



UNIVERSITAT_{DE}
BARCELONA

El dominio de la escritura en el género de investigación: de la redacción guiada a la publicación reconocida

Ana M^a Pujol Dahme

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

El dominio de la escritura en el género de investigación: de la redacción guiada a la publicación reconocida

Ana M^a Pujol Dahme

Tesis presentada para optar
al grado de **Doctor en Lingüística** en el programa de doctorado
Lingüística Teórica, Computacional y Aplicada,
Departamento de Lingüística,
Universidad de Barcelona

bajo la supervisión de
Dr. Moisés Selfa
Universidad de Lérida

Tutora: **Dra. Liliana Tolchinsky**
Universidad de Barcelona

Universidad de Barcelona
Marzo de 2016

Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.

Los límites de mi lenguaje significan los límites de mi mundo.

(Wittgenstein, 1922)

Si yo no soy para mí mismo, ¿quién será para mí?

Si yo soy para mí solamente, ¿qué soy yo?

Y si no ahora, ¿cuándo?

Mishná, Pirke Avot (1:14)

A la memoria de mi padre, Dr. Pere Pujol Amat,
cuya curiosidad intelectual siempre fue ejemplo y estímulo.

Agradecimientos

Esta tesis ha llegado a materializarse gracias a la colaboración de muchas personas. Por falta de espacio (*léase*: memoria) no mencionaré a todos. Si tú, estimado lector, no te quedaste inmóvil al borde del camino, no congelaste el júbilo, no te quedaste sin labios, no te salvaste, ni antes, ni ahora, ni nunca, entonces...sírrete de mi gratitud.

(adaptado de M. Benedetti, 1973-74)

Debo agradecer a la Dra. Liliana Tolchinsky la confianza que me mostró, desde el principio, al darme total autonomía en el desarrollo de esta tesis. De sus correcciones y comentarios, he aprendido mucho sobre escritura y demás. Al Dr. Moisés Selfa (Universidad de Lleida), aunque solo participó del tramo final de este largo recorrido, llegó en el momento oportuno. Su empatía y presencia de ánimo siempre han sido alentadoras.

Al coordinador, Dr. Lloberas, y tutores del Màster en Immunologia Avançada (UB-UAB), así como al Dr. Ferrer, coordinador del Màster d'Estudis Històrics (UB), por permitirme recopilar el corpus de universidad. A los profesores del Instituto Àusias March y especialmente a Lluïsa Castañeda del Instituto Sant Just Desvern, cuya disposición a facilitarme los trabajos de bachillerato fue ejemplar. A la Dra. Berlanga del Dpt. de Mètodes d'Investigació i Diagnòstic en Educació por su asesoramiento e interés. Tanto a la Dra. Pérez-Llantada de la revista *Ibérica*, como al Dr. Galbraith, de la revista *Journal of Writing Research*, por su amabilidad y profesionalidad, así como a los revisores anónimos de ambas revistas, cuyos comentarios han sido de gran valúa para este trabajo.

A todos mis compañeros de doctorado, pero en especial a Laia, gracias por compartir inquietudes. A Jin Tao, agradecerte las conversaciones tan entusiastas que hemos tenido sobre "futuras investigaciones", espero que se materialicen. A las investigadoras del Dpt. de Lingüística, que acabaron el recorrido mucho antes que yo y que han sido siempre ejemplo de amabilidad y buen humor, Gloria y Marta, gracias.

A Montse y Santi, del Servei de Tecnologia Lingüística, por su buena disponibilidad para ayudar.

A las compañeras, profesoras asociadas del Dpto. de Didáctica de la Lengua y la Literatura, en especial a Sole, Eva y Mireia por las distendidas charlas sobre el doctorado y las clases.

A Jordi, que siempre está disponible para relativizar, con su humor, todo. A Isabel, ejemplo de fortaleza, simpatía y ánimo. A Toni, experto en programación, sin su ayuda con los "shell scripts" aún estaría en ruta.

A mi familia, hermanos y sobrinos, y en especial a mi madre, "meine liebste Mutter", gracias por estar siempre aquí.

A Oscar, Tera y Nika, cuya compañía siempre fue y es alegría.

Pero mi más sentido agradecimiento es para una persona muy especial, gracias por acompañarme en el camino más importante, el de la vida.

Resumen

La alfabetización o cultura escrita está inexorablemente vinculada a la práctica de cualquier disciplina académica, así que para integrarse en una comunidad disciplinar, es necesario dominar la escritura académica. Es decir, adoptar las formas particulares que dicha comunidad utiliza, no solo para crear e interpretar significado, sino también para transmitir el conocimiento. La comunicación del conocimiento disciplinar, que se realiza generalmente a través del artículo de investigación, se caracteriza por tener un registro específico y un metadiscurso. El concepto *registro* refiere a los recursos léxicos y gramaticales prevalentes en contextos discursivos particulares. Por su parte, *metadiscurso* se define como el uso de los recursos lingüísticos para referirse al texto, al escritor y al lector. Por lo tanto, adquirir destreza en el registro y en el metadiscurso es crucial para lograr el éxito académico y el desarrollo profesional en la carrera de los investigadores.

Esta investigación tiene dos objetivos generales. En primer lugar, la creación de un corpus de trabajos académicos de investigación escritos en catalán, de bachillerato y máster universitario, en las disciplinas de biología e historia. Este corpus llamado TARBUC (*Treballs Acadèmics de Recerca de Batxillerat i Universitat en Català*) está digitalizado y disponible para la comunidad investigadora. En segundo lugar, caracterizar a partir de este corpus el desarrollo de la escritura académica en momentos diferentes de la historia educativa de los estudiantes. Para ello presentamos dos estudios.

En el estudio 1, analizamos transversalmente los trabajos de investigación de bachillerato y de final de máster universitario en las disciplinas de biología e historia, escritos en catalán. El estudio 2 es longitudinal. En él comparamos algunos de los trabajos de investigación de final del máster universitario, en catalán, con los artículos publicados en inglés, en la disciplina de biología, escritos por los mismos sujetos. En concreto, examinamos el registro y el metadiscurso. Analizamos el registro por medio de las siguientes características léxico-gramaticales, que son indicadoras de un mayor desarrollo de cultura escrita: (i) la diversidad léxica, (ii) la densidad léxica y (iii) la complejidad sintáctica.

Respecto al metadiscurso, estudiamos la función metadiscursiva interaccional a través de los marcadores lingüísticos, que señalan, por una parte, la voz textual del escritor para involucrarse con su texto, la postura (*stance*), y, por otra, los que utiliza el escritor para involucrar al lector, la afiliación (*engagement*). Si bien hay gran cantidad de investigaciones que han analizado el registro y el metadiscurso en el género de investigación, son pocos los estudios (sobre todo para el catalán) que abordan estos aspectos conjuntamente en un corpus de aprendices y desde una perspectiva evolutiva.

Por consiguiente, esta tesis doctoral llena un vacío en la investigación actual, porque amplía el panorama sobre variación lingüística en géneros académicos desde un enfoque evolutivo y en una lengua románica como es el catalán. Además, tiene un potencial para diseñar estrategias didácticas en la enseñanza de la escritura académica.

En la educación secundaria y universitaria los estudiantes han de desarrollar las competencias asociadas a la investigación, habilidades que van más allá de saber el contenido de la materia. Su aprendizaje del conocimiento disciplinar incluye participar de un contexto social, en el que uno de los géneros más utilizado para la creación y transmisión de conocimiento es el género de investigación. Por lo tanto, el hecho de afrontar este género les permite asimilar las prácticas de la comunidad disciplinar y desarrollar una cultura escrita avanzada, que les posibilitará contribuir a la evolución del conocimiento.

Así, con el propósito de cumplir con el objetivo general de caracterizar el desarrollo de la escritura académica y abordar el estudio 1, compilamos el primer corpus académico-científico de bachillerato y universidad en catalán, corpus de *Treballs Acadèmics de Recerca de Batxillerat i de Universitat en Català* (TARBUC), que consta de un total de 110 textos, 57 trabajos de investigación del nivel de bachillerato (edad 17-18 años) y 53 trabajos de final de máster universitario, de las disciplinas de biología e historia. Se seleccionaron estas dos disciplinas porque son representativas de dos tradiciones metodológicas de investigación diferenciadas, la experimental-cuantitativa y la exploratoria-cualitativa.

Para poder llevar a cabo el estudio 2, compilamos un sub-corpus de los mismos ocho sujetos en dos períodos diferentes de tiempo: como estudiantes de universidad y como investigadores profesionales. Este corpus consta de 16 textos, está integrado por trabajos final de máster en catalán y por artículos publicados en revistas con factor de impacto, escritos en inglés y de la disciplina de biología. Al haber pasado una revisión por pares, estos textos constituyen prototipos de cómo se presenta y se negocia el conocimiento en esta comunidad discursiva disciplinar.

Examinar el registro lingüístico y el metadiscurso supone hacer un análisis cuantitativo y cualitativo. Los indicadores del registro, es decir, las características léxico-gramaticales los examinamos cuantitativamente. Por su parte para los índices del metadiscurso realizamos un análisis más cualitativo, con el propósito de comprobar las funciones desempeñadas por las formas lingüísticas. A efectos de ello, utilizamos un analizador multilingüe, *FreeLing* (Padró & Stanilovsky, 2012), un *shell script* para el sistema operativo Linux y un software de concordancias, *AntConc 3.2.4w* (Anthony, 2011). Finalmente los indicadores del registro y del metadiscurso se evaluaron estadísticamente. Todas las herramientas de análisis utilizadas son de código abierto, lo que facilita la replicabilidad de este estudio.

Los resultados del estudio 1 revelan cambios evolutivos significativos, en los textos, a través de los niveles académicos, es decir, la escuela secundaria y la universidad. Los trabajos de universidad muestran una mayor complejidad sintáctica y densidad léxica respecto a los de secundaria, mientras que la diversidad léxica no varía. Los resultados de la prueba discriminante sugieren que la densidad léxica es la que mejor predice la clasificación de un texto en un nivel u otro. Por otro lado, los resultados sobre la función interaccional indican que, en cuanto a los marcadores de evaluación (*stance*), concretamente el uso de los mitigadores o atenuadores (*hedges*) aumenta en los textos universitarios, y se correlaciona con un tipo específico de marcador que involucra al lector (*engagement*), aquel que indica a los lectores cómo interpretar un argumento. En cambio los textos de secundaria muestran un número significativamente mayor de apelación al conocimiento compartido, donde se les pide a los lectores reconocer algo como aceptado por la mayoría.

Así, con un mayor conocimiento de la cultura escrita y mayor madurez cognitiva, se moviliza la capacidad para elegir entre las opciones retóricas posibles y así guiar al lector a través de múltiples puntos de vista hacia su propio argumento. La mención del autor (*self-mention*) es la estrategia retórica más utilizada por los estudiantes, tanto de secundaria como de universidad, y esta característica se ajusta a los requisitos de la comunidad discursiva disciplinar. Esto sugiere que la convención del género del artículo de investigación condiciona la selección de los recursos de interacción, ya a partir de la escuela secundaria.

Por el contrario, los resultados respecto a la mayoría de las características léxico-gramaticales, es decir, la complejidad sintáctica y la densidad léxica, no muestran diferencias significativas en las disciplinas de biología e historia. Sí hay diferencias de longitud, que es mayor en la disciplina de historia, y hay una mayor diversidad léxica en la disciplina de biología. En cambio, la frecuencia de uso de los marcadores interaccionales es significativamente diferente entre biología e historia. Los resultados indican que en biología hay un mayor porcentaje de intensificadores, atenuadores, marcadores de auto-mención, directivas al texto, al argumento y conocimiento compartido, comparado con la disciplina de historia.

En el estudio 2, en el que comparamos longitudinalmente un corpus de trabajos de investigación de final del máster universitario, en catalán, y de publicaciones en inglés en la disciplina de biología, los resultados muestran diferencias significativas. Hay una variación de registro entre los artículos publicados y los trabajos de investigación universitarios. Las publicaciones tienen una mayor diversidad léxica que los trabajos universitarios. Los trabajos de universidad son más complejos, según la primera medida de complejidad sintáctica y son más densos, es decir, tienen mayor densidad léxica, que las publicaciones. En cambio, las publicaciones tienen mayor complejidad sintáctica, entendida esta como media de palabras antes del primer verbo finito.

En cuanto a los marcadores metadiscursivos interaccionales, los resultados revelan que hay un mayor uso de los marcadores de postura en las publicaciones respecto a los trabajos de universidad. Asimismo, las publicaciones tienen más marcadores de actitud y más marcadores de auto-mención que los textos universitarios. Por el contrario, los trabajos universitarios tienen más porcentaje de pronombres del lector que las publicaciones. Para los demás marcadores, tales como, los intensificadores, atenuadores, directivas al texto, al argumento y marcadores de conocimiento compartido, no hay diferencias significativas entre los trabajos universitarios y las publicaciones.

Esta investigación ha aportado un corpus de textos de aprendices, es decir, unas muestras de lengua en uso de estudiantes, que están en el proceso de aprendizaje de la escritura académica y que estará disponible para la comunidad investigadora. Asimismo, ha permitido ampliar un poco más el panorama y ha ofrecido un marco evolutivo de la escritura académica, en concreto, del registro y del metadiscurso, en el género de investigación desde bachillerato hasta la publicación. Es decir, ha posibilitado describir el desarrollo de algunas características léxico-gramaticales y averiguar qué recursos retóricos utilizan los estudiantes cuando participan en una de las actividades de su comunidad de práctica.

ÍNDICE GENERAL

Agradecimientos	6
Resumen	8
Índice general	11
Índice de tablas y gráficos	15
CAPÍTULO 1	17
INTRODUCCIÓN	17
1.1 La alfabetización académica avanzada	20
1.2 Los trabajos de investigación y los artículos de investigación: diferentes géneros, algunas similitudes	25
1.3 El contexto académico y cultural del corpus TARBUC	33
CAPÍTULO 2	37
2. Objetivos	37
2.1 Predicciones	40
CAPÍTULO 3	41
3. Estructura general de la tesis	41
CAPÍTULO 4	42
4. Antecedentes teóricos y empíricos	42
4.1 La modalidad oral y escrita	42
4.2 La alfabetización: Una perspectiva multidimensional	46
4.2.1 La dimensión lingüística	49
4.2.2 La dimensión cognitiva	55
4.2.3 La dimensión sociocultural	58
4.2.4 La dimensión de desarrollo	60
4.2.5 La dimensión multimodal	61
4.3 Género	65
4.3.1 Comunidades discursivas versus comunidades de práctica	72
4.4 La escritura	74
4.5 La escritura académica	77
4.5.1 Funciones de la escritura	78

4.6 Registro	80
4.6.1 Léxico-gramática	81
4.7 Metadiscurso	85
CAPÍTULO 5	91
5. El análisis del discurso y la lingüística de corpus	91
5.1 El análisis del discurso	92
5.1.1 Enfoques del análisis del discurso	93
5.2 El análisis del discurso especializado	96
5.3 La lingüística de corpus	97
5.3.1 Entre la teoría y la metodología	98
5.3.2 El enfoque corpus-driven versus corpus-based	100
5.4 Definición de corpus	103
5.5 La lingüística de corpus y el funcionalismo	105
CAPÍTULO 6	108
6. Método	108
6.1 Descripción del corpus	108
6.2 Razones para seleccionar los textos	109
6.3 Procedimiento de recogida de textos.	111
6. 4 Categorías de análisis	112
6. 4. 1 Características léxico-gramaticales	112
6.4. 2 Características metadiscursivas interaccionales	116
6.5 Procesamiento y herramientas de análisis	121
6.5.1 Procesamiento	122
6.5.2 Herramientas de análisis	125
CAPÍTULO 7	131
7. Resultados del estudio 1	131
7.1 Resultados del registro académico	131
7.1.1 Resultados descriptivos de las características léxico-gramaticales	132
7.1.2 Resultados inferenciales de las características léxico-gramaticales por nivel	134
7.1.3 Resultados inferenciales de las características léxico-gramaticales según la disciplina	139
7.2 Resultados de los marcadores metadiscursivos interaccionales	141
7.2.1 Resultados descriptivos de los marcadores interaccionales	141
7.2.2 Resultados inferenciales de los marcadores interaccionales por niveles	143
7.2.3 Resultados inferenciales de los marcadores interaccionales por disciplinas	145

CAPÍTULO 8	147
8. Resumen de los resultados del estudio 1	147
CAPÍTULO 9	149
9. Discusión	149
CAPÍTULO 10	158
10. Resultados del estudio 2	158
10.1 Resultados descriptivos de las características léxico-gramaticales y los marcadores metadiscursivos interaccionales	159
10.2 Resultados inferenciales de las características léxico-gramaticales	160
10.3 Resultados inferenciales de los marcadores metadiscursivos interaccionales	162
CAPÍTULO 11	165
11. Resumen de los resultados del estudio 2	165
CAPÍTULO 12	167
12. Discusión	167
CAPÍTULO 13	174
13. Conclusiones	174
Bibliografía	178
Anexos	
Autorizaciones para hacer uso de los trabajos de máster universitarios	209
Títulos de los trabajos que conforman el sub-corpus TARBUC	212

ÍNDICE DE TABLAS Y GRÁFICOS

Tabla 6.1 Descripción del corpus	109
Tabla 7.1 Resultados descriptivos de los trabajos de investigación por nivel, bachillerato y universidad	133
Tabla 7.2 Resultados descriptivos de los trabajos de investigación por disciplina, biología e historia	133
Tabla 7.3 Resultados léxico-gramaticales según el nivel académico	135
Tabla 7.4 Correlaciones de las características léxico-gramaticales en bachillerato	136
Tabla 7.5 Correlaciones de las características léxico-gramaticales en universidad	137
Tabla 7.6 Resultados del análisis discriminante mediante bootstrapping	138
Tabla 7.7 Resultados de las características léxico-gramaticales según las disciplinas	140
Tabla 7.8 Resultados descriptivos de los marcadores interaccionales según los niveles	142
Tabla 7.9 Resultados descriptivos de los marcadores interaccionales según las disciplinas	142
Tabla 10.1 Resultados descriptivos de las características léxico-gramaticales	159
Tabla 10.2 Resultados descriptivos de los marcadores interaccionales	160
Tabla 10.3 Resultados características léxico-gramaticales entre universidad y publicaciones	161
Gráfico 6.1 Visión de conjunto de las unidades y herramientas de análisis	121
Gráfico 7.1 Resultados inferenciales de los marcadores interaccionales por niveles	144
Gráfico 7.2 Resultados inferenciales de los marcadores interaccionales según disciplinas	146
Gráfico 10.1 Resultados marcadores interaccionales en universidad y publicaciones	163
Figura 4.6.1 Léxico-gramática en un continuum	83
Figura 4.6.2 Líneas de concordancia del programa AntConc con eje sintagmático y paradigmático	84
Cuadro 6.4.2.1 Cuadro resumen de las características metadiscursivas interaccionales	119-120

CAPÍTULO 1

INTRODUCCIÓN

Llegar a dominar el registro académico y las convenciones retóricas del género de investigación es importante para lograr el éxito académico y la promoción de los investigadores (Berkenkotter, Huckin, & Ackerman, 1991; Schleppegrell, 2001; Snow & Uccelli, 2009; Tolchinsky, et al., 2013). El registro académico escrito suele presentar una alta diversidad léxica y una gran densidad informativa. Esta densidad se consigue por medio del uso de términos técnicos, de las nominalizaciones y de la ampliación del grupo nominal. Esta última se relaciona con la complejidad sintáctica propia del texto académico (Biber & Gray, 2010; Halliday, 1985; Snow & Uccelli, 2009; Teberosky, 2007).

