, F. (2003). Productos y organización técnica del trabajo en Azapa durante el siglo XVIII: Poniendo chacra de ají, cogiendo aceitunas. <i>Chungará, Revista de Antropología Chilena</i> , 35(1), 125-140.

185	CHU5-2003 35(1)	Araya, M. (2003). Fiscalidad y economía regional: Arica 1759-1799. <i>Chungará, Revista de Antropología Chilena</i> , 35(1), 141-157.
186	INE1-2000 24(1)	Gutiérrez, M. (2000). Efectos del endeudamiento sobre los resultados de las empresas en sectores en crisis. <i>Investigaciones Económicas</i> , 24(1), 75-116.
187	INE2-2000 24(1)	Vázquez, F. & León, C. (2000). Modelización del aprendizaje en valoración contingente. <i>Investigaciones Económicas</i> , 24(1), 117-138.
188	INE3-2000 24(1)	Blanco, R. (2000). Efectos sobre la volatilidad del mercado bursátil de la introducción de los contratos de futuros y opciones sobre el IBEX-35. <i>Investigaciones Económicas</i> , 24(1), 139-175.
189	INE4-2000 24(1)	Maudos, J., Pastor, J. & Serrano, L. (2000). Crecimiento de la productividad y su descomposición en progreso técnico y cambio de eficiencia: Una aplicación sectorial y regional en España (1964-93). <i>Investigaciones Económicas</i> , 24(1), 177-205.
190	INE5-2000 24(2)	Conesa, J. & Garriga, C. (2000). Reforma del sistema de seguridad social y adquisición de formación. <i>Investigaciones Económicas</i> , 24(2), 271-295.
191	INE6-2000 24(2)	Muñoz, M. (2000). Estructura demográfica y sistemas de pensiones. Un análisis de equilibrio general aplicado a la economía Española. <i>Investigaciones Económicas</i> , 24(2), 297-327.
192	INE7-2000 24(2)	Ropero, M. (2000). Adquisición de acciones propias: Factores explicativos para el caso español. <i>Investigaciones Económicas</i> , 24(2), 329-355.
193	INE8-2000 24(2)	González, V. & González, F. (2000). Procedimientos de resolución de insolvencia financiera en España: Costes de insolvencia y transferencias de riqueza. <i>Investigaciones Económicas</i> , 24(2), 357-384.

194	INE9-2000 24(2)	Corredor, P. & Santamaría, R. (2000). La estructura temporal de las volatilidades implícitas en la opción sobre el IBEX-35. <i>Investigaciones Económicas</i> , 24(2), 385-417.
195	INE10-2000 24(2)	Ferri, J. & Uriel, E. (2000). Multiplicadores contables y análisis estructural en la matriz de contabilidad social. Una aplicación al caso español. <i>Investigaciones Económicas</i> , 24(2), 419-453.
196	INE11-2000 24(3)	Campos, J. (2000). Responsabilidad limitada, estructura financiera y comportamiento de las empresas españolas. <i>Investigaciones Económicas</i> , 24(3), 585-610.
197	INE12-2000 24(3)	De Andrés, P., Azofra, V. & Rodríguez, J. (2000). Endeudamiento, oportunidades de crecimiento y estructura contractual: Un contraste empírico para el caso español. <i>Investigaciones Económicas</i> , 24(3), 385- 417.
198	INE1-2001 25(1)	Álvarez, B. (2001). La demanda atendida de consultas médicas y servicios urgentes en España. <i>Investigaciones Económicas</i> , 25(1), 93-138.
199	INE2-2001 25(1)	Gómez, J. (2001). Rentabilidad y liquidez alrededor de la fecha de desdoblamiento de las acciones. <i>Investigaciones Económicas</i> , 25(1), 171-202.
200	INE3-2001 25(1)	Alonso, A. (2001). Impacto sobre el mercado bursátil del vencimiento de los contratos de derivados sobre el Ibex 35. <i>Investigaciones Económicas</i> , 25(1), 203-234.
201	INE4-2001 25(2)	Goicolea, A., Lisandro, O. & Maroto, R. (2001). Picos de inversión y productividad del trabajo en los establecimientos industriales madrileños. <i>Investigaciones Económicas</i> , 25(2), 255-288.
202	INE5-2001 25(2)	Escribá, F. & Murgui, M. (2001). Tecnología, cambio estructural y convergencia en las regiones españolas (1980-1995). <i>Investigaciones Económicas</i> , 25(1), 335-357.