A continuación ilustramos estas características mediante un ejemplo extraído de un artículo en inglés publicado en una revista de impacto del sub-corpus TARBUC. Lo hemos seleccionado, porque representa un extremo del continuum de la evolución de la escritura académica. Es decir, representa un ejemplar modelo del registro académico en una disciplina. En el siguiente ejemplo podemos apreciar el uso de términos técnicos (*en cursiva*), de nominalizaciones (**en negrita**), y de ampliación del nombre (subrayado) o grupo nominal, en este caso mediante sintagmas preposicionales [entre corchetes].

Recent findings provide evidence [of the essential role] [of *type B scavenger receptors*] [in the innate immune response] [to *exogenous pathogens*], in particular, to bacteria and their inflammatory compounds.

Además, no hay duda que el discurso académico especializado es dialógico e interactivo (Flowerdew, 2014; Swales, 2004). El escritor ha de expresar adecuadamente sus aserciones para que su trabajo sea apto para la publicación. Ha de adecuar su discurso a los requerimientos retóricos de su comunidad disciplinar. Esto repercutirá en la percepción de calidad de su trabajo y, por lo tanto, en su aceptación por parte de la comunidad académica específica (Hyland, 2005; Crismore, Markkanen, & Steffensen, 1993; Lafuente, 2008; Oliver del Olmo, 2004; Salager-Meyer, 1994; Vázquez & Giner, 2008).

Otros ejemplos extraídos del sub-corpus ilustran algunos elementos retóricos que se utilizan para la interacción, como son los atenuadores (*hedges*), con los que el escritor presenta la aserción como plausible; o intensificadores (*boosters*), con los que el escritor expresa su seguridad.

Our results **suggest** that as the monocyte matures, it becomes more sensitive to TLR signals, losing rapidly the expression of scavenger receptors.

Despite the relatively small sample size, we **clearly demonstrated** that the administration of rGH to HIV-infected patients before vaccination was associated with a significant recovery of cellular immunity.

Acabamos de ver algunas características léxico-gramaticales del registro académico y algunos marcadores que realizan la función metadiscursiva interaccional. En este sentido, ¿qué nos podría revelar el análisis de estas particularidades, a través de los niveles académicos, sobre el desarrollo del registro académico en el género de investigación? ¿Y sobre el conocimiento que tienen los estudiantes de la cultura escrita... y sobre el conocimiento de los requerimientos retóricos específicos de su comunidad de práctica? En las líneas que siguen intentaremos dilucidar estos y otros asuntos relacionados con la escritura.

El punto de partida de esta tesis doctoral es contribuir al mejor conocimiento del desarrollo del lenguaje escrito, desde una visión integradora de las perspectivas lingüísticas, cognitivas, sociales y culturales. Por eso, suscribimos la propuesta de Olson (2009a): “writing and literacy have become essential aspects of understanding of language, mind and society” (2009a, p.5). Esta investigación se enmarca dentro de la lingüística aplicada, en particular, en la intersección entre la lingüística y la educación. Caracterizamos desde una perspectiva lingüística un producto de la escritura académica, el trabajo de investigación, con la finalidad última de poder aportar conocimiento que pueda ser útil a la enseñanza de la escritura académica.

El objeto de estudio de esta investigación es un producto de la escritura académica, en concreto, el trabajo de investigación. El trabajo de investigación es un género que se sitúa entre los textos académicos y los científicos (Castelló & Iñiesta, 2012). *Académicos* porque se refiere al contexto de uso de esta escritura, es decir, a los textos escritos en entornos formales como escuelas, universidades o institutos de investigación. *Científicos* porque, como veremos más adelante, presentan algunas características comunes a los artículos de investigación publicados (Russell & Cortés, 2012; Snow & Uccelli, 2009). En consecuencia, consideramos oportuno empezar por presentar el tema desde una visión general para ir estrechando el foco al objeto de estudio.

Antes de centrarnos en el trabajo de investigación hemos de ampliar nuestra perspectiva hasta la alfabetización por tres motivos: el primero, porque esta ofrece los antecedentes para comprender las recientes teorías sobre escritura y su enseñanza; el segundo, para tener una visión de conjunto de lo que supone la escritura; y el tercero, porque la escritura académica es una muestra del saber de la cultura escrita. Por lo tanto, desde la alfabetización académica, las características del texto académico (trabajo de investigación), las singularidades del texto científico (el artículo de investigación) y el contexto académico y cultural se desprenden las justificaciones necesarias para abordar este estudio y plantear los objetivos que han guiado esta investigación.

Relevante para la perspectiva de desarrollo adoptada en este trabajo es la concepción de alfabetización propuesta por Ravid & Tolchinsky (2002). Es decir, el saber de cultura escrita implica la capacidad de utilizar el propio repertorio lingüístico y adaptarlo de forma flexible a diferentes destinatarios y contextos comunicativos, esto es, en diferentes modalidades, registros y géneros. En esta investigación la modalidad sería la escrita; el registro, el académico; y el género, el trabajo de investigación. Pero comencemos primero por la alfabetización académica avanzada, hasta llegar, paulatinamente, hasta el trabajo de investigación.

1.1 La alfabetización académica avanzada

La alfabetización es un sistema semiótico social, una forma de acción social para crear significado con un sentido para la comunidad (Halliday, 1978; Lemke, 1990). Las personas utilizamos los recursos de una variedad de sistemas semióticos, es decir, signos para crear significado. Elegimos un modo u otro, según los objetivos que queremos conseguir en los contextos interpersonales o institucionales en el que nos encontramos. Por ejemplo, escribimos el signo lingüístico (en nuestro caso, las palabras, el texto) con el objetivo que sea relevante para, y reconocible por, los miembros de alguna comunidad (Lemke, 2002).

La alfabetización es un proceso que se desarrolla en la sociedad y en el individuo. En el contexto social, esta evolución es especialmente dinámica, ya que sobre todo, los cambios tecnológicos se suceden rápidamente y requieren de nuevas formas de alfabetización para poder adaptarse a estos cambios. En el ámbito académico, se crea y se comparte nuevo conocimiento a través de significados complejos, con diferentes sistemas de notación como, por ejemplo, el escrito, el numérico y el visual, entre otros (Schleppegrell & Colombi, 2002).

En el contexto académico, a partir de la educación secundaria, los estudiantes se inician en las prácticas necesarias para poder participar de la cultura discursiva de las disciplinas especializadas. Es decir, los estudiantes aprenden modos particulares de crear e interpretar significado; de argumentar y criticar; de integrar diferentes significados impresos, visuales, electrónicos. Esto requiere un nivel avanzado de alfabetización.

Denominamos *alfabetización avanzada* a la forma de crear e interpretar significado a partir de la educación secundaria (Schleppegrell & Colombi, 2002). La alfabetización académica se extiende más allá de esta etapa educativa y es la que se requiere para poder participar en diferentes contextos profesionales (Schleppegrell & Colombi, 2002).

Desde una perspectiva psicolingüística se considera la adolescencia, 16-18 años, como un punto de inflexión en el conocimiento y uso lingüístico, tanto a nivel léxico y sintáctico como a nivel de posicionamiento discursivo (Berman, 2007; Berman & Verhoeven, 2002). La alfabetización académica avanzada consiste en desarrollar el dominio de nuevos géneros y las formas que cada comunidad disciplinar utiliza para crear y compartir el conocimiento (Halliday, 2004; Christie & Derewianka, 2008; Schleppegrell & Colombi, 2002). La alfabetización académica, por lo tanto, debería fomentar que los estudiantes participen de las prácticas discursivas de las disciplinas para poder apropiarse del conocimiento (Carlino, 2013). Es necesario favorecer que escriban y lean, que elaboren y comuniquen el conocimiento como lo hacen los especialistas, los miembros de cada comunidad disciplinar, ya que cada comunidad tiene sus propias prácticas discursivas. En otras palabras, sus miembros utilizan el lenguaje escrito de acuerdo con ciertos propósitos y valores epistémicos (Bazerman, 2005; Carlino, 2003, 2005; Ivanic, 1998; Prior, 2006). Por consiguiente, no podemos hablar de una única alfabetización académica, al contrario, cada ámbito tiene su propia cultura escrita.

Asimismo, la cultura escrita es una parte fundamental de todas las disciplinas académicas, también de las científicas. Podríamos pensar que lo que constituye la ciencia es su contenido y que la cultura escrita solo es una herramienta auxiliar para almacenar o transmitir conocimientos. No obstante, lo que caracteriza el discurso científico no es solamente la terminología técnica, sino la gramática técnica. Así lo afirma Halliday (1985, 2004): “The distinctive quality of scientific language lies in the lexicogrammar” (p. 201). Debido a los significados que se tienen que expresar en ciencia, se han tenido que utilizar nuevos recursos gramaticales, como, por ejemplo, la metáfora gramatical, recurso potente para agrupar significados en tema y rema.

Este hecho tiene una implicación sustancial: nuevas formas de utilizar el lenguaje conlleva nuevas formas de pensamiento. Es decir, la evolución de la ciencia está relacionada con la evolución de los recursos gramaticales, que a su vez, construyen el pensamiento científico (Halliday, 1993). La cultura escrita es intrínseca a la ciencia, como lo es la medición u observación (Norris & Phillips, 2009). De ahí que sea importante desarrollar un conocimiento letrado para lograr el éxito académico y el desarrollo profesional (Tolchinsky & Simó, 2001; Snow & Biancarosa, 2003).

Desarrollar un conocimiento letrado supone llegar a dominar nuevos géneros, como, por ejemplo, el género expositivo; y nuevos registros, como, por ejemplo, el lenguaje de la ciencia (Uccelli, Dobbs, & Scott, 2013). Por *género* se entiende un grupo de textos, que representa cómo los escritores usan el lenguaje para situaciones recurrentes. En un sentido amplio, el término género se puede caracterizar de la siguiente manera: cada género tiene un propósito específico, una estructura general, algunos rasgos lingüísticos específicos, y es compartido por los miembros de una comunidad (Hyland, 2004b). Vemos que en la definición de género, a parte de las características lingüísticas, se da importancia al contexto y a su dimensión social. Así, los géneros también reflejan las prácticas de cada comunidad, permiten a los usuarios realizar acciones situadas lingüística y retóricamente. En este sentido, Miller (1984) nos recuerda que el género nos permite entender cómo participar en las acciones de una comunidad. Por otra parte, Berkenkotter & Huckin (1995) señalan la *situacionalidad* de los géneros, por lo que la participación en las actividades comunicativas de la vida cotidiana y profesional facilita el conocimiento y asimilación de

los diferentes géneros. Por consiguiente, entendemos género también como un acto social, que contiene las formas de conocer y actuar propias del discurso de una comunidad, y que son dinámicas, cambian con el tiempo (Bazerman, 1988; Berkenkotter & Huckin, 1993; Miller, 1984). Cuando los estudiantes, por ejemplo, aprenden a usar el género de investigación, asimilan los valores y normas de la comunidad científica.

La capacidad de utilizar eficazmente los géneros es importante en niveles avanzados de escritura, debido a que proporcionan el marco del conocimiento. Llegar a dominar diferentes géneros supone alcanzar destreza en el uso de la lengua, en el propósito que se pretende conseguir con la escritura, en el contenido y en el contexto (Grabe & Kaplan, 1996; Swales, 1990). Cada género tiene su propio registro, es decir, una serie de rasgos o repertorios lingüísticos. Sin embargo, algunos rasgos lingüísticos son comunes a todas las disciplinas, son característicos del lenguaje académico (Schleppegrell, 2001; Uccelli, Dobbs, & Scott, 2013).

Basándonos en la revisión exhaustiva que Snow & Uccelli (2009) hacen de las características del lenguaje académico, consideramos relevantes para nuestra investigación las siguientes: la diversidad léxica, la densidad léxica y la complejidad sintáctica. La diversidad léxica es una medida que refleja la variedad de las palabras utilizadas en un texto (Johansson, 2008). La densidad léxica, es la proporción de palabras léxicas en relación con el número total de palabras. Las palabras léxicas transmiten la mayor parte del contenido semántico y la información proposicional, por lo tanto, una alta densidad léxica se considera un buen indicador de la riqueza textual (Halliday, 1985). La complejidad sintáctica refiere a la sofisticación y al rango o variedad de las formas sintácticas (Lu, 2011; Ortega, 2003). El lenguaje académico en comparación con el lenguaje coloquial tiene una gran diversidad léxica. También condensa mucha información, tiene una alta densidad léxica, mediante pocas palabras se da mucha información, esto se consigue habitualmente por medio de la nominalización, es decir, el cambio de la categoría gramatical protípica o congruente.

No obstante, un uso excesivo de nominalizaciones da la sensación de un texto duro, sin fluidez. Este hecho dificulta su legibilidad (Pinker, 2014; Ventola, 1996). Además, la sintaxis del lenguaje académico es compleja, pero esta complejidad se consigue no tanto por la subordinación, sino por la ampliación del grupo nominal (Biber & Gray, 2010). Por otro lado, el posicionamiento discursivo en la escritura académica suele ser impersonal y objetivo (Berman & Ravid, 2009; Chafe & Danielewicz, 1987; Snow & Uccelli, 2009; Uccelli, Dobbs & Scott, 2013).

Muchas de estas particularidades también se dan en el lenguaje de la ciencia (Halliday, 2004a; Halliday & Martin, 1993), cuya evolución mantiene ciertos paralelismos con la forma en que el lenguaje se desarrolla desde la infancia. Por ejemplo, la densidad léxica es un indicador de los textos más abstractos, es decir, más alejados del *aquí y ahora*, esta es una característica de los textos científicos. De manera similar, la densidad tiende a aumentar desde la niñez hasta la edad adulta, especialmente en el discurso escrito (Christie & Derewianka, 2008; Halliday, 2004b; Veel, 1997). Tanto el llegar a dominar el lenguaje académico como el hecho de componer un texto son tareas complejas y suponen un reto para los estudiantes (Castelló, Corcelles, Iñiesta, Bañales, & Vega, 2011; Llamas Saíz, 2014; Snow & Uccelli, 2009; Schleppegrell, 2012).

Las características que hemos visto hasta aquí tienen implicaciones directas en el campo de la investigación y de la enseñanza. Por un lado, el tema de la alfabetización sigue suscitando interés como objeto de investigación, ya que por su propiedad dinámica se abren constantemente nuevas perspectivas de análisis. Por otro lado, adquirir cierta competencia en gestionar, interpretar y utilizar el lenguaje académico, así como los diferentes modos semióticos es un proceso complejo para el alumno. Además, supone un reto para el profesorado que todos los alumnos lleguen a adquirir una alfabetización avanzada (Schleppegrell & Colombi, 2002). No obstante, para poder dotar a los estudiantes de las habilidades y conocimientos necesarios para hacer frente a la difícil tarea de utilizar el lenguaje para nuevos propósitos y en diferentes contextos, se necesita una descripción lingüística comprensiva de los registros que se dan en el entorno académico (Biber, 2006).

1.2 Los trabajos de investigación y los artículos de investigación: diferentes géneros, algunas similitudes

Nuestra investigación se centra en la escritura de investigación académica, en concreto, los trabajos de investigación de los estudiantes y los artículos de investigación de los científicos, que son considerados los más avanzados de todos los géneros académicos (Schleppegrell, 2001), ya que son el resultado de un proceso complejo (Swales, 2004). Aunque los trabajos de investigación de los estudiantes y los artículos de investigación no pertenecen al mismo género, tal como explicamos más adelante, el objetivo de la mayor parte de la escritura académica es que los estudiantes desarrollen la capacidad de escribir profesionalmente como, por ejemplo, en los artículos de investigación (Conrad, 1996; Gardner & Nesi, 2013; Russell & Cortés, 2012). Así lo afirman Gardner & Nesi (2013, p. 28): “The research article is a model for the sort of writing that students aspire to produce”.

Podríamos considerar también los trabajos de investigación de los estudiantes de bachillerato y los de universidad como géneros diferenciados, ya que el contexto y situación no son los mismos en el instituto que en la universidad (Giltrow, Gooding, Burgoyne & Sawatsky, 2005). Las expectativas y demandas en un nivel y otro son también diferentes, por ejemplo, en bachillerato se espera que el conocimiento sea reproducido, mientras que en la universidad se espera un pensamiento más analítico y crítico en relación a la información recibida (Carlino, 2003). Aun reconociendo que existen ciertas divergencias entre los textos producidos en un nivel académico y en otro, hemos decidido considerarlos como un mismo género, por la perspectiva de desarrollo en la que nos basamos y por las características comunes que tienen los trabajos de investigación de los estudiantes de ambos niveles. Estas características comunes también se explican en oposición al artículo de investigación. A continuación explicamos esta cuestión con más detalle.

Como ya hemos mencionado, los textos escritos por los estudiantes, tales como los trabajos de investigación, se diferencian de los artículos de investigación en varios aspectos. Los textos producidos por los estudiantes varían no solo en la audiencia, generalmente son leídos solo por el profesorado, sino que también tienen unos propósitos distintos y requieren de unas habilidades y conocimientos diferentes que los artículos de investigación. Mientras los textos de estudiantes tienen como objetivo demostrar la adquisición de las habilidades requeridas en determinados niveles académicos, los artículos de investigación tienen como finalidad informar de nuevos experimentos y persuadir al lector de la validez y la importancia de los nuevos hallazgos (Gardner & Nesi, 2013). A pesar de estas diferencias, también hay algunas similitudes entre ambos textos. Ambos géneros son el resultado de un proceso de investigación. Ambos incluyen unas preguntas de investigación, referencias a otras investigaciones y posible relevancia para la disciplina. Además, los trabajos de investigación comparten la estructura fija de los artículos de investigación en sus niveles de organización macroestructural, es decir, los apartados de Introducción, Método, Resultados y Discusión (IMRD), que son el reflejo del proceso de investigación y que dan coherencia en el nivel retórico.

Estudiantes y profesionales han compartido una forma tipificada de actuar en el proceso de investigación, es decir, que han escrito sus investigaciones (Mauranen, 1993). Por lo tanto, en ambos casos, la etapa final del proceso es un producto escrito que tiene como finalidad comunicar el conocimiento. Por consiguiente y debido a sus similitudes, hemos decidido analizar los trabajos de investigación y compararlos con los artículos de investigación, ya que consideramos que esto nos puede ofrecer una visión de la evolución de la escritura en unas disciplinas específicas.

Los enfoques basados en el género utilizan el criterio de *finés comunicativos* para la identificación de géneros (Bhatia, 1993; Swales, 1990). El artículo de investigación es una de las maneras de comunicación para los académicos, otras son presentaciones en congresos, conferencias, etc., para poder progresar en el conocimiento y comunicar sus hallazgos.

El artículo de investigación es también un medio que permite a los académicos noveles integrarse en diferentes comunidades disciplinarias (Bazerman, 1988; Hyland, 2000). Al unirse a una comunidad de práctica de investigación, los estudiantes deberían llegar a dominar no solo las características léxico-gramaticales, sino también las características discursivas que se dan en el género de la investigación (Berkenkotter, Huckin & Ackerman, 1991; Schleppegrell, 2001; Snow & Uccelli, 2009). Así lo demuestran diferentes estudios en lingüística textual que han identificado las características comunes de los artículos de investigación. Ya hemos mencionado que el registro académico se caracteriza por presentar una gran diversidad léxica. Pues bien, debido a que los artículos de investigación tienen un enfoque informativo y un gran porcentaje de sustantivos, uno esperaría también encontrar en ellos una alta diversidad léxica. Sin embargo, no siempre es así, por ejemplo, en artículos de investigación de ecología hay una gran repetición léxica, porque esto da cohesión al texto y asegura la exactitud del significado (Biber & Finegan, 1994; Conrad, 1996; Halliday & Hasan, 1976). En cambio, la densidad léxica, que es una particularidad del lenguaje académico, también está presente en los artículos de investigación, como, por ejemplo, en biomedicina (Dubois, 1982; Guillén Galve, 1998; León & Divasson, 2006).

En cuanto a la complejidad sintáctica, estudios basados en corpus de artículos de investigación en cuatro disciplinas (ciencia, educación, psicología e historia) han demostrado que la complejidad sintáctica se asocia con sintagmas nominales complejos en lugar de con cláusulas dependientes incrustadas (Biber, Gray & Poonpon, 2011).

Previamente, hemos mencionado que el posicionamiento discursivo en la escritura académica suele ser impersonal y objetivo. En cambio, hay un amplio consenso en que el discurso académico especializado es dialógico e interactivo (Bakhtin, 1981; Bazerman, 2004; Flowerdew, 2014; Swales, 2004). Es por esto por lo que en los textos académicos escritos, tales como artículos de investigación, los escritores tratan de presentar sus argumentos de manera dialógica.

Podríamos decir que la investigación científica sigue en un sentido amplio dos pasos: el primero es un proceso de construcción del conocimiento; y el segundo es un modo discursivo de argumentación, que está ligado a los objetivos epistémicos valorados por la disciplina, esto es, cómo se genera y se valida el conocimiento (Sandoval & Reiser, 2004). Este último paso se relaciona con las características retóricas y la destreza en las relaciones interpersonales. El escritor se identifica con una comunidad discursiva particular, por lo que utilizará recursos textuales pertenecientes a esa comunidad. Por lo tanto, la identidad social (por ejemplo, como investigador o autor académico) se construye discursivamente. Los estudiantes deberían manifestar una voz de autoría si quieren comprometerse con su comunidad de práctica.

La voz de autoría o discursiva se construye por medio de alinear el propio trabajo con ciertos autores, textos y, en especial, por la elección de formas particulares del discurso (Berkenkotter & Huckin, 1995; Ivanic, 1998; Lillis, 2001; Nelson & Castelló, 2009). Esto se suele lograr mediante la función metadiscursiva interaccional, es decir, a través de los marcadores lingüísticos que señalan la voz textual del escritor para involucrarse con su texto, la postura (*stance*) y los marcadores que utiliza el escritor para involucrar al lector, la afiliación (*engagement*) (Hyland, 2008a).

Aunque escribir de forma efectiva, presentar una voz de autoría y dialogar con la audiencia, sea relevante para lograr el éxito académico y el desarrollo profesional, a los estudiantes les resulta difícil emplear adecuadamente los recursos de interacción de una comunidad disciplinaria (Cortés, 2002; Henderson & Barr, 2010; Hyland, 2008b; Nelson & Castelló, 2012).

Para hacer frente a este reto, es útil que los escritores noveles tomen conciencia de los marcadores de interacción, es decir, cómo las interacciones escritor-lector se encarnan en un texto (Thompson, 2001). Esto es una necesidad, ya que cada vez es más habitual pedir a los estudiantes universitarios que adopten una posición explícita sobre las cuestiones sobre las que escriben, con el fin de construir su propia voz (Castelló et al., 2011a). Sin embargo, la mayoría de los manuales sobre escritura académica sugieren, en general, que en la escritura académica, no deben aparecer referencias al escritor o al lector. Asimismo, los manuales científicos no se ocupan de este tipo de marcadores discursivos (Mendiluce, 2005).

No podemos obviar el desarrollo de las habilidades de escritura explicadas desde los enfoques de la psicología cognitiva (Bereiter & Scardamalia, 1987; Kellogg, 2008), ya que la perspectiva adoptada en nuestro trabajo es la del *desarrollo tardío* del lenguaje. *Desarrollo* se refiere a la progresión más allá de la primera infancia, hacia la adolescencia y la edad adulta; o bien, la transición que se realiza de un contexto académico a otro; con *desarrollo tardío* nos referimos a la evolución del lenguaje más allá de la primera infancia (Berman, 2004). Bereiter & Scardamalia (1987), por ejemplo, mostraron que las habilidades de los aprendices diferían mucho de las de los escritores expertos en la escuela: la fase de *knowledge-telling* caracteriza el proceso seguido por un aprendiz, que simplemente explica lo que sabe sobre un tema; esto va acompañado de poca planificación y una revisión muy superficial del texto. Por el contrario, en la fase de *knowledge-transforming* el escritor considera la escritura como una tarea de resolución de problemas, en la cual se han de resolver dos problemas: uno está relacionado con el contenido (ideas) y el otro con los aspectos retóricos (la audiencia, la organización, etc.).

El modelo de Kellogg (2008) suma a estas fases una tercera, la *knowledge-crafting*, que se refiere a los niveles profesionales de la escritura. En esta fase, el escritor tiene que tener en cuenta las ideas de los autores consultados y el contenido de su propio texto, así como anticipar las diferentes interpretaciones de un lector imaginario. Además, también tiene que conocer las prácticas discursivas de la comunidad específica.

En definitiva, los estudiantes, cuando escriben un trabajo de investigación, tienen que hacer frente al pensamiento complejo, en el que se requiere una elección adecuada y una gestión de recursos lingüísticos para adaptarse al contexto del discurso (Nippold, 2004), ya que el significado del texto es el producto de lo que se escribe (contenido), de cómo se organiza el discurso (organización) y de cómo se comunica el conocimiento (interacción) (Snow & Uccelli, 2009).

1.3 La variación lingüística y discursiva en la escritura académica: las disciplinas de biología e historia

Aunque solemos utilizar el término *escritura académica* de forma general, la escritura académica es un registro que presenta una gran variación lingüística (Gray, 2015). Así, podemos distinguir, por ejemplo, la escritura académica de aprendices de la de profesionales. Asimismo, el lenguaje académico varía entre diferentes disciplinas (Gray, 2015; Conrad, 2001). Si bien una gran cantidad de investigación se ha llevado a cabo en analizar aspectos del registro en una disciplina, es menor el número de estudios que han comparado las características léxico-gramaticales entre diferentes disciplinas (e.g., Biber & Gray, 2010; Conrad, 2001; Fang, Schleppegrell & Cox, 2006; Gray, 2015; Hardy & Römer, 2013). Así los estudios de Biber & Gray (2010) y Fang, Schleppegrell & Cox (2006) han revelado que la escritura académica, independientemente de la disciplina, tiene una marcada extensión del grupo nominal, también se caracteriza por el uso de adjetivos y sintagmas preposicionales. Los otros estudios citados han utilizado la metodología multidimensional, propuesta por Biber (1988), que permite analizar los patrones lingüísticos que co-ocurren en un registro. Los resultados sugieren que hay diferencias en las dimensiones analizadas, por ejemplo, en la dimensión involucrada vs. informativa los textos de las disciplinas de filosofía y educación tienen una función comunicativa más involucrada que las de biología y física, que son más informativas.

Además, las características lingüísticas que reflejan una escritura orientada al proceso son más comunes en ciencias, en cambio en humanidades reflejan una escritura dependiente de la situación, es decir, hay más reflexión y síntesis de lecturas relacionadas (Hardy & Römer, 2013). Lo cierto es que el conocimiento en las distintas disciplinas se produce de acuerdo a determinadas normas, es decir, hay unas convenciones propias de cada disciplina para interactuar, comunicar y representar el conocimiento. Para relacionarlo con las dos disciplinas que conforman nuestro objeto de estudio, biología e historia, nos remitimos a las prácticas normativas de ciencias e historia señaladas por Moje (2008).

Así, por ejemplo, en la disciplina de las ciencias naturales, una práctica normativa es que los problemas que se investigan deben ser cuidadosamente definidos, estudiados de manera sistemática y repetidamente antes de poder hacer aserciones acerca de los fenómenos observados. En este sentido, se requieren pruebas empíricas que se derivan de un estudio experimental para poder hacer aserciones. En historia, por el contrario, las prácticas normativas son algo distintas. Los historiadores, al igual que los científicos, estudian los problemas sistemáticamente, pero los medios de obtención de evidencias y las formas que proporcionan las garantías para las aserciones difieren. El contexto temporal es una dimensión que un investigador de textos históricos debe conocer, descubrir o examinar en profundidad, para poder dar explicaciones verosímiles sobre los hechos. En consecuencia, el aprendizaje en las diferentes disciplinas requiere no solamente conocer estas prácticas normativas, sino también ponerlas en práctica (Berkenkotter, Huckin & Ackerman, 1991; Gee, 2001; Lemke, 1990; Moje, 2008; Schleppegrell, 2001; Snow & Uccelli, 2009).

La forma de citar, por ejemplo, varía según la disciplina, cada disciplina tiene sus propias convenciones. Así, por ejemplo, en historia, las citas que reúnen información original, como podrían ser entrevistas, documentos, etc., suelen presentarse como transcripciones literales de fragmentos de esta información. Para referenciar estas citas se pone una nota a pie de página (Estrada, 2013). En cambio, en biología, la convención es no utilizar transcripciones literales de fragmentos (Cañedo-Argüelles, 2013). Vemos, por lo tanto, que los textos presentan convenciones académicas en general, pero también propias de las disciplinas especializadas (Nelson & Castelló, 2012).

Por otra parte, según la *lingüística sistémico-funcional*, el lenguaje es esencialmente funcional. Toda experiencia humana se puede transformar en significado. Así cada signo sirve esencialmente a tres metafunciones simultáneamente, la *ideacional*, la *interpersonal* y la *textual*. La *ideacional* contiene recursos para crear contenido; la *interpersonal*, para interactuar; y la *textual* contiene recursos para organizar los textos (Halliday, 1978). Según este autor, las disciplinas científicas como la biología privilegian la función *ideacional*, mientras que las humanísticas se centran más en la función *interpersonal* (Halliday, 2002). Según la revisión bibliográfica que hemos realizado, son pocos los estudios que han centrado su análisis en la comparación de publicaciones y escritos producidos por estudiantes en las disciplinas de biología e historia, excepto Cortes (2004), que ha estudiado las colocaciones (*lexical bundles*) que cumplen con las metafunciones descritas por Halliday (1978). Sus resultados indican que en las publicaciones de biología hay colocaciones con una función de postura, en concreto, las que cumplen con una función atenuadora, mientras que en historia no se utilizan colocaciones con esta función. Asimismo, el estudio de Cortes (2004) muestra diferencias notables en la frecuencia y uso de las colocaciones al comparar las publicaciones con los escritos de los estudiantes. Sus resultados sugieren que la simple exposición a las colocaciones que cumplen una determinada función, a través de la lectura académica, no da lugar a la adquisición de estas por parte de los estudiantes. Por otra parte, Hyland (2005) ha analizado la postura y la afiliación en artículos de investigación, en ocho disciplinas distintas, entre ellas sociología y microbiología. Sus resultados muestran que las expresiones de postura y afiliación se dan en todas las disciplinas y que los marcadores de postura, sobre todo los atenuadores, son proporcionalmente más frecuentes que los de afiliación.

En este apartado hemos visto las diferencias del registro y el metadiscurso entre disciplinas. A continuación nos detenemos en el contexto académico y cultural de los textos que conforman nuestro corpus.

1.4 El contexto académico y cultural del corpus TARBUC

En Cataluña, los estudiantes de secundaria de 17/18 años de edad deben completar un trabajo de investigación como requisito académico para acceder a su título de bachillerato. Este trabajo, que es el resultado de una investigación personal guiada, debe presentarse en forma escrita, lo que permite a los estudiantes desarrollar las habilidades propias de la investigación, la discusión y las habilidades expresivas, similar a lo que ocurre en las comunidades de discurso científico. Por otra parte, a nivel universitario, los estudiantes deben presentar un trabajo de investigación con la típica estructura retórica (IMRD) del artículo de investigación, al finalizar sus estudios de máster. Así vemos que los alumnos tienen que enfrentarse, no solo a las habilidades de investigación, sino también al lenguaje especializado de su disciplina ya desde la educación secundaria.

Cataluña es una comunidad bilingüe, que tiene el español y el catalán como lenguas oficiales. El catalán es la lengua de escolarización de toda la educación obligatoria y es estructuralmente muy similar a la lengua española.

En el marco de la lingüística cognitiva, los estudios contrastivos sugieren que la tipología gramatical tiene un efecto sobre la tipología retórica (Slobin, 1996). Por otra parte, desde la perspectiva de la retórica contrastiva, la escritura académica se ve como culturalmente determinada y se pueden distinguir dos grandes estilos de escritura: las culturas orientadas a la forma (por ejemplo, inglés) y las culturas orientadas al contenido (por ejemplo, alemán, español, catalán) (Clyne, 1994; Connor, 2002; Cuenca, 2003; Hinds, 1987; Kaplan, 1966).

En el caso de las culturas orientadas a la forma se intenta facilitar el trabajo a los lectores; por lo tanto, la principal preocupación es cómo transmitir conocimiento. Por otro lado, en el caso de las culturas orientadas al contenido se considera que el lector es responsable de la interpretación; por lo tanto, la principal preocupación es la cantidad de conocimientos transmitidos (Cuenca, 2003).