203	INE6-2001 25(2)	Serrano, G. (2001). Economías externas y crecimiento desigual en las regiones españolas. Un modelo dinámico de datos de panel. <i>Investigaciones Económicas</i> , 25(1), 359-389.
204	INE7-2001 25(2)	María-Dolores, R. (2001). Asimetrías en los efectos de la política monetaria en España (1977-1996). <i>Investigaciones Económicas</i> , 25(1), 391-415.
205	INE8-2001 25(3)	García, R. & Doménech, R. (2001). Estructura fiscal y crecimiento económico en la OCDE. <i>Investigaciones Económicas</i> , 25(1), 441-472.
206	INE9-2001 25(3)	Castelló, A. (2001). Desigualdad en la distribución de la renta, políticas impositivas y crecimiento económico en los países de la OCDE. <i>Investigaciones Económicas</i> , 25(1), 473-514.
207	INE10-2001 25(3)	Carnicero, I. (2001). Comercio y convergencia regional: Un análisis sectorial del caso español. <i>Investigaciones Económicas</i> , 25(1), 603-624.
208	INE1-2002 26(2)	Martínez, M. & Navarro, E. (2002). El modelo de Mccallum. Evidencia empírica en la estructura temporal de los tipos de interés española. <i>Investigaciones Económicas</i> , 26(2), 323-357.
209	INE2-2002 26(2)	Del Río, C. (2002). Desigualdad intermedia paretiana. <i>Investigaciones Económicas</i> , 26(2), 299-321.
210	INE3-2002 26(3)	López, V. (2002). ¿Ha seguido el Banco de España una regla de Taylor con información en tiempo real? <i>Investigaciones Económicas</i> , 24 (3), 475-496.
211	INE4-2002 26(2)	Blázquez, J. & Martín, J. (2002). Capital privado y competencia internacional: La armonización fiscal nunca es suficiente. <i>Investigaciones Económicas</i> , 26(2) 359-372.
212	INE5-2002 26(3)	Esteller, A. (2002). La configuración de una tasa judicial: Análisis teórico. <i>Investigaciones Económicas</i> , 26(3), 525-550.

213	INE6-2002 26(3)	Lozano, M., De Miguel, A. & Pindado, J. (2002). Papel de política de dividendos en las empresas reguladas. <i>Investigaciones Económicas</i> , 26(3), 447-474.
214	INE1-2003 27(1)	Segura, J. & Toledo, L. (2003). Tamaño, estructura y coste de financiación de las empresas manufactureras españolas. <i>Investigaciones Económicas</i> , 27(1), 39-69.
215	INE2-2003 27(1)	Perales, N. & Vázquez, L. (2003). Determinantes de la intensidad franquiciadora: Un enfoque de agencia. <i>Investigaciones Económicas</i> , 27(1), 151-172.
216	INE3-2003 27(2)	López, J. & Ridruejo, Z. (2003). Pensiones, crecimiento económico y envejecimiento poblacional. <i>Investigaciones Económicas</i> , 27(2), 343-367.
217	INE5-2003 27(3)	Gómez, J. & Yzaguirre, J. (2003). Presión sobre los precios en las revisiones del índice ibex35. <i>Investigaciones Económicas</i> , 27(3), 491-531.
218	INE6-2003 27(3)	Massot, M. & Nave, J. (2003). La hipótesis de las expectativas en el largo plazo: Evidencia en el mercado español de deuda pública. <i>Investigaciones Económicas</i> , 27(3), 533-564.
219	INE7-2003 27(3)	Pérez, P. (2003). Fuentes de variabilidad en las principales economías occidentales. <i>Investigaciones Económicas</i> , 27(3), 565-591.
220	NS1-2000 165	Rojas, J. (2000). La sociedad chilena posdictatorial. Entre la modernización y el imaginario democrático. <i>Nueva Sociedad</i> , 165, 32-43.
221	NS2-2000 167	Del Olmo, R. (2000). Ciudades duras y violencia urbana. <i>Nueva Sociedad</i> , 167, 74-86.
222	NS3-2000 168	Ocampo, J. (2000). Hacia una reforma financiera internacional: Cinco temas esenciales. <i>Nueva Sociedad</i> , 168, 72-78.