En un estudio comparativo basado en un corpus de textos expositivos en inglés, español y catalán, Cuenca (2003) ha analizado los marcadores del discurso (reformuladores) y la gramática. Sus resultados muestran que la estructura de los marcadores es similar en las tres lenguas, aunque hay diferencias en la variedad de formas para expresar reformulación. Así sugiere que ciertas formas son elegidas por su función de discurso, ya que hay una conexión entre el discurso y la sintaxis.

Desde una perspectiva de desarrollo, esto se relaciona con la idea básica de que el progreso en el conocimiento lingüístico implica la selección de una variedad de formas lingüísticas para expresar una función particular del discurso (Berman, 2004). Por otra parte, las diferencias observadas entre la gramática de estas tres lenguas podrían estar relacionadas con los dos estilos de escritura. Con respecto a la sintaxis, las culturas orientadas a la forma utilizan frases cortas y repiten términos léxicos; mientras que en las culturas orientadas al contenido se suelen utilizar oraciones complejas, se es menos conciso y se evita repetir elementos léxicos, ya que esto es una señal del estilo formal (Cuenca, 2003). Estas características son también una característica de la prosa académica del portugués (Bennett, 2010).

Actualmente, la lengua para la difusión de resultados de investigación, especialmente en ciencias biomédicas, es el inglés (Swales, 2004), por lo que enseñar la escritura académica, que suele seguir el modelo angloamericano, a los escritores continentales es, cuanto menos, complejo, debido precisamente a las diferentes tradiciones de escritura (Rienekker & Stray, 2003). En sus etapas iniciales, los estudiantes deben escribir sus trabajos de investigación en catalán, mientras que en los niveles avanzados (por ejemplo, en tesis doctorales, o como investigadores profesionales) se requiere el inglés. Por lo tanto, también es un desafío para los estudiantes adaptarse a la tradición de escritura angloamericana. En el contexto académico español, por ejemplo, Martín, Rey-Rocha, Burgess & Moreno (2014) afirman que la comunicación en inglés es sin duda un obstáculo para los académicos de las ciencias médicas y estos se preocupan por aprender estrategias para resaltar de forma adecuada la contribución de sus hallazgos.

En resumen, cuando los estudiantes tienen que redactar un trabajo de investigación académico, tienen que haber asimilado unas formas lingüísticas específicas para comunicar el conocimiento y también las prácticas y normas de una comunidad disciplinaria específica (Bawarshi & Reiff, 2010). Esto implica desarrollar unas habilidades culturales, cognitivas y lingüísticas, sobre todo en los estudiantes cuyo idioma materno no es el inglés, ya que tienen que encajar con la tradición de escritura académica angloamericana (Hyland & Hamp-Lyons, 2002; Rienecker & Stray, 2003). En consecuencia, proporcionar evidencias para explicar cómo los estudiantes utilizan las características léxico-gramaticales y los marcadores metadiscursivos del discurso es crucial para poder mejorar las habilidades de escritura.

A pesar de la afirmación de Aull & Lancaster (2014) que pone de relieve la importancia de que los estudiantes aprendan la escritura específica de la disciplina lo antes posible, ha habido pocos estudios basados en corpus que hayan analizado el desarrollo de la función metadiscursiva interaccional en diferentes niveles académicos.

Por ejemplo, la postura se ha examinado en textos universitarios y en artículos publicados en las disciplinas de biología e historia (Cortes, 2004), en ensayos de primer y último año de universidad (e.g., Aull & Lancaster, 2014), en textos de historia en secundaria (e.g., Coffin, 2002), y en la sección introductoria de disertaciones de primer año de universidad en L2 (e.g., Hood, 2004). Por otra parte, el posicionamiento del escritor y la representación del lector también ha sido estudiado desde la psicología cognitiva (Bereiter & Scardamalia, 1987; Kellogg, 2008), en artículos publicados en diferentes disciplinas (e.g., Kuo, 1999; McGrath & Kuteeva, 2012) y en diferentes culturas (e.g., Mur, 2007; Vázquez, 2010).

La mayoría de estos estudios se han centrado en textos escritos en inglés, por lo que hay poca investigación que examine desde una perspectiva evolutiva el género de investigación producido en ambientes académicos cuyo idioma de instrucción no sea el inglés.

Respecto al registro, estudios previos han analizado algunas características léxico-gramaticales, por ejemplo, la diversidad léxica y la densidad léxica, desde una perspectiva de desarrollo, en el género narrativo y expositivo, en lenguas como el inglés, sueco, hebreo, islandés (e.g., Strömqvist et al., 2002), el catalán (e.g., Cutillas, Tolchinsky, Rosado & Perera, 2014) y en bilingües alemanes-turcos (e.g., Woerfel & Yilmaz, 2011).

La complejidad sintáctica, por ejemplo, se ha estudiado en textos argumentativos escritos por estudiantes universitarios (estudiantes de primer y cuarto año) en español desde una perspectiva contrastiva, inglés-español (Neff & Prieto, 1994).

En consecuencia, suponemos que el análisis basado en corpus de las características léxico-gramaticales y los marcadores metadiscursivos de interacción, que se dan en los trabajos de investigación nos puede ilustrar aspectos del continuum, de principiante a experto que participa en una comunidad discursiva específica, como es la de biología e historia.

Por lo tanto, esta tesis contribuye a caracterizar, desde un enfoque evolutivo, que tiene en cuenta factores cognitivos y socio-culturales a la hora de interpretar los resultados, el desarrollo lingüístico de un género especializado, como es el trabajo de investigación. Para ello adoptamos básicamente el marco de la *teoría sistémico-funcional*, porque supone una herramienta útil para el análisis de textos, ya que permite la conexión de opciones léxico-gramaticales con sus contextos sociales (Schleppegrell, 2000).

CAPÍTULO 2

2. Objetivos

De los argumentos aducidos hasta aquí, consideramos justificado el abordar el estudio de la escritura académica a través del trabajo de investigación. El supuesto que subyace en el presente estudio es que los estudiantes de la universidad han desarrollado una mayor alfabetización académica, por lo tanto, las características léxico-gramaticales y sus formas discursivas podrían ser las propias del registro de su comunidad de práctica.

Esta investigación tiene como objetivo principal caracterizar el desarrollo de la escritura académica del género de investigación en el continuum que va desde la educación secundaria y universitaria, hasta la publicación profesional.

Para el estudio 1, en el que comparamos transversalmente un corpus de trabajos de investigación de bachillerato y de final de máster universitario en las disciplinas de biología e historia, escritos en catalán, establecemos dos objetivos generales de los que se derivan unos objetivos específicos:

- (1) Estudiar el registro académico en los trabajos de investigación por medio de un análisis cuantitativo de las siguientes características léxico-gramaticales: (i) la diversidad léxica, (ii) la densidad léxica y (iii) la complejidad sintáctica, esta con dos medidas diferentes. La diversidad léxica mide la variedad de uso del vocabulario. La densidad léxica indica la riqueza textual e informativa. La complejidad sintáctica refiere a la sofisticación y a la variedad de las formas sintácticas.

- 1.1 Examinar las características léxico-gramaticales, que son propias del registro académico. Por lo tanto, examinar si el factor nivel educativo influye.

- 1.2 Analizar si hay variaciones de estas características entre disciplinas (biología e historia).

(2) Analizar, cuantitativamente y cualitativamente, los recursos retóricos que utilizan los estudiantes, en concreto, los que tienen una función metadiscursiva interaccional, por medio de: (i) los marcadores lingüísticos que señalan la voz textual del escritor para involucrarse con su texto (*stance*), y (ii) los que utiliza el escritor para involucrar al lector (*engagement*).

2.1 Identificar los marcadores metadiscursivos interaccionales que utilizan los estudiantes y su distribución por nivel educativo.

2.2 Examinar si hay variaciones entre disciplinas.

Para el estudio 2, en el que analizamos longitudinalmente un corpus de trabajos de investigación de final del máster universitario, en catalán, y de publicaciones en inglés en la disciplina de biología y escritos por los mismos sujetos, examinamos el registro y el metadiscurso.

(3) Estudiar el registro académico en los trabajos de investigación y en los artículos publicados, por medio de un análisis cuantitativo de las siguientes características léxico-gramaticales: (i) la diversidad léxica, (ii) la densidad léxica y (iii) la complejidad sintáctica, esta con dos medidas diferentes.

3.1 Examinar las características léxico-gramaticales propias del registro académico. Por lo tanto, determinar si hay variaciones entre los dos géneros.

(4) Analizar, cuantitativamente y cualitativamente, los recursos retóricos que aparecen en los trabajos de investigación y los artículos publicados, en concreto, los que tienen una función metadiscursiva interaccional, por medio de: (i) los marcadores lingüísticos que utiliza el escritor para involucrarse con su texto (*stance*), y (ii) para involucrar al lector (*engagement*).

4.1 Identificar los marcadores metadiscursivos interaccionales de los trabajos de investigación universitarios y de las publicaciones y determinar si hay variaciones en la frecuencia de uso.

Por lo tanto, el registro lingüístico y los recursos retóricos se examinan en función de los factores objeto de estudio, el nivel académico y la disciplina, para intentar responder a las siguientes preguntas:

1. ¿Qué características léxico-gramaticales, a través de los niveles, pueden revelar el desarrollo del registro académico en los artículos de investigación? ¿Hasta qué punto pueden, las diferencias esperadas, ser significativas?
2. ¿Qué diferencias, si las hay, se observan entre las disciplinas de biología e historia?
3. ¿Cómo se distribuye la función metadiscursiva de interacción a través de los niveles?
4. ¿Qué diferencias, si las hay, se observan entre las disciplinas de biología e historia?
5. ¿En la distribución de esta función en la disciplina de biología, hay similitudes entre el nivel universitario y las publicaciones?

2.1 Predicciones

Para los objetivos del estudio 1, establecemos las siguientes predicciones.

- 1.1 Respecto al registro, partimos de la asunción que hay diferencias en el registro en función de la edad y el nivel de alfabetización (e.g., Bar-Ilan & Berman, 2007; Ravid & Tolchinsky, 2002), por lo tanto, esperamos encontrar diferencias entre niveles académicos. Esperamos encontrar una mayor diversidad léxica, densidad léxica y complejidad sintáctica en los textos universitarios.
- 1.2 El registro varía según la disciplina (e.g., Biber & Conrad, 2001), esto parece ser más evidente entre ciencias e historia (Halliday & Martin, 1993), ya que las disciplinas científicas privilegian la función *ideacional*, mientras que las humanísticas la *interpersonal* (Halliday, 2002). En consecuencia, prevemos encontrar diferencias de registro entre las disciplinas de biología e historia.
- 2.1 En relación a los recursos retóricos, suponemos que habrá un patrón evolutivo en cuanto a la frecuencia de aparición de los marcadores interaccionales entre niveles académicos (Aull & Lancaster, 2014), en nuestro caso, entre bachillerato y universidad, ya que partimos del supuesto que hay un desarrollo en la expresión de la postura influenciado por factores como el cognitivo (Berman, 2007). Por lo tanto, esperamos encontrar un uso mayor, sobre todo, de atenuadores en los textos universitarios.
- 2.2 Cada disciplina utiliza los marcadores de la postura y la afiliación de forma diferente, de acuerdo a sus convenciones (Hyland, 2005), por lo tanto, esperamos encontrar diferencias significativas entre las disciplinas de biología e historia.

Para los objetivos del estudio 2, realizamos las siguientes predicciones.

- 3.1 En cuanto al registro, esperamos encontrar diferencias significativas entre trabajos y publicaciones, ya que estamos analizando dos géneros diferentes, tipologías de lenguas diferentes, y dos tradiciones culturales de escritura, la orientada a la forma y la orientada al contenido (e.g., Cuenca, 2003).
- 4.1 Respecto a los marcadores interaccionales, suponemos que habrá diferencias significativas, ya que los géneros que analizamos son diferentes (Hyland, 2005).

CAPÍTULO 3

3. Estructura general de la tesis

En las páginas que siguen, presentamos, en primer lugar, los antecedentes teóricos y empíricos (capítulo 4) de los que partimos y que tratan las siguientes temáticas: *la modalidad oral y escrita, la alfabetización académica, el género, la escritura, la escritura académica, el registro y el metadiscurso*. En segundo lugar, presentamos la metodología (capítulo 5), es decir, los postulados teóricos con los que abordamos el estudio. En el capítulo 6, describimos el método, *los motivos para seleccionar los textos, el procedimiento de recogida de los textos, las categorías de análisis, el procesamiento y herramientas de análisis*. En el capítulo 7 presentamos los resultados del estudio 1, posteriormente el resumen de estos resultados (capítulo 8) y la discusión (capítulo 9). El capítulo 10 recoge los resultados del estudio 2, el resumen de estos resultados (capítulo 11) y su discusión (12). Finalmente en el capítulo 13, a modo de cierre presentamos las conclusiones, las limitaciones de nuestra investigación y líneas futuras.

CAPÍTULO 4

4. Antecedentes teóricos y empíricos

Este capítulo ofrece una visión general de los antecedentes teóricos, así como de los estudios empíricos más relevantes para nuestra investigación. En el primer apartado (4.1) nos referimos brevemente a la diferencia entre modalidad oral y escrita, ya que nuestra investigación tiene en la modalidad escrita su objeto de estudio. En el apartado 4.2, justificamos la necesidad de atender a las diferentes dimensiones que componen la alfabetización académica, porque son las que le confieren su complejidad. En el tercer apartado (4.3) discutimos las diferentes nociones de género, por ser este un concepto estrechamente vinculado a la cultura escrita. La escritura y sus funciones las exponemos en el apartado 4.4 y 4.5. Finalmente, en el apartado 4.6 y 4.7 establecemos el marco analítico para nuestra investigación, además de repasar estudios previos que se han basado en este marco. Explicamos el análisis del registro (4.6), el enfoque en el que nos basamos para la investigación de la variación lingüística; y en el apartado 4.7 nos centramos en el metadiscurso, perspectiva para examinar la interacción social en textos académicos.

4.1 La modalidad oral y escrita

El título de este sub-apartado podría sugerir el debate imperante durante mucho tiempo sobre la dicotomía entre oralidad y escritura. No es esta nuestra intención al abordar este apartado, ya que una división binaria de estas modalidades asume diferentes tipos de sociedades, por ejemplo, las iletradas y letradas de forma simplista, o lo que se conoce por *Great divide theory*. Más bien, consideramos que una revisión breve de la oralidad y la escritura, tanto de sus propiedades distintivas como desde una perspectiva diacrónica, ayudan a comprender mejor nuestro objeto de estudio, la escritura académica.

A primera vista parece evidente que no hay diferencias entre la oralidad y la escritura, todo lo que se puede decir se puede escribir y viceversa (Olson, 2009a). De hecho, las dos modalidades se influyen mutuamente (Ravid & Tolchinsky, 2002). Además, hay quien sugiere que estamos en una cultura escrita secundaria, en la que la modalidad oral que nos llega desde los diferentes medios, por ejemplo, la televisión, está basada en la escritura y que los escritos que leemos están influenciados por las formas orales y mezclados con elementos visuales, sonoros, etc. (Ong, 1982; Lunsford, 2006). Pero según Kellogg (2008), si bien adquirimos el lenguaje oral de forma natural y relativamente rápida, a la edad de 5 años la oralidad está ya bastante desarrollada; la capacidad para escribir un texto coherente requiere más esfuerzo y un desarrollo mucho más prolongado. “Literacy, on the other hand, is a purely cultural achievement that may never be learned at all” Kellogg (2008, p. 2).

Asimismo, un análisis más profundo sugiere que en la modalidad oral y escrita hay un potencial de variación lingüística atribuible a las circunstancias de producción (Biber, 2009). En la modalidad oral, casi cualquier interacción comunicativa incluye los interlocutores y el contexto. La circunstancia de producción es inmediata, la comunicación es un proceso efímero y dinámico, apenas se puede planificar ni corregir cuidadosamente. En cambio, en la modalidad escrita, no hay un contexto situacional compartido, la comunicación entre emisor y receptor es en diferido, por lo que no pueden negociarse los significados, ni los contextos. La circunstancia de producción permite la planificación y revisión y esto implica una codificación lingüística diferente a la modalidad oral (Cantero & Mendoza, 2003). Algunos géneros especializados, como los trabajos de investigación, se basan en la modalidad escrita; por consiguiente, si uno quiere participar en una disciplina específica debería aprender a leer y escribir esos textos especializados (Olson, 2009b).

No podemos hablar de modalidad escrita sin hacer referencia a la lectura. El hecho de que la modalidad escrita sea un producto estable y sinóptico, es decir, que tenga una disposición gráfica que facilita la visión aislada y al mismo tiempo conjunta de las formas lingüísticas, tiene consecuencias cognitivas. La forma lingüística se vuelve visible para un sujeto alfabetizado, lo que le permite reflexionar sobre las propiedades del lenguaje, es decir, posibilita desarrollar la capacidad metalingüística (Olson, 2009b).

Pero además, facilita diferenciar el sentido literal del sentido intencional de las palabras, por lo que el texto escrito pone a disposición múltiples interpretaciones, tal como explican Teberosky & Tolchinsky (1994, p. 10) citando a Lévinas (1982), el lector puede ir “ más allá del texto: el primer sentido se *dice* en el texto, el segundo se *quiere decir*, el tercero se *puede querer decir* “. Tal como vemos, el acto de leer implica interpretar los diferentes sentidos del texto y esto es una acción, entendida esta como una actividad donde se manifiesta nuestra identidad, nuestra singularidad (Arendt, 1958, p. 179). Ouaknin (1994), influenciado por las ideas de estos dos filósofos, Lévinas y Arendt, concluye que la lectura, en tanto que interpretación, no es sino un reflejo de la propia singularidad, por lo que es productiva y creativa (Rodríguez, 2012).

Justamente la historia de la lectura es en parte una historia de una interpretación soslayada. Por ejemplo, los cristianos interpretaron las antiguas escrituras de un modo diferente al que lo hicieron los hebreos, los protestantes hicieron lo propio con las escrituras católicas (Simpson, 2007 *apud* Olson, 2009a). La historia de la humanidad se puede explicar a través del lenguaje y las acciones, como puede ser la escritura (Goody & Watt, 1968), ya que es en el lenguaje y la acción donde las personas revelan sus identidades y aparecen, por tanto, en el mundo (Arendt, 1958, p. 179). A través del lenguaje y la escritura se originaron formas de organizaciones sociales complejas. Antes de la aparición de la escritura se transmitía la herencia cultural, las formas de actuar del grupo, así como las ideas o creencias de cómo interpretar y actuar en el mundo, es decir, la *Weltanschauung*, de forma oral. Lo que se consideraba importante para el grupo social se transmitía oralmente, el resto se iba olvidando. Con el tiempo y el nacimiento de los diferentes sistemas de escritura, esta tradición oral se iría fijando (Goody & Watt, 1968), pero lo cierto es que las primeras manifestaciones escritas que tenemos servían únicamente para la contabilidad de la antigua ciudad de Uruk (la bíblica Erech en Mesopotamia, actual Irak) (Tusón, 2006). Ya ha pasado cierto tiempo desde que la escritura se utilizaba casi exclusivamente para organizar la vida en grandes asentamientos, y desde entonces ha contribuido a cambios sociales y culturales: "la escritura es una revolución que revoluciona" (Tusón, 2006, p. 74).

Así, por ejemplo, la invención de la imprenta permitió la propagación de la cultura y, la expansión de la posibilidad de lectura individual, lo que facilitó los avances en todas las áreas de conocimiento y en definitiva, la democratización en el acceso a la cultura (Tusón, 2006).

Esta herencia junto con otros factores sociales, económicos y culturales derivó en un incremento, a partir de 1960 en Estados Unidos y Gran Bretaña, no solamente del número de universidades, sino sobre todo del número de estudiantes que se decidían por una educación superior. Esta educación, que dejaba de ser privilegio de las clases altas, evidenció que muchos estudiantes presentaban pocas habilidades para la escritura académica. Para hacer frente a este reto se crearon programas de escritura generales, por ejemplo el conocido como WAC (*Writing Across the Curriculum*) y más específicos de cada disciplina, el WID (*Writing in the Disciplines*), que han influenciado las perspectivas e investigación en escritura académica y permitieron el posterior desarrollo del movimiento llamado alfabetizaciones académicas ACLYTS (*Academic Literacies*) en los años 90, que se sitúa en la intersección entre la teoría y la práctica de la alfabetización (Castelló, 2014; Russell, Lea, Parker, Street & Donahue, 2009).

Como acabamos de ver, el tema de la escritura se puede abordar desde diferentes perspectivas, entre ellas la filosófica o la histórica. Aunque sean sumamente interesantes, estas se alejan de nuestro ámbito de estudio, por eso a continuación expondremos con detalle los aspectos de la escritura relacionados con el ámbito académico, es decir, la alfabetización.

4.2 La alfabetización: una perspectiva multidimensional

El término tan corriente en inglés *literacy* no tiene una traducción unívoca y no está exenta de cierta controversia. Algunos autores la traducen como *literacidad* (e.g., Cassany & Castellá, 2010), *cultura escrita* (e.g., Ferreiro, 1999) o *alfabetización académica* (e.g., Carlino, 2003). Esta última autora en una primera definición del concepto de *alfabetización académica* toma como punto de vista a los estudiantes, al proceso de aprendizaje. Así *alfabetización académica*:

Señala el conjunto de nociones y estrategias necesarias para participar en la cultura discursiva de las disciplinas, así como en las actividades de producción y análisis de textos requeridas para aprender en la universidad. Apunta, de esta manera, a las prácticas de lenguaje y pensamiento propias del ámbito académico. Designa también el proceso por el cual se llega a pertenecer a una comunidad científica y/o profesional [...] (Carlino, 2003, p. 410).

En una posterior redefinición del concepto (Carlino, 2013), incluye en la perspectiva el proceso de enseñanza, la función de los docentes y las instituciones académicas, unos objetivos relacionados con una orientación socio-cultural. En este sentido, la alfabetización académica:

[..] Conlleva dos objetivos que, si bien relacionados, conviene distinguir: enseñar a participar en los géneros propios de un campo del saber y enseñar las prácticas de estudio adecuadas para aprender en él. En el primer caso, se trata de formar para escribir y leer como lo hacen los especialistas; en el segundo caso, de enseñar a leer y a escribir para apropiarse del conocimiento producido por ellos. De acuerdo con las teorías sobre aprendizaje situado y sobre géneros como acciones sociales, alfabetizar académicamente equivale a ayudar a participar en prácticas discursivas contextualizadas, lo cual es distinto de hacer ejercitar habilidades desgajadas que fragmentan y desvirtúan esas prácticas. Porque depende de cada disciplina y porque implica una formación prolongada, no puede lograrse desde una única asignatura ni en un solo ciclo educativo. Así, las “alfabetizaciones académicas” incumben a todos los docentes a lo ancho y largo de la universidad [...] Carlino (2013, p. 270).

Aunque compartimos la idea que se da en esta redefinición del concepto de alfabetización, para el propósito de nuestro trabajo adoptamos la primera definición (Carlino, 2003), es decir, la que se centra en el proceso de aprendizaje, las habilidades y estrategias necesarias para participar en el ámbito académico. En este sentido podemos equiparar el concepto de *alfabetización* con *cultura escrita* para incluir todos los sentidos de *literacy*, que se refiere a disponer de una gran variedad de recursos del lenguaje como la lectura, la escritura, los números y las imágenes para poder adaptarlos a los diferentes contextos sociales.

Tradicionalmente, se entendía la alfabetización como tener las habilidades para leer y escribir para fines concretos. Esta concepción ha sido superada, ya que los continuos cambios sociales y tecnológicos también imponen cambios en la alfabetización. Las personas han de crear y extraer el significado, ya no solo del lenguaje técnico y abstracto, sino también de medios visuales y electrónicos (Blue, 2003; Schleppegrell & Colombi, 2002). Por lo tanto, ya no tenemos una única alfabetización, más bien habría que hablar de múltiples alfabetizaciones o alfabetización multimodal (Carlino, 2003; Kress & Van Leeuwen, 1996; Kucer, 2001; Street, 2012). En este contexto de cambio, se concibe la alfabetización como un proceso dinámico, una práctica social que evoluciona en la sociedad y en el individuo (Barton, Hamilton & Ivanic, 2000; Schleppegrell & Colombi, 2002).

Desde la perspectiva sociocultural, la alfabetización se entiende como un conjunto de prácticas sociales, situadas en un contexto social, y estas prácticas se observan en eventos (actividades), que son mediados por textos escritos (Barton, Hamilton & Ivanic, 2000, Bazerman, 1989, Gee, 1987). Asimismo, desde el funcionalismo también se entiende la alfabetización dentro de un contexto sociocultural, “la realidad social es un constructo semiótico” (Halliday, 1978, p. 2). Se trata de una actividad social en que el discurso y el contexto crean conjuntamente el significado (Lemke, 2002).

Kucer (2001) entiende la *alfabetización* como multidimensional, formada por cuatro dimensiones (lingüística, cognitiva, sociocultural y de desarrollo); y el estar alfabetizado significa aquella capacidad de controlar de forma efectiva, eficiente y simultánea el sistema lingüístico y otros sistemas de signos, la dimensión cognitiva, sociocultural y el desarrollo del lenguaje escrito. La dimensión lingüística se refiere al sistema comunicativo a través del cual se transmite el significado; la cognitiva a las estrategias y procesos para crear significado; la sociocultural a las identidades y prácticas sociales, cómo los diferentes grupos usan la alfabetización para diferentes propósitos y de forma diferente; y la de desarrollo refleja si el escritor/lector domina las dimensiones en cada contexto, su proceso de aprendizaje, “becoming literate rather than being literate” (Kucer, 2001).

Nos hemos basado en estas dimensiones de Kucer (2001) y hemos añadido una más, la dimensión multimodal. Asimismo, hemos adaptado el contenido al contexto académico en el que se enmarca nuestro trabajo, es decir, hemos adaptado las dimensiones a la *alfabetización avanzada*, que refiere a la manera de crear significado a partir de la adolescencia, es decir, en los niveles de bachillerato y de universidad (Schleppegrell & Colombi, 2002). El término adolescente tiene un significado diferente según las culturas. Además la adolescencia no es únicamente un período cronológico en el ciclo de vida, tampoco es solamente una etapa de desarrollo hacia la vida adulta (Christenbury, Bomer & Smagorinsky, 2009). Por ejemplo, desde una perspectiva sociocultural se ha querido ir más allá del determinismo biológico definiendo la adolescencia como organizada simultáneamente en el pasado, presente y futuro, lo que destaca su complejidad (Lesko, 2001). A pesar de esto, lo cierto es que comúnmente se acepta la etapa que va más o menos desde los 11 hasta los 20 años como la adolescencia (Moje, Overby, Tysvaer, & Morris, 2008). La alfabetización adolescente se entiende como la capacidad de leer, interpretar, criticar y producir el discurso de una área disciplinaria, por lo que a partir de la educación secundaria se espera que los alumnos participen en los discursos de las disciplinas, para así poder aprender las convenciones de la producción de los conocimientos disciplinarios (Moje, Overby, Tysvaer, & Morris, 2008; Schleppegrell & Colombi, 2002; Wineburg & Martin, 2004). A continuación presentamos las dimensiones que conforman esta alfabetización.

4.2.1 La dimensión lingüística

Según Kucer (2001) los escritores/lectores conocen y utilizan los sistemas del lenguaje para crear o decodificar el significado escrito. Estos van desde el más local, grafémico, grafofonémico, ortográfico, morfémico, sintáctico, semántico, estructura del texto, género, hasta el más global, pragmático. Todos estos sistemas interactúan en la construcción del texto. Como nuestro objeto de estudio es un tipo de texto, es decir, el trabajo de investigación, a continuación nos centraremos en el texto.

Durante muchos años se consideraba la unidad básica de la comunicación la oración y asimismo era la unidad mayor analizable por parte de los lingüistas. Hartman (1971) causó cierta revolución al argumentar que era el texto y no la oración la unidad básica de la comunicación (Heinemann & Heinemann, 2002). De hecho, el resultado de cualquier evento de alfabetización es el texto (Kucer, 2001). Este autor cita la definición de *texto* de Beaugrande (1980): una unidad con sentido del lenguaje con intención comunicativa; y comenta que Gee (1990) hace una definición similar de *discurso*. Lo cierto es, que a menudo, los términos *texto* y *discurso* se utilizan de manera intercambiable e incluso como sinónimos, ya que la delimitación de cada concepto ha sido objeto de discusión (Bonilla, 1997). Según la posición teórica adoptada, el término *texto* se refiere a los aspectos lingüísticos, mientras que el término *discurso* se refiere a algo más general, que incluye los aspectos de la pragmática (Atienza, 1992).

La lingüística textual tiene en el texto (en forma oral o escrita) la unidad fundamental de análisis del lenguaje humano, es decir, excede la gramática de la oración y contempla como unidad de análisis secuencias de oraciones o fragmentos más grandes de textos (Dressler, 1978). Con esta definición se va más allá del paradigma que dice que la oración representa la unidad lingüística mayor con sentido. El texto, como objeto de análisis, es una entidad multidimensional porque aglutina la dimensión funcional, situacional, temática y gramatical de la lengua (Heinemann, 2000). Pero esto no supone considerar el *texto* como autónomo, más bien, es un *artefacto en actividad*. Cualquier escritor consulta y utiliza otros recursos que trascienden el momento concreto de la transcripción.

Así lo explica Prior (2006): “Text, as artifacts in activity, and the inscription of linguistic signs in some medium are parts of stream of mediated, distributed and multimodal activity” (p. 58). Cada texto se relaciona con otros textos. El escritor utiliza recursos de otros textos y estos, a su vez, están relacionados con los textos anteriores. La escritura es el producto de su contexto social, no es una actividad autónoma, toda escritura es el producto de una colaboración y mediación social, hay una co-autoría en el texto. La *intertextualidad* (término acuñado por Kristeva en 1980) no es solo una referencia a otros textos, también indica cómo se posiciona el escritor para hacer su propia declaración (Bazerman, 2004; Prior, 2006).¹

Si se define un texto como un evento comunicativo en un contexto, un discurso sería un conjunto de textos interconectados, donde la instancia principal es la conversación (Beaugrande, 1997, p. 21). Calsamiglia & Tusón (1999, p. 15) consideran “el discurso una práctica social, articulada a partir del uso lingüístico contextualizado, oral o escrito”. *Discurso* también se puede utilizar para referirse al uso de la lengua en instituciones o profesiones (Bhatia, 2004). Para Zorraquino y Portolés (1999, p. 4057) el *discurso* “es la acción y el resultado de utilizar las diferentes unidades, que facilitan la gramática de una lengua en un acto concreto de comunicación; por ello, todo discurso se compone de una parte puramente gramatical y de otra pragmática.”

Pero ¿qué características tiene un texto académico? A continuación nos detendremos solamente en algunos aspectos.²

Desde el punto de vista semántico, el texto se considera el resultado de un progreso de la secuencia de información. Por lo tanto, la coherencia, entendida como conexión o compatibilidad de los elementos textuales, es la que dará la continuidad de sentido tanto en un nivel de microestructura (organización oracional) como en un nivel de macro-estructura (organización global del texto) (Kintsch & Van Dijk, 1978).

¹Véase Bazerman (2004) para una detallada revisión de la intertextualidad.

² Para más detalle véase Castelló et al., 2007, y Tolchinsky et al., 2013.

Esto se consigue por medio de conectores gramaticales, lexicalización (isotopía: campos semánticos que dan homogeneidad de significado al texto), progresión temática e integración proposicional (predicado y argumentos) (Heinemann & Heinemann, 2002). Por lo tanto, y tal como recogen Teberosky (2007) y Llamas Saíz (2014) una característica de los textos académicos es su alto porcentaje de conectores o *marcadores del discurso* que tienen una función cohesiva del texto, ya que establecen una relación lógico-semántica entre dos o más oraciones o entre segmentos textuales y que guían al lector a una interpretación determinada (Montolío, 2001; Zorraquino & Portolés, 1999).

Asimismo, como ya hemos señalado en el apartado de la introducción, los textos académicos han de condensar en pocas palabras mucha información, esto se consigue por medio del uso de términos técnicos, las nominalizaciones y la ampliación del grupo nominal. Los términos técnicos permiten la precisión léxica, ya que sus significados remiten de forma unívoca y precisa a la realidad a la que hacen referencia y tienen un significado preciso y consensuado en cada disciplina (Teberosky, 2007; Llamas Saíz, 2014). Las nominalizaciones se realizan por medio de la metáfora gramatical. Este proceso fue descrito por Halliday (1993) y consiste en cambiar la categoría gramatical congruente en otra categoría menos congruente. Por ejemplo, las cualidades, procesos y acciones, que habitualmente se expresan mediante adjetivos y verbos, se transforman en sustantivos. Otro recurso gramatical es la ampliación del grupo nominal, es decir, un nombre con pre- y post-modificadores (Biber & Gray, 2010). Todos estos recursos hacen que el léxico sea muy denso en los textos académicos, que haya una alta densidad léxica, que se mide habitualmente como la proporción de palabras, léxicas o de contenido partido por el total de palabras (Halliday, 1985; Mc Namara, Crossley & Mc Carthy 2010, Snow & Uccelli, 2009; Uccelli, Dobbs & Scott, 2013; Teberosky, 2007).

El progreso de la información en un texto se consigue por medio de lo que Halliday (1994) describió como progresión temática. Para tal fin, emplea el concepto de *tema* (información ya conocida o tópico), que está situado más a la izquierda de la frase, y *rema* (información nueva) es lo que sigue y lo que el emisor expone sobre el tema.

Lo que se selecciona para que vaya en primer término en una oración influirá en la interpretación que haga el lector. Por medio de esta estructura la información textual puede ir avanzando, es lo que se denomina progresión temática.

Por otra parte, y como ya hemos mencionado, la intertextualidad es un aspecto característico de los textos académicos, ya que siempre aparecen referencias a otros textos. Así es como avanza el conocimiento (Bazerman, 1988). En la disciplina científica hay que conocer qué problemas se han tratado hasta el momento, hacia dónde se dirige la disciplina, qué autores son los más influyentes, etc. Pero la intertextualidad también indica cómo se posiciona el escritor para hacer su propia declaración (Bazerman, 1988). Esto se consigue por medio de las citas que pueden ser integradas o no integradas. Las integradas se caracterizan porque el nombre del autor tiene un papel sintáctico en el enunciado, mientras que en la cita no integrada el autor y año están enmarcados entre paréntesis (Teberosky, 2007).