223	NS4-2000 169	Ruiz-Tagle, J. (2000). Las organizaciones sindicales frente a la exclusión social en el Mercosur. <i>Nueva Sociedad</i> , 169, 62-78.
224	NS1-2001 171	Rabotnikof, N. (2001). La caracterización de la sociedad civil perspectiva de los bancos multilaterales de desarrollo. <i>Nueva Sociedad</i> , 171, 101-119.
225	NS2-2001 172	Acosta, A. (2001). El falso dilema de la dolarización. <i>Nueva Sociedad</i> , 172, 66-84.
226	NS3-2001 173	Ramos, A. (2001). Viejo y nuevo. Partidos y sistemas de partidos en las democracias andinas. <i>Nueva Sociedad</i> , 173, 65-75.
227	NS4-2001 174	Rubio, B. (2001). La agricultura latinoamericana: Una década de subordinación excluyente. <i>Nueva Sociedad</i> , 174, 54-65.
228	NS5-2001 175	García, N. (2001). Por qué legislar sobre industrias culturales. <i>Nueva Sociedad</i> , 175, 60-69.
229	NS6-2001 176	Messner, D. (2001). Globalidad y gobernabilidad global. <i>Nueva Sociedad</i> , 176, 48-66.
230	NS1-2002 177	Lander, E. (2002). Los civilizados y los bárbaros. <i>Nueva Sociedad</i> , 177, 55-59.
231	NS2-2002 178	Duany, J. (2002). Nación, migración, identidad. Sobre el transnacionalismo a propósito de Puerto Rico. <i>Nueva Sociedad</i> , 178, 56-69.
232	NS3-2002 179	Cheresky, I. (2002). Autoridad política debilitada y presencia ciudadana de rumbo incierto. <i>Nueva Sociedad</i> , 179, 112-129.
233	NS4-2002 180	Aricó, J. (2002). Marx y América Latina. <i>Nueva Sociedad</i> , 180, 71-86.
234	NS5-2002 182	Delamata, G. (2002). De los "estallidos" provinciales a la generalización de las protestas en Argentina. Perspectiva y contexto en la significación de las nuevas protestas. <i>Nueva Sociedad</i> , 182.

235	NS1-2003 183	Acosta, A. & Ugarteche, O. (2003). Una propuesta global para un problema global: Hacia un tribunal internacional de arbitraje de deuda soberana. <i>Nueva Sociedad</i> , 183, 119-133.
236	NS2-2003 184	Pérez, J. (2003). Globalización, riesgo y empleabilidad. Algunas hipótesis. <i>Nueva Sociedad</i> , 184, 68-85.
237	NS3-2003 184	Hualde, A. (2003). ¿Existe un modelo maquilador? Reflexiones sobre la experiencia mexicana y centroamericana. <i>Nueva Sociedad</i> , 184, 86-101.
238	NS4-2003 185	Tokatlian, J. (2003). El orden sudamericano después de Irak. <i>Nueva Sociedad</i> , 185, 102-114.
239	NS5-2003 185	Dilla, H. (2003). Cuba, la compleja coyuntura. <i>Nueva Sociedad</i> , 185, 4-14.
240	NS6-2003 186	De la Ossa, A. (2003). Cooperación e integración entre gobiernos en Centroamérica. <i>Nueva Sociedad</i> , 186, 60-68.
241	NS7-2003 186	Giacalone, R. (2003). Integración Norte/Sur y tratamiento especial y diferenciado en el contexto regional. <i>Nueva Sociedad</i> , 186, 69-85.
242	NS8-2003 187	Knoop, J. (2003). El Brasil con Lula ¿Más de lo mismo? <i>Nueva Sociedad</i> , 187, 44-58.
243	NS9-2003 187	Santos, R. (2003). Lula y el Movimiento de los Sin Tierra. En la hora de la política. <i>Nueva sociedad</i> , 187, 128-138.
244	NS10-2003 188	Costa, S. (2003). Derechos humanos en el mundo posnacional. <i>Nueva Sociedad</i> , 188, 52-65
245	NS11-2003 188	Laserna, R. (2003). Bolivia: Entre populismo y democracia. <i>Nueva Sociedad</i> , 188, 4-14.