Aunque se suele caracterizar el posicionamiento discursivo en la escritura académica como distanciado y objetivo (Berman & Ravid, 2009; Chafe & Danielewicz, 1987; Snow & Uccelli, 2009; Uccelli, Dobbs & Scott, 2013), esa objetividad no significa una ausencia total de implicación personal hacia el contenido (Llamas Saíz, 2014). Todo texto presenta un posicionamiento del escritor. Al igual que la mente se forma dialógicamente, los textos también adoptan una postura de polifonía o dialogismo (Bakhtin, 1981). Hay un amplio consenso de que el discurso académico especializado es dialógico e interactivo (Bazerman, 2004; Flowerdew, 2014; Swales, 2004). Cada texto refleja una interacción entre escritor y lector (Hyland, 2005). El escritor proyecta su voz en el texto y esta identidad autorial se refleja en la función metadiscursiva interaccional, con la postura y la afiliación (Hyland, 2008a; Hyland & Sancho-Guinda, 2012; Tardy, 2012).

Hasta aquí hemos visto algunas características que presentan los textos académicos. Hemos iniciado esta revisión diciendo que el texto es el resultado de un progreso de la secuencia de información. Para cerrar este sub-apartado, nos centraremos en cómo se organiza la información.

Para ello nos referimos a la noción de género, que será tratada extensamente en otro apartado (4.3). No obstante, diremos que la definición de género que da Bruce (2008) contempla la dimensión cognitiva, es decir, la que se refiere a la orientación cognitiva general de un texto en términos de un propósito retórico particular, algo que se refleja en la manera en que la información está organizada internamente (Bruce, 2008). El propósito retórico de los textos académicos suele ser el de generar y difundir el conocimiento disciplinar (Tolchinsky et al., 2013). Como señalan Sampson & Clark (2008) la investigación científica es un proceso de construcción de conocimiento en el que el desarrollo de las explicaciones es esencial para dar sentido a los datos, para su debate, revisión y crítica dentro de la comunidad científica (Driver, Newton, & Osborne, 2000; Duschl, 2000). Una vez que el proceso de investigación ha terminado, los resultados del proceso deben ser comunicados a la comunidad científica para su debate. Esta comunicación se realiza por medio de la argumentación.

El modo discursivo de la ciencia es básicamente argumentativo, en el sentido que su estructura está formada para apoyar las conclusiones sobre la base de razones y evidencias (Norris & Phillips, 2009; Sampson & Clark, 2008; Sandoval & Reiser, 2004). Los argumentos, como otros modos discursivos, es decir, la narración y exposición, tienen una estructura específica, que cumplen una función particular, en este caso, convencer a la comunidad científica. El conocimiento teórico de la disciplina y el uso de métodos apropiados hacen que el discurso sea coherente, porque responden a las expectativas del lector y de la comunidad disciplinaria (Sandoval & Milwood, 2005).

Un argumento tiene un significado tanto individual como social y existe una relación entre ellos (Kuhn, 1993). El significado individual se refiere a cualquier pieza de discurso razonado, es decir, el desarrollo de un punto de vista (Billig, 1987). Este pensamiento reflexivo desarrolla el pensamiento crítico, porque es una manera de buscar evidencias para justificar creencias (Siegel, 1992). También tiene un componente para la emancipación, ya que permite a los estudiantes entenderse a sí mismos y al mundo, lo que les proporciona la capacidad de transformar la sociedad (Freire, 1970). El significado social se refiere al debate de puntos de vista diferentes entre las personas. La argumentación que se produce en un diálogo social mejora el pensamiento de orden superior, ya que se externaliza el razonamiento interno (Jiménez-Aleixandre & Erduran, 2008).

Si tenemos en cuenta el modelo de argumentación propuesto por Toulmin (1958), la argumentación pone su énfasis en las aserciones (*claims*), que se apoyan en los datos específicos (*data*) a través de garantías (*warrants*). Por lo tanto, la argumentación permite a los estudiantes desarrollar criterios racionales epistémicos, porque tienen que escoger entre teorías o posiciones y esto forma la base de la enculturación en la comunidad científica (Sampson & Clark, 2008). Las prácticas epistémicas son las formas en que los miembros de una comunidad proponen, justifican, evalúan y legitiman las afirmaciones de conocimiento dentro de un marco de disciplina (Kelly, 2008). Por todo ello, hay un amplio reconocimiento de la importancia de argumentar como parte del proceso educativo formal, ya que como hemos visto, desarrolla el pensamiento crítico, y forma comunidades de conocimiento (Bazerman, 1988; Erduran, Simon, & Osborne, 2004; Kelly & Crawford, 1997; Sandoval & Reiser, 2004).

En este apartado hemos detallado algunas características de los textos y cómo se organiza la información en ellos. Entendemos el *texto* como una entidad multidimensional que aglutina la dimensión funcional, situacional, temática y la gramática de la lengua (Heinemann, 2000). Esta definición nos lleva a considerar el texto como un producto poliédrico, susceptible, por tanto, de ser analizado desde diferentes perspectivas como, por ejemplo, la cognitiva. En el siguiente apartado explicaremos la dimensión cognitiva de la alfabetización.

4.2.2 La dimensión cognitiva

La dimensión cognitiva de la alfabetización se centra en los procesos mentales, estrategias y procedimientos para crear significado. A continuación describiremos brevemente los procesos implicados en la lectura y la escritura, ya que no podemos considerar la alfabetización sin la lectura. En cambio, el tema de la escritura, al constituir nuestro objeto de estudio, se aborda con más detalle a lo largo de los diferentes apartados.

Leer “es un proceso de interacción entre un lector y un texto, proceso mediante el cual el lector intenta satisfacer los objetivos que guían su lectura” (Solé, 1992, p.21). De esta definición se desprende, por una lado, que hay diferentes niveles de lectura, según los objetivos que tiene el lector; y, por otro, que el aprendizaje de la lectura es un proceso, al igual que la escritura, es decir, que no se completan en un momento determinado (Carlino, 2003; Solé, 2014). Asimismo, esta definición implica una perspectiva constructiva, no meramente reproductiva, que reconoce las diversas interpretaciones que pueden suscitar la lectura de un texto. Desde este enfoque constructivista, no obstante, se advierte que en el entorno académico, no todas las interpretaciones son adecuadas. Por lo tanto, pone el énfasis en un lector crítico con la argumentación y que busca en el texto la interpretación más exacta posible (Solé, 2014).

La lectura es un proceso en que interviene la percepción y la memoria. Las dos fuentes de información de la lectura son la información visual y la no visual que interactúan y en la que interviene la memoria (Smith, 1990). Este proceso es activo y constructivo, en el que el cerebro influye en la selección que hacen los ojos ante los estímulos visuales. La información se capta a través de los sentidos y se almacena en los esquemas de conocimiento, este proceso es común a todos los procesos de obtención de información (Colomer & Camps, 1996). Por lo tanto, cuanto más información tenga el cerebro menos dependerá de la selección visual y recíprocamente. Así, el objetivo que guía la lectura ayudará a la comprensión del texto.

Gracias a los avances tecnológicos como el *eye-tracker* se ha comprobado que los ojos se desplazan con pequeños saltos, movimientos saccádicos que avanzan y retroceden y focalizan en determinadas palabras, son las fijaciones. Sobre todo en palabras léxicas con contenido semántico el tiempo de la fijación es mayor. La información abarcada en cada fijación se envía al cerebro para su procesamiento, los lectores inexpertos abarcan pocas palabras en cada fijación, mientras que los expertos tienen un campo más amplio y se detienen con menos frecuencia. En todo este proceso también interviene la memoria a corto y largo plazo (Colomer& Camps, 1996; Kucer, 2001).

Un texto con sentido, organizado en conjuntos significativos, ayudará en el proceso de comprensión lectora. Los tipos de texto, que tienen una estructura convencional, permiten al lector tener un esquema que le facilitará la comprensión (Kintsch & Van Dijk, 1978). Así los trabajos de investigación también se caracterizan, en general, por tener una estructura fija: introducción, método, resultados y discusión (IMRD) (Swales, 1990).

Por otro lado, escribir es una forma de crear significado por medio de símbolos (Kellogg, 1994), o más detalladamente, es “el proceso mediante el cual se produce un texto escrito significativo” (Correig, 2008, p. 146). Si nos basamos en los modelos cognitivos tradicionales (e.g., Flower & Hayes, 1980) el proceso de escritura se considera un problema a resolver, esto es, la actividad de realizar determinados pasos secuenciados para conseguir un objetivo (Hayes, 1996). En este proceso hay un problema de contenido (qué decir) y un problema retórico (cómo decirlo), que requiere pensar estratégicamente la mejor manera de conseguir el objetivo comunicativo propuesto (Galbraith, 2001; Kellogg, 2008). Desde este conocido modelo teórico se considera que la escritura consta de tres elementos de nivel superior, la situación de comunicación, la memoria a largo plazo del escritor y el proceso de escritura, que a su vez, tiene unos procesos básicos: planificar, traducir las ideas a texto y revisar, que son recursivos, no son secuencias lineales. La situación de comunicación contiene elementos externos, el problema retórico, tema, audiencia y objetivos; y el texto que se va escribiendo.

En la memoria a largo plazo el escritor ha almacenado conocimientos, no solo del tema, sino también de la audiencia y las diferentes tipologías y estructuras textuales. El proceso de escritura, a su vez, tiene subprocesos, así en la planificación se generan ideas, se organizan estas y se formulan objetivos. En el proceso de traducir las ideas o redacción intervienen conocimientos de bajo nivel (e.g., composición de palabras), que si bien en adultos se ha convertido en un proceso automático, en estadios iniciales de la escritura, requiere habilidades motrices y gran demanda memorística. Asimismo, intervienen conocimientos de alto nivel (e.g., sintácticos, cohesión, etc.). En el proceso de revisión intervienen los subprocesos de evaluar y rehacer. Todo este proceso de escritura está controlado por un monitor, que es el responsable de coordinar este proceso, por ejemplo, el escritor decide cuándo pasar al siguiente proceso (Flower & Hayes, 1980).

Hay que tener en cuenta que, según algunos estudios (Breetveld, van den Bergh & Rijlaarsdam, 1994), la mayoría de las actividades cognitivas (planificación, traducción, revisión, etc.) muestran un patrón distinto de ocurrencia durante el proceso de escritura. Además, las diferentes combinaciones de estas actividades afectan de manera diferente y en momentos diferentes el proceso de escritura y así también la calidad del texto (Rijlaarsdam & van den Bergh, 2006). Es decir, la misma actividad cognitiva que muestra un efecto positivo en la calidad de texto durante una fase determinada, muestra un efecto negativo durante otra fase o ningún efecto en absoluto (Breetveld, van den Bergh & Rijlaarsdam, 1994). Esto conlleva una revisión de la unidad de análisis para establecer teorías sobre el proceso de escritura, ya que las actividades cognitivas tienen una relación funcional (Rijlaarsdam & van den Bergh, 2006). Todo ello revela, una vez más, la complejidad del proceso de escritura, que no solamente impone importantes retos a la cognición, tanto a nivel de memoria de trabajo como a nivel de razonamiento (Kellog, 2008), sino también a su análisis y estudio: "In writing process studies, we only have access to the tip of an iceberg. The more we know, the more complex the writing process seems to be" (Rijlaarsdam, et al. 2005, p. 135).

4.2.3 La dimensión sociocultural

Existen diferentes teorías socioculturales sobre la alfabetización y la mayoría se originan a partir del trabajo de Vigotsky (1978), por eso, todas muestran un especial interés en el contexto social y cultural donde se lleva a cabo la alfabetización. Es decir, entienden la alfabetización como una práctica social, ya que el uso del lenguaje ilustra o ejemplifica la cultura (Perry, 2012; Street, 1984). La teoría sociocultural no considera las acciones de las personas como reguladas por reglas o normas, sino que entiende cada actividad como situada en interacciones concretas, y mediada por prácticas y herramientas, tales como por ejemplo, los modos semióticos (Prior, 2006). Así desde la lingüística funcional, Halliday (1978) nos recuerda que el lenguaje solo surge en un contexto de situación, en un contexto social, asimismo, la cultura es un sistema de significados, "and the linguistic system is one mode of realization of these meanings" (Halliday, 2004b, p. 301).

Tal como recoge Perry (2012), las diferentes perspectivas socioculturales en la alfabetización se centran, por ejemplo, en cómo el lenguaje está relacionado con las relaciones sociales, valores y actitudes, o con el contexto cultural (e.g., Gee, 1996; Halliday, 1973), las formas en que el uso del lenguaje varía de acuerdo a los contextos (e.g., Bakhtin, 1986), la relación entre el uso de la lengua y el poder (e.g., Bourdieu, 1991), y la etnografía de la comunicación (e.g., Hymes, 1994). Nosotros nos detendremos en dos perspectivas: las que consideran la alfabetización como práctica social y las que la consideran como multi-alfabetizaciones (Perry, 2012).

La alfabetización como práctica social se vio influenciada inicialmente sobre todo por la obra de Street (1984), y tiene su representación en los *estudios de nueva literacidad*, *New Literacy Studies* (NLS). La alfabetización desde esta perspectiva está respaldada por la pragmática (Prior, 2006), ya que tal como resume Perry (2012, p. 54): "literacy is what people do with reading, writing, and texts in real world contexts". Los autores que se adscriben a este movimiento se oponen al paradigma psicológico que considera la alfabetización únicamente como un logro cognitivo e individual, sino que la consideran un fenómeno sociocultural.

Por lo tanto, desde esta perspectiva, la alfabetización no sería solamente un proceso mental de descodificación, inferencia, revisión etc., más bien, es considerada como el aprendizaje de la lectura o la escritura mediante la participación en las prácticas sociales y culturales distintivas de los diferentes grupos sociales y culturales. Esto implica que el lenguaje escrito se utiliza de forma diferente en las prácticas de los diferentes grupos sociales, esto es, de acuerdo con ciertos propósitos y valores epistémicos (Bazerman, 2004; Carlino, 2005; Ivanic, 1998; Prior, 2006). De ahí que, ya no se pueda hablar de una alfabetización, sino de múltiples. Consecuentemente, el estudio de la alfabetización necesita tener en cuenta los diferentes contextos, no solo factores cognitivos, sino también el contexto social, cultural, histórico e institucional (Gee, 2010). Aunque la teoría de la alfabetización como práctica social no puede explicar el proceso de cómo las personas aprenden a leer y escribir, puede ayudar a describir qué tipo de conocimientos son necesarios para poder participar de manera efectiva en las prácticas de alfabetización (Perry, 2012).

Por otro lado, la teoría de la multi-alfabetización se desarrolló por el *New London Group* (Cope & Kalantzis, 2000). Este movimiento, como los *estudios de nueva literacidad*, también enfatiza los contextos sociales de alfabetización. En lo que difiere con el anterior, es que no solamente se centra en el lenguaje impreso como modo de representación, sino que va más allá y considera otros modos semióticos para crear significado, por ejemplo, el visual, auditivo, gestual, etc. Es lo que se conoce como multimodalidad (Perry, 2012; Kress, 2000).

Para resumir hemos comentado que desde una perspectiva sociocultural la lectura y la escritura no son solamente actos individuales de pensamiento y lenguaje, sino que también configuran los patrones de los actos sociales de los grupos. La alfabetización se da porque la pertinencia al grupo lo requiere. Es por este motivo que las diferentes maneras de usar la cultura escrita por parte de los miembros de un grupo, refleja la naturaleza del grupo o la posición del grupo dentro de la sociedad (Kucer, 2001).

4.2.4 La dimensión de desarrollo

La dimensión de desarrollo considera cómo los individuos entienden y utilizan los principios y patrones del lenguaje durante su desarrollo. Como en los demás apartados, solo nos centraremos en el desarrollo de la alfabetización más allá de la primera infancia, concretamente en el período adolescente, ya que una parte de nuestro corpus está escrito por alumnos de educación secundaria. Esta etapa se considera como un momento crucial en el conocimiento lingüístico y uso del lenguaje (Berman, 2011). Por ejemplo, la lectura se basa en la conversión rápida y precisa de caracteres ortográficos, grafemas, en sus representaciones fonológicas, fonemas. Con el desarrollo de la alfabetización, los lectores tienen un acceso más rápido y automático de estos.

Desde la neurociencia se ha comprobado que hay diferencias de maduración cerebral en la sinaptogénesis, el metabolismo de la glucosa, la mielinización y en la materia blanca; todas estas aumentan, especialmente en las regiones prefrontales, durante la pubertad (10-20 años). Este hecho está correlacionado con la automatización de la lectura (Schlaggar et al., 2002 *apud* Proverbio & Zani, 2005). Igualmente, el incremento en la densidad y complejidad de las estructuras cerebrales se relaciona, en general, con una mayor explicitación y accesibilidad al conocimiento (Karmiloff-Smith, 1992).

Desde una perspectiva psicolingüística se ha constatado, que a partir de la adolescencia se da un mayor y más elaborado control lingüístico, tanto en la modalidad oral como en la escrita (Berman & Ravid, 2009). Así, se constata una mayor complejidad sintáctica y extensión del grupo nominal. Además se da un mayor control en la organización del texto expositivo y en el procesamiento y comprensión de diferentes géneros del discurso (Berman, 2011). En un proyecto internacional translingüístico (Berman & Verhoeven, 2002) se observó un importante cambio lingüístico en el grupo de los adolescentes respecto a las etapas evolutivas anteriores. Este proyecto examinó en holandés, inglés, francés, hebreo, islandés, español y sueco el desarrollo de la producción de textos en la modalidad oral y escrita en los géneros narrativo y expositivo desde los 9 años hasta la edad adulta. Los dominios analizados iban desde un nivel local lingüístico, léxico y sintaxis, hasta un nivel global de calidad textual, contenido y posicionamiento (Berman, 2007).

En cuanto al tema de posicionamiento o postura (*stance*), los resultados mostraron que antes de la adolescencia, en los textos expositivos, se expresan generalidades sin ningún tipo de evaluación personal hacia los hechos expuestos. En cambio, a partir de la adolescencia, se combinan una perspectiva genérica y abstracta con comentarios personales de evaluación de los hechos. Esta evaluación se expresa por medio del uso de verbos modales con un sentido epistémico (e.g., podría, debería, etc.) que cumplen también una función de mitigadores o atenuadores del grado de compromiso hacia la proposición, ya que están relacionados con estados hipotéticos o múltiples perspectivas (Berman, 2007).

Así, se constató que en el progreso de la alfabetización, en general, los adolescentes tienen desarrolladas las capacidades metalingüísticas, aunque las capacidades metatextuales requieren de una mayor madurez cognitiva. En cuanto a la lectura, esto significa tener la capacidad de ir más allá del texto, aplicar en él los conocimientos previos del mundo y sistema de valores para conseguir interpretar el sentido de manera bien informada e individual. En relación a la escritura, no es hasta la etapa adulta en que se añaden comentarios personales a las ideas o eventos que se describen, generalizando desde la propia experiencia vital. Por otro lado, un mayor desarrollo de la alfabetización influye en la oralidad, así el lenguaje propio de la modalidad escrita se llega a utilizar en la modalidad oral (Jisa, 2004).

4.2.5 La dimensión multimodal

La semiótica social tiene como objetivo explicar cómo creamos significado con todos los recursos a nuestra disposición: los lingüísticos, pictóricos, gestuales, musicales, etc. (Kress & Van Leeuwen, 1996; Lemke, 1998a). Si consideramos que toda comunicación humana es multimodal, ya que se interrelacionan diferentes modos, como por ejemplo, el textual, visual, auditivo para crear significado (Lemke, 1998b), también hemos de entender que la alfabetización, en nuestro caso concreto, el aprendizaje de las disciplinas especializadas, implica tener en cuenta otros modos semióticos. El discurso verbal en el aula no es del todo comprensible sin activar otros sistemas semióticos, como el visual o el gestual.

De hecho, no es posible activar solo un sistema semiótico concreto (Kress, Jewitt, Ogborn & Tsatsarelis, 2001; Márquez, Izquierdo & Espinet, 2003; Lemke, 1998b). Movilizamos siempre más de un sistema, por consiguiente, siempre hay más de un posible significado en lo que expresamos. En el caso de las ciencias esta multimodalidad se hace más evidente. Es cierto que en las clases de ciencias la mayor parte de la información conceptual tiene una forma lingüística, oral o escrita. El uso del lenguaje verbal permite adquirir cierta familiaridad con los conceptos técnicos científicos, pero el aprendizaje también incluye la adquisición de la forma de representación del razonamiento científico.

Los científicos no producen solamente un texto verbal lineal cuando se comunican. Usan representaciones gráficas (figuras, mapas, tablas, gráficos y fotografías) y expresiones matemáticas, porque piensan en espacios imaginarios visuales. Si miramos un artículo científico podemos constatar este hecho, no veremos solo un texto verbal lineal, sino que la información y la argumentación científica está construida a través de expresiones matemáticas, de gráficas, de diagramas, de imágenes, etc. El uso de esos recursos multimodales les permite multiplicar y modular el significado (Lemke, 1998b). En otras palabras, un concepto científico no incluye todas las representaciones posibles, de hecho ningún tipo de concepto lo incluye, pero en el caso del concepto científico, este se articula a través de diferentes modos de representación para conseguir un significado completo. Además, no hay una traducibilidad directa e inequívoca entre los diferentes modos, por lo que la combinación de los mismos modula el significado.

El ser capaz de utilizar un concepto científico, y por lo tanto para entenderlo en la forma en que un científico lo hace, requiere saber coordinar e interpretar con soltura los aspectos verbales, visuales y matemáticos. Y estas prácticas de coordinación deben ser aprendidas en cada caso como una forma especializada de alfabetización multimodal (Lemke, 1998b). Así, esta alfabetización multimodal puede ser analizada en el texto, por ejemplo, desde la lingüística sistémica-funcional, en la cual las tres macro-funciones que se proyectan en el texto (Halliday, 1980) - la *ideacional*, *interpersonal* y *textual*- se pueden hacer corresponder con la función *presentacional*, *orientacional* y *organizacional* en los recursos no verbales (Lemke, 1998b; Pujol Dahme, 2014).

También puede ser estudiada como actividad de aprendizaje de los estudiantes, esto es, el proceso de estudio y comprensión de representaciones gráficas, como por ejemplo diagramas, por medio de protocolos *think-aloud* y *eye-tracking* (Kragten, Admiraal & Rijlaarsdam, 2015).

Para finalizar este apartado queremos resaltar que, si tenemos en cuenta que una de las funciones del lenguaje es la *metalingüística*, la capacidad de reflexionar sobre el lenguaje, consideramos fundamental el desarrollar también la función *metamultimodal*, esto es, la capacidad de reflexionar sobre otros modos semióticos, sobre todo cuando se trabaja con híbridos semióticos como son los trabajos de investigación (Pujol Dahme, 2014, p. 47).

En este capítulo hemos visto las diferentes dimensiones que componen la alfabetización: la lingüística, la cognitiva, la sociocultural, la de desarrollo y la multimodal. Todas ellas nos dan una idea de la complejidad que supone, por un lado, adquirir la alfabetización y, por el otro, los paradigmas que configuran la alfabetización y que evolucionan constantemente para poder avanzar en el conocimiento de esta y así facilitar su adquisición.

Tal como señala Johns (2002), hubo un importante cambio de paradigma en los estudios teóricos y pedagógicos sobre la alfabetización. En los años 70-80 predominaba una perspectiva psicolingüística y cognitiva que tenía su foco en el proceso, *process approach*, que se centra en el proceso que un escritor hace en la construcción de un texto y este puede incluir etapas identificadas del proceso de escritura como las siguientes: pre-escritura, la escritura y reescritura (Montague, 1995). Fue partir de los años 90 en que la perspectiva se focalizó hacia el contexto, a las situaciones que dan lugar a la escritura, cobraba importancia el escritor como ser social. Por lo tanto, la idea de *género*, entendida como forma retórica tipificada de interactuar en situaciones recurrentes (Miller, 1984), influyó notablemente en los estudios de la alfabetización. Esta autora suscitó una nueva conceptualización del término en el que el contexto, el contenido, los roles de escritor-lector y los valores de cada comunidad discursiva cobraban interés (Johns, 2002).

Obviamente, la noción de *género* no surgió en esa época, sino que el hecho de considerar los textos como géneros data de mucho antes. Desde la definición de *género* como herramienta para clasificar textos, a una forma de conocimiento cultural y hasta su utilización como formas de resistencia y de cambio, las definiciones se han sucedido a lo largo del tiempo.

Bawarshi & Reiff (2010) lo explican así:

At various times and in various areas of study, genre has been defined and used mainly as a classificatory tool, a way of sorting and organizing kinds of texts and other cultural objects. But more recently and, again, across various areas of study, genre has come to be defined less as a means of organizing kinds of texts and more as a powerful, ideologically active, and historically changing shaper of texts, meanings, and social actions. From this perspective, genres are understood as forms of cultural knowledge that conceptually frame and mediate how we understand and typically act within various situations. This view recognizes genres as both organizing and generating kinds of texts and social actions, in complex, dynamic relation to one another (p.4)

Por su importancia, en el siguiente apartado nos detendremos en el concepto de *género*.

4.3 Género

El *género* es un concepto importante en los estudios sobre escritura y educación, ya que el *género* está ligado a todas las culturas escritas (Chitez & Kruse, 2012; Hyland, 2004b). En los sistemas complejos de actividad, incluidas las de educación formal, normalmente hay muchos géneros escritos, que los participantes combinan para organizar sus interacciones (Russell, 1995). Además, es a través de los géneros que se establece el conocimiento disciplinario (Chitez & Kruse, 2012). La alfabetización avanzada implica avanzar hacia el dominio de nuevos géneros (Uccelli, Dobbs, & Scott, 2013).

Los estudiantes deben llegar a dominar nuevas formas de comunicación, deben adquirir *competencia comunicativa* (Hymes, 1972) cuando pasan de un contexto discursivo a otro, de la escuela secundaria a la universidad y de la universidad al mundo profesional (Berkenkotter, Huckin & Ackerman, 1991). Esta competencia comunicativa está ligada al conocimiento de diferentes géneros, dado que los miembros de cada comunidad discursiva utilizan unos géneros y un léxico específico para lograr sus fines (Swales, 1990).

No obstante, no hay un consenso generalizado sobre el término *género* (Bawarshi & Reiff, 2010, p. 3; Biber & Conrad, 2009, p. 21-23) y las aportaciones al concepto son constantes desde diferentes marcos teóricos, ya que la noción es intrínsecamente compleja (Ciapuscio, 2005). Incluso hay quien utiliza *género* y *registro* indistintamente para referirse a “variedades asociadas a situaciones particulares de uso y propósitos comunicativos particulares” (Biber & Conrad, 2009, p. 21).

En un sentido amplio *género* refiere a distinciones típicas entre los textos y a las formas de utilizar el lenguaje en situaciones recurrentes (Chitez & Kruse, 2012). Más concretamente, cada género tiene un propósito específico, una estructura general, rasgos lingüísticos específicos y es compartido por los miembros de una comunidad (Hyland, 2004b). Por ello, son formas fácilmente reconocidas por los usuarios de la lengua, ya que los géneros se distinguen en el caso de los textos por las similitudes textuales, o bien por el contexto social en que se dan (Hyland, 2004b).

Aunque desde la perspectiva de la lingüística de texto, por ejemplo, se relacione la categoría de *tipo de texto* con la noción de *género* (Beaugrande & Dressler, 1981), hay quien propone hacer una distinción entre *tipo de texto* y *género* basándose en criterios internos y externos (Biber, 1988). Por lo tanto, podríamos definir *tipo de texto* basándonos en criterios internos, esto es, en las características lingüísticas y definir *género* basándonos en criterios externos, en el uso, es decir, en la actividad misma, el propósito comunicativo y la audiencia a la que va dirigida la comunicación³.

De esta definición se desprende la importancia del contexto social en la construcción del género. El género discursivo se realiza en las interacciones sociales. El hablante selecciona un género, forma típica relativamente estable, según la esfera discursiva en la que se encuentra (Bakhtin, 1986). Así, en un sentido estricto, un género es una estrategia de repetición socialmente reconocida para lograr objetivos similares en situaciones socialmente percibidas como similares (Miller, 1984, *apud* Bazerman, 1988).

Pero, tal como señala Bazerman (2004), si se consideran los géneros solo como un conjunto fijo de características textuales, serían atemporales e iguales para todos. No se tendría en cuenta el papel de las personas en el uso y creación de significados. El conocimiento general cambia con el tiempo, al igual que cambian las situaciones, las personas y los géneros. Los géneros, pues, surgen en los procesos sociales de las personas por sus efectos prácticos, para coordinar actividades y compartir significados.

Un número considerable de autores han descrito el concepto de *género*. Prior (2009), por ejemplo, afirma que los géneros han sido descritos en términos de cadenas (Swales, 2004; Fairclough, 2004), colonias (Bhatia, 2002a), repertorios (Orlikowski & Yates, 1994), o conjuntos y sistemas (Bazerman, 1994).

³Para una discusión detallada de estos términos véase Lee (2001).

Nosotros solo revisaremos sumariamente algunas definiciones de *género*. Estas se suelen clasificar en tres tradiciones, inicialmente descritas por Hyon (1996) y posteriormente utilizadas por otros autores como, por ejemplo, Chitez & Kruse (2012), Bawarshi & Reiff (2010), Hyland (2004b) y Johns (2002). Las tres principales tradiciones son: la *nueva retórica*, la *lingüística sistémico-funcional* (LSF) o también denominada la escuela de Sydney y la de *inglés para fines específicos* (*English for Specific Purposes*). Aunque la mayoría de autores utiliza esta clasificación basada en las tradiciones y por tanto, en su orientación teórica, Flowerdew (2002) sugiere que estas también se pueden clasificar, en un sentido amplio, en perspectivas lingüísticas o no lingüísticas, es decir, si los investigadores priorizan el texto o el contexto en sus estudios (Woodward-Kron, 2005). Según Flowerdew (2002), la *lingüística sistémico-funcional* y el *inglés para fines específicos* tienen un enfoque principalmente lingüístico. Se centran en la realización retórica y léxico-gramatical del propósito comunicativo que se da en los géneros. En cambio, un enfoque no-lingüístico, como es el caso de la nueva retórica, se centra más en la situación del contexto, en las actitudes y comportamientos de una comunidad de discurso.

Para nuestro propósito consideramos útil esta clasificación, ya que nuestro trabajo tiene una perspectiva lingüística, es decir, priorizamos el texto más que el contexto, nos centramos en los aspectos léxico-gramaticales y retóricos del género de investigación. No obstante, reconocemos la característica dinámica de *género* postulada por la nueva retórica, además coincidimos con Miller (1984, p. 165) cuando afirma que el género sirve para entender cómo se participa en las acciones de una comunidad.

A continuación veremos con un poco más de detalle las tres tradiciones previamente mencionadas, ya que, como Gardner (2012); Gardner & Nesi (2013); Schleppegrell (2000) entre otros, afirmamos que la tradición de la *lingüística sistémico-funcional* ofrece herramientas útiles para el análisis de los textos. Tiene el potencial de vincular las opciones léxicas y gramaticales con los contextos sociales. En las líneas que siguen describimos las características de las tres tradiciones: la *nueva retórica*, la *lingüística sistémico-funcional* y el *inglés para fines específicos*.

1. La *nueva retórica* tiene una perspectiva social e ideológica y está basada en estudios literarios y retóricos. En este sentido, el crítico literario Bakthin (1986) ha tenido una gran influencia en la tradición de la nueva retórica, no solo por enfatizar el *dialogismo* en el discurso, sino también porque su unidad de análisis es el enunciado y este ya se define como social, es decir, siempre se da en relación con otros hablantes (Freedman & Medway, 1994). El análisis más minucioso de género como un fenómeno social, sin embargo, proviene de los estudios retóricos y no de los literarios (Bazerman, 1988).

Miller (1984), cuyo trabajo ha tenido una gran repercusión en lo que al estudio del concepto de *género* se refiere, propone que una definición de *género* debe centrarse no en la forma del discurso sino en la acción o propósitos sociales que se utilizan para llevarlo a cabo. Así define *género* como “typified rhetorical actions based in recurrent situations” (Miller, 1984, p.159).

Por un lado, la tipificación de acciones retóricas implica la aparición de tipos de texto reconocibles marcados por elementos formales repetidos (Bazerman, 1988). Por otro, las acciones retóricas deben estar conectadas a la cognición, ya que lo que sabemos y cómo actuamos están relacionados entre sí (Bawarshi & Reiff, 2010). Berkenkotter & Huckin (1995) sugieren que el conocimiento del *género* es una forma de cognición situada en actividades disciplinarias: “Genre knowledge is a form of situated cognition, inextricable from professional writers' procedural and social knowledge” (p.13). El conocimiento procedural refiere cómo reconocer y negociar los problemas, y dónde, cómo y cuándo producir conocimiento dentro de contextos disciplinarios. El conocimiento social es el que tiene el escritor sobre el contexto conceptual y retórico (Bawarshi & Reiff, 2010). El conocimiento de los géneros se deriva de, y está arraigado en la participación en las actividades comunicativas de la vida cotidiana y profesional. Por lo tanto, el género se entiende como un acto social, que contiene las formas de conocer y de actuar de una comunidad discursiva, que cambian con el tiempo (Berkenkotter & Huckin, 1993; Miller, 1984). Desde este punto de vista, estos autores no efectúan un análisis lingüístico tan preciso como en las tradiciones *sistémico-funcionales* y del *inglés para fines específicos*, ya que privilegian una metodología etnográfica (Bawarshi & Reiff, 2010; Hyland, 2004b).

2. Desde la *lingüística sistémico-funcional* el concepto de *género* es visto como un fin social con una naturaleza básicamente funcional. Según Martin (1984), el género es un proceso social orientado a un objetivo y organizado en etapas. Es un proceso social porque para generar un género los miembros de una cultura interactúan entre ellos. Está orientado a unos objetivos porque las fases que se desarrollan están diseñadas para lograr un fin. Está organizado en etapas, porque se da habitualmente más de una fase de significado para lograr un fin (Martin, 2009).

Este enfoque pone de relieve la relación entre texto y contexto. Cada texto se enmarca en un contexto y este opera en dos niveles: a nivel de registro y a nivel de género. El registro vincula un tipo de situación con determinados patrones semánticos y léxico-gramaticales, por lo que el registro describe lo que se lleva a cabo (*campo/ field*), que incluye el tópico y el propósito del emisor; cómo los participantes se relacionan entre sí (*tenor*); y cómo se utiliza el lenguaje en la interacción (*modo/mode*), que incluye el canal de la comunicación, oral o escrito, y el modo retórico, por ejemplo, persuasivo, didáctico, etc. El parámetro de campo caracteriza textos en términos de su dominio de especificidad, cómo se describe en términos de léxico, terminología especializada, etc., por consiguiente, el campo varía según las disciplinas. El tenor del discurso se refleja en los textos en términos de la interacción entre los participantes involucrados en la situación de discurso, por ejemplo, de estudiante a experto en los trabajos de investigación de los estudiantes y de experto a experto en los artículos publicados. El parámetro de modo se refleja en los textos en que es un escrito para ser leído, es denso, técnico, abstracto y acompañado con imágenes o gráficos (Holtz, 2011; Martin & Rose, 2007). Estos tres parámetros, campo, tenor y modo, son abstractos y difícilmente operacionalizables, es decir, son variables latentes. No obstante, diferentes autores (e.g., Steiner, 2004; Neumann, 2014) han establecido indicadores lingüísticos concretos de estos constructos.