“Referencias bibliográficas de los artículos utilizados en la muestra de investigación”

CIENCIAS EXACTAS		
ID	Código	Referencia Bibliográfica
1	BSCQ1-2000 45(3)	Alvarado, Y., Cubillán, N., Soscun, H., Soscun, H., Osorio, E. & Vallejo, R. (2000). Estudio teórico de la polarizabilidad molecular en triafulveno y compuestos análogos. <i>Boletín de la Sociedad Chilena de Química</i> , 45(3), 339-346.
2	BSCQ2-2001 46(1)	Dimitrijewits, M. I., Arciprete, C., Becerra, A. & Castro-Luna, A. (2001). Reformado de metano con Co ₂ sobre catalizadores de níquel: efectos de la técnica de preparación. <i>Boletín de la Sociedad Chilena de Química</i> , 46(1), 65-75.
3	BSCQ5-2001 46(4)	Herrera, C., Tello, M., Ibáñez, I., Velásquez, L., Mancini, M., & Álvarez, S. (2001). Digestión proteica con microondas y su aplicación en preparación de muestras para análisis de harina de pescado. <i>Boletín de la Sociedad Chilena de Química</i> , 46(4), 487-494.
4	RPQ2-2000 3(1)	Pizarro, R. & Salas, G. (2000). Diseño y construcción de un sistema de flotación por aire disuelto, de efluentes líquidos, a nivel de laboratorio. <i>Revista Peruana de Química e Ingeniería Química</i> , 3(1), [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/ing_quimica/Vol3_Nº1/diseneno.htm

5	AVC3-1999 50(1)	Rivas, H. & Gutiérrez, X. (1999). Los surfactantes: comportamiento y algunas de sus aplicaciones en la industria petrolera. <i>Actas Venezolanas de Ciencias</i> , 50(1), 54-65.
---	-----------------	--

CIENCIAS BIOLÓGICAS

ID	Código	Referencia Bibliográfica
1	AMV3-2002 34(1)	Sievers, G., Jara, M., Cárdenas, C. & Núñez, J. (2002). Estudio anual de la eliminación de huevos y ooquistes de parásitos gastrointestinales y larvas de nemátodos pulmonares en ovinos de una estancia en Magallanes, Chile. <i>Archivos de Medicina Veterinaria</i> , 34(1), 37-47.
2	GC3-2002 66(2)	Lara, G., Parada, E. & Peredo, S. (2002). Alimentación y conducta alimentaria de la almeja de agua dulce <i>Diplodon chilensis</i> (bivalvia: hyriidae). <i>Gayana (Concepc.)</i> , 66(2), 107-112.
3	GC3-2003 67(1)	Alarcón, M. (2003). Sifonapterofauna de tres especies de roedores de Concepción, VI Región, Chile. <i>Gayana (Concepc.)</i> , 67(1), 16-24.
4	RCHA4-2001 19(3)	Vásquez, B. (2001). Presencia de CBG en el estroma ovárico de mamíferos. <i>Revista Chilena de Anatomía</i> , 19(3), 279-284.
5	RCHA5-2001 19(3)	Castro, A., Ghezzi, M., Alzota, R., Lupidio, M. & Rodríguez, J. (2001). Morfología del hígado de llama (<i>Lama glama</i>). <i>Revista Chilena de Anatomía</i> , 19(3), 291-296.
6	RCHA5-2002 20(2)	Briones, F., Calderón, M., Muñoz, J., Venegas, F. & Araya, N. (2002). El anticuerpo monoclonal Ki-67 como elemento de valor diagnóstico y pronóstico en neoplasias mamarias caninas. <i>Revista Chilena de Anatomía</i> , 20(2), 165-168.
7	RCHA2-2002 20(3)	Babinski, M., Chagas, M., Costa, W. & Pereira, M. (2002). Morfología y fracción del área del lumen glandular de la zona de transición en la próstata

		humana. <i>Revista Chilena de Anatomía</i> , 20(3), 255-262.
7	RCHN6-2000 73(4)	Véliz, D. & Vásquez, J. (2000). La Familia Trochidae (Mollusca: Gastropoda) en el norte de Chile: consideraciones ecológicas y taxonómicas. <i>Revista Chilena de Historia Natural</i> , 73(4), 757-769.
9	RCHN7-2000 73(4)	Martínez, G. & Montecino, V. (2000). Competencia en Cladocera: implicancias de la sobreposición en el uso de los recursos tróficos. <i>Revista Chilena de Historia Natural</i> , 73(4), 787-795.
10	RCHN1- 2002 75(2)	Canals, M., Atala, C., Olivares, R., Novoa, F. & Rosenmann, M. (2002). La asimetría y el grado de optimización del árbol bronquial en <i>Rattus norvegicus</i> y <i>Oryctolagus cuniculus</i> . <i>Revista Chilena de Historia Natural</i> , 75(2), 271-282.
11	AMV2-2001 33(1)	Díaz, D., Picco, E., Encinas, T. Rubio, M. & Litterio, N. (2001). Residuos tisulares de nicotinato de norfloxacin administrado por vía oral en cerdos. <i>Archivos de Medicina Veterinaria</i> , 33(1), 37-42.
12	AMV15-2002 34(2)	Perfumo, C., Sanguinetti, H., Giorgio, N., Armocida, A., Machuca, M., Massone, A., Risso, M., Aguirre, J. & Idiart, J. (2002). Constrictura rectal en cerdos necropsiados en una granja de ciclo completo en confinamiento. Consideraciones sobre su prevalencia, hallazgos anatomopatológicos y etiopatogenia. <i>Archivos de Medicina Veterinaria</i> , 34(2), 245-252.

CIENCIAS SOCIALES

ID	Código	Referencia Bibliográfica
1	AMB1-2001 6	Emanuelli, P. (2001). Dominante cultural y productos televisivos: Géneros que homogeneizan preferencias. <i>Ámbitos</i> , 6, 7-20.
2	AMB11-2001 6	Barrero, A. (2001). Juicios paralelos y Constitución: Su relación con el

		Periodismo. <i>Ámbitos</i> , 6, 171-189.
3	AMB14-2001 6	Egea, C. (2001). La carrera por la comunicación local (1998-2000) “Los grandes” se atreven con “lo pequeño”. <i>Ámbitos</i> , 6, 237-260.
4	AD14-2002 5	Moreira, J. (2002). Aplicaciones al análisis automático del contenido provenientes de la teoría matemática de la información. <i>Anales de Documentación</i> , 5, 273-286.
5	CHU2-2002 34(1)	Schiappacasse, V. & Niemeyer, H. (2002). Ceremonial Inca provincial: El asentamiento de Sagura (cuenca de Camarones). <i>Chungará, Revista de Antropología Chilena</i> , 34(1), 53-84.
6	INE4-2001 25(2)	Goicolea, A., Lisandro, O. & Maroto, R. (2001). Picos de inversión y productividad del trabajo en los establecimientos industriales madrileños. <i>Investigaciones Económicas</i> , 25(2), 255-288.
7	INE2-2002 26(2)	Del Río, C. (2002). Desigualdad intermedia paretiana. <i>Investigaciones Económicas</i> , 26(2), 299-321.
8	NS3-2003 184	Hualde, A. (2003). ¿Existe un modelo maquilador? Reflexiones sobre la experiencia mexicana y centroamericana. <i>Nueva Sociedad</i> , 184, 86-101.
9	NS7-2003 186	Giacalone, R. (2003). Integración Norte/Sur y tratamiento especial y diferenciado en el contexto regional. <i>Nueva Sociedad</i> , 186, 69-85.
10	NS10-2003 188	Costa, S. (2003). Derechos humanos en el mundo posnacional. <i>Nueva Sociedad</i> , 188, 52-65