Todas estas variables influyen en la elección lingüística, que el hablante hace, de todas las posibilidades presentes en el sistema lingüístico. En el nivel de género, el propósito social también determina las elecciones que se hacen del sistema (Christie, 1990, *apud* Johns, 2002). El registro funciona en el nivel de contexto de la situación y el género funciona en el nivel de contexto de la cultura. La situación es el contexto del texto. La cultura es el total potencial de todas las situaciones (Halliday, 2004b). El registro se realiza a través de la léxico-gramática y el género se realiza a través del registro.

Así los géneros son moldeados por su objetivo cultural, por ejemplo, en la escritura académica los estudiantes, en el género del ensayo, tendrían que demostrar la capacidad de construir un argumento o ser capaces de, en el género de trabajo de investigación, realizar una investigación independiente (Gardner, 2012). Desde esta perspectiva no hay una distinción entre género y tipo de texto, como por ejemplo, el narrativo, el descriptivo y el expositivo, por lo que se utiliza habitualmente el término *macrogénero* para referirse a géneros complejos que son una combinación de varios géneros, por ejemplo, unas instrucciones (Hyland, 2004b).

3. Desde la tradición del *inglés para fines específicos* (ESP), el género es visto como una competencia profesional, se considera una herramienta para el aprendizaje y la enseñanza del inglés como L2. El trabajo de Swales (1990) ha sido fundamental en el desarrollo de la teoría de *género* en esta perspectiva, ya que enfatiza tanto la función social y la forma de la lengua hablada y escrita en el ámbito académico y de investigación (Flowerdew, 2005). Para Swales (1990) un género comprende una clase de eventos comunicativos, cuyos miembros comparten un conjunto de propósitos comunicativos, que a su vez, establecen los marcos en términos de sus propiedades formales y estructuras esquemáticas. Es decir, un análisis desde este enfoque se iniciará mediante la identificación de un género dentro de una comunidad de discurso y la definición del propósito comunicativo que pretende lograr ese género. A partir de allí, el análisis se detiene en la organización, en la estructura esquemática que se caracteriza habitualmente por los movimientos retóricos (*moves*).

Un *movimiento* es una unidad discursiva o retórica que realiza una función comunicativa coherente, en un discurso hablado o escrito (Swales, 2004). Posteriormente, se examinarían los rasgos textuales y lingüísticos (estilo, la gramática, la sintaxis) de esos movimientos retóricos. En definitiva, se parte primero del contexto y posteriormente se analiza el texto (Bawarshi & Reiff, 2010).

La noción de *género* desde la *lingüística sistémico-funcional* y desde el *inglés para fines específicos* tiene en común, que las características lingüísticas están conectadas tanto a la función como al contexto social. No obstante, tienen objetivos pedagógicos diferentes. La *lingüística sistémico-funcional* tiene como objetivo de estudio generalmente a los niños en edad escolar de entornos culturales y económicos desfavorecidos en Australia. En cambio el *inglés para fines específicos* se dirige a niveles más avanzados de escolaridad, es decir, a estudiantes internacionales de universidades anglófonas. Por lo tanto son estudiantes cuyas capacidades lingüísticas están en desventaja respecto a los nativos anglófonos.

Las diferencias en el objetivo al que van dirigidos hace que la noción de contexto difiera entre *lingüística sistémico-funcional* y el *inglés para fines específicos*. La perspectiva de la *lingüística sistémico-funcional* se centra generalmente en pre-géneros, por lo que, tal como ya hemos visto, la definición de *contexto de cultura* es muy amplia. En cambio, desde el *inglés para fines específicos*, el *género* define dentro de contextos más específicos, lo que Swales (1990) denomina "comunidades discursivas" (Bawarshi & Reiff, 2010).

4.3.1 Comunidades discursivas versus comunidades de práctica

Podríamos pensar que los términos *comunidad discursiva* y *comunidad de práctica* se utilizan indistintamente cuando hablamos de discurso académico, no obstante, igual que en la definición de género hay cierta controversia entre estos dos conceptos (Bamford & Bondi, 2005). Por ejemplo, el concepto de *comunidad discursiva* ha sido criticado por ser impreciso e inexacto, por tener una connotación de uniformidad y de suponer unas relaciones simétricas entre sus miembros (Bazerman & Prior, 2005).

Asimismo, Prior (2003) se pregunta si las comunidades de práctica son realmente tan diferentes a las comunidades discursivas. La *comunidad de discurso* es un término estrechamente asociado con la interpretación que hace Swales (1990) de género: “A genre comprises a class of communicative events, the members of which share some set of communicative purposes.” (p. 58.).

Para Swales (1990) la comunidad de discurso identifica un grupo de individuos con un conjunto de objetivos públicos comunes, unos mecanismos de intercomunicación establecida, se mantienen informados a través de mecanismos de participación, tienen uno o más géneros para promover los objetivos de la comunidad, utilizan un léxico específico y sus miembros incluyen tanto aprendices como expertos con diferentes grados de experiencia.

A pesar de esta definición, las investigaciones situadas sobre escritura académica han arrojado resultados empíricos que complican o cuestionan la noción de comunidad discursiva disciplinaria (Prior, 1998). Como señala Donahue (2011), la crítica que hace Prior (1998, 2003) se basa en los fundamentos estructuralistas de las diferentes versiones de *comunidad* de las disciplinas, en este sentido las disciplinas se han considerado como *territorios sociales unificados* o *sistemas abstractos de conocimientos codificados*.

No obstante, tal como sugiere Prior (1998) "... participation in disciplinary practices point to doing things rather than having something or being someplace; they suggest a process view of disciplines" (p.xii). Por lo que los estudiantes, cuando aprenden una disciplina, no entran en los espacios sociales y cognitivos de las comunidades discursivas, sino que se involucran en relaciones activas con comunidades de práctica, que son dinámicas y abiertas (Prior, 1998). Por lo cual, algunos autores (e.g., Donahue, 2011; Prior, 1998) sugieren utilizar el término de *comunidad de práctica* (Lave & Wenger, 1991) en vez de *comunidad discursiva*.

La *comunidad de práctica* es definida como "a set of relations among persons, activity, and world, over time and in relation with other tangential and overlapping communities of practice" (p.98), pone el énfasis en el proceso, en la actividad, más que en el producto, y consta de las siguientes dimensiones: un compromiso mutuo, un objetivo conjunto y un repertorio compartido de acciones, discursos, etc. (Wenger, 1998). En cambio la *comunidad discursiva* implica unas convenciones unificadas en cuanto a estructura lingüística, retórica y epistémica (Prior, 1998).

Tal como resume Johns (1997) las *comunidades discursivas* se centran en el texto y el lenguaje, que son los géneros y el léxico que permiten a los miembros lograr sus objetivos y regular la incorporación de miembros. Las *comunidades de práctica* se refieren a los géneros y el léxico, pero especialmente a los valores y prácticas que mantienen a sus miembros unidos. En definitiva, para llegar a ser un miembro de una comunidad de práctica hay que adaptar el propio discurso al de la comunidad a la que se quiere pertenecer, a las convenciones propias de la comunidad específica (Schleppegrell & Colombi, 2002).

Nosotros, en este trabajo utilizamos el término *comunidad disciplinaria* en un sentido más amplio y neutro, *comunidad discursiva* cuando nos referimos sobre todo a las características textuales y *comunidad de práctica* cuando queremos resaltar la participación activa de los estudiantes.

En el apartado siguiente nos detendremos en la escritura, que es el resultado de la participación de las prácticas disciplinarias, un producto discursivo que tiene su concreción en el texto.

4.4 La escritura

La escritura es un logro de la humanidad relativamente reciente, si lo comparamos con otros logros, como por ejemplo, la adquisición del lenguaje; reciente también, en cuanto a su aparición (3.300 a.C. las tablas encontradas en Mesopotamia); y más reciente aun si consideramos su uso generalizado en todas las sociedades (Coulmas, 1989; Grabe & Kaplan, 1996). De lo que no cabe duda es que resulta difícil imaginar algún aspecto de nuestra vida que no esté ligado a la escritura, de hecho el acceso al conocimiento pasa por la escritura, solo hay que pensar en la cantidad de registros escritos conservados en bibliotecas o bancos de datos. Pero a través de la escritura no solamente obtenemos información del pasado, sino que esta es necesaria para nuestra proyección futura, pensemos sino en los ordenadores, que operan básicamente en la modalidad escrita (Coulmas, 2003).

El término *escritura* es polisémico. En un sentido amplio, es un sistema de notación, como los gráficos, los números o las partituras musicales. Todos ellos son representaciones externas, altamente convencionalizadas (Coulmas, 2003; Tolchinsky, 2003). La escritura, en tanto que sistema notacional, representa las unidades de una lengua específica. El alfabeto (Romano), el silabario (e.g., Cherokee o Etíope), el abjad (e.g., Árabe), son tipos de sistemas de escritura (Schmandt-Besserat & Erard, 2008). También se entiende por *escritura* la actividad de poner en práctica este sistema. Por otro lado, *escritura* en un sentido estricto es el resultado de esa actividad, es un producto discursivo que tiene su concreción en el texto (Coulmas, 2003).

Existen muchos tipos de escritura con diferentes funciones, así escribimos para comunicar, recordar, identificar, etc. Grabe & Kaplan (1996, p. 4-5) proponen distinguir la escritura que requiere composición de la que no la requiere (e.g., la lista de la compra), esta distinción es especialmente útil para nuestro propósito, ya que la escritura asociada al contexto académico presupone la composición. Así la composición es un proceso en que se combinan unidades sintácticas en estructuras mayores coherentes y cohesionadas (Grabe & Kaplan, 1996). Estos autores, asimismo, proponen dividir la composición en: escritura en la que se refiere algo, *telling*, como narraciones o descripciones; y la escritura que transforma, *transforming*, tanto la información como el conocimiento, como la que se da en los textos argumentativos/persuasivos, expositivos y creativos.

Si la escritura ha suscitado interés y su estudio todavía presenta retos es básicamente porque esta permite conocer el funcionamiento de la mente. Nos detendremos en dos enfoques que tienen el estudio de la mente como objeto de análisis: la psicología cognitiva, que explora los procesos mentales internos del individuo, asume ciertos principios universales y uniformidad en el funcionamiento de la arquitectura de la mente; y la sociocultural o psicología cultural-histórica, heredera de la escuela soviética, representada por Vigotsky, Leontiev y Luria, y que enfatiza la diversidad al relacionar los procesos mentales con sus contextos históricos y culturales (Wertsch, 1993).

Desde la psicología cognitiva, Kellogg (1994) sostiene que el estudio de la escritura puede ofrecer una idea sobre la psicología del pensamiento, ya que escribir es una forma de pensar. La escritura captura la dificultad de resolver un problema, el de crear significado. Al escribir se activan cuatro operaciones fundamentales en cualquier tarea de razonamiento: generar información, planificar, traducir las ideas a texto y revisar ideas y texto (Kellogg, 1994). En consecuencia, desde este enfoque el concepto de escritura se define por los procesos cognitivos de planificar, textualizar y revisar (Castelló, Bañales & Vega, 2010). Pero además, la resolución de este problema pone en juego otras actividades cognitivas complejas en cuyo producto final, el texto escrito, se reflejan conocimientos distintos: el conocimiento del dominio (ámbito conceptual que se ha expresado en el texto); conocimiento lingüístico (las reglas gramaticales y el léxico); conocimiento pragmático (adaptación del contenido conceptual y de la forma lingüística a la audiencia); y

el conocimiento procedimental (ejecutar estratégicamente los anteriores conocimientos) (Alamargot & Chanquoy, 2001).

Por lo tanto, el producto final, el texto, incluye mucha y diversa información, que da al estudio de la escritura más riqueza metodológica si lo comparamos con tareas deductivas tradicionales de laboratorio, utilizadas, generalmente, para dilucidar el proceso de razonar. Además, el hecho de analizar trabajos que crean significado, sea a través de la escritura, la música u otros medios tiene un valor añadido, el cultural (Kellogg, 1994).

Otro motivo de interés de la escritura parte de una aproximación sociocultural de la mente, cómo su funcionamiento está mediado por modos semióticos y por contextos culturales e institucionales (Newell, 2006; Englert, Mariage & Dunsmore, 2006). El análisis sociocultural se centra en la acción humana, ya hemos apuntado en el apartado 4.1 el concepto de *acción* desde una perspectiva filosófica, en la que la acción permite revelar la identidad del sujeto y por consiguiente situarse en el mundo (Arendt, 1958). Otra perspectiva de este término la da Wertsch (1993), destacado representante de la teoría sociocultural neovigotskiana y que basa su concepto de *acción* en el de *acción teleológica* de Habermas (1984), es decir, el actor intenta conseguir un objetivo y para ello debe elegir entre diferentes medios que se presumen exitosos para tal fin. No obstante, desde el enfoque sociocultural no se asume que haya una separación nítida entre objetivo y medios, por eso se utiliza el concepto de *acción mediada*. Los individuos no actúan solos, utilizan instrumentos de mediación, como el lenguaje, y estos moldean la acción. Por lo que la acción puede ser interna o externa y realizada tanto por el individuo como por el grupo (Wertsch, 1993). Así, desde la perspectiva de la última teoría de la actividad (e.g., Engeström, 1987; Cole, 1996) y sobre la base de la tradición socio-cultural soviética, se considera que todas las acciones humanas están mediadas por elementos culturales o artefactos. Desde este marco teórico la escritura se considera una forma de acción social, ya que contribuye a la constitución de instituciones y culturas, además no se la considera una actividad autónoma, toda escritura es colaborativa y socialmente mediada (Prior, 2006; Bazerman, 2004).

Por lo tanto, desde este enfoque el concepto de escritura se define como "una actividad discursiva del escritor mediada, dialógica y situada dentro de una comunidad discursiva y ubicada en un contexto social, histórico y cultural determinado" (Castelló, Bañales & Vega, 2010, p. 1275).

De los argumentos aducidos hasta aquí podemos concluir que la escritura es un constructo multifacético y, por lo tanto, requiere de múltiples perspectivas para su análisis y comprensión. El proceso de escritura no está dissociado del producto y tampoco se puede separar del contexto inmediato, institucional o sociocultural. Por consiguiente, la escritura se concibe como un acto de identidad personal y sociocultural, en la que se expresa la individualidad creativa y, al mismo tiempo, se señala la pertenencia a una comunidad de práctica (Candlin & Hyland, 1999).

Por otro lado, la enseñanza de la escritura en todos sus sentidos y, sobre todo, la composición escrita se lleva a cabo generalmente en instituciones académicas, como escuelas, universidades, etc. La división institucionalizada de la escolaridad en educación primaria, secundaria y universitaria ha provocado que la investigación sobre la escritura también se centre en estos diferentes niveles (Bazerman, 2008); y en la escritura como lengua primera o segunda, o bien, como actividad o como producto. En el siguiente apartado nos centraremos en la escritura académica.

4.5 La escritura académica

Una de las funciones principales de las instituciones académicas es el de crear conocimiento y difundirlo a la sociedad para fomentar el bienestar y cohesión social. Esa transferencia de conocimiento se realiza básicamente a través de la escritura, ya que esta es un medio significativo para poder participar en la sociedad (Bazerman, 2004; Ivanic, 1998; Prior, 2006; Tolchinsky et al., 2013). Por lo tanto, podemos definir escritura académica como aquellos textos escritos en contextos formales, como por ejemplo, escuelas, universidades o institutos de investigación (Russell & Cortés, 2012; Snow & Uccelli, 2009).

El objeto de estudio de la escritura académica serán, en consecuencia, los textos producidos por estudiantes y profesores, principiantes y expertos de estas instituciones. Así, uno de los objetivos de enseñar a principiantes la escritura académica es prepararlos para la escritura de textos científicos, para ser profesionales competentes (Conrad, 1996; Gardner & Nesi, 2013; Russell & Cortés, 2012).

4.5.1 Funciones de la escritura

Tal como recoge Iñiesta (2009), y Castelló & Iñiesta (2012), la escritura académica tiene diversas funciones: comunicativa, epistémica, dialógica, de construcción de identidad social y, establecimiento y mantenimiento de redes de influencia.

La función **comunicativa** de la escritura académica consiste en la transmisión de conocimientos a otros miembros de una comunidad específica como, por ejemplo, la comunidad científica (Swales, 1990, 2004). Las personas pueden comunicarse entre sí de manera informal, con otros colegas de profesión, o formalmente a través de la publicación de obras en distintos medios de comunicación que persisten en el tiempo (Bazerman, 2005).

Es bien conocido que la escritura permite adquirir conocimiento, tiene un efecto **epistémico**, aunque este efecto no es exclusivo de la escritura (Alamargot & Chanquoy, 2001; Alamargot & Andriessen, 2002). Tal como señalan estos autores, la creación de nuevos conocimientos durante la escritura depende de factores tales como restricciones pragmáticas (Bereiter & Scardamalia, 1987), la reflexión y planificación (Hayes, 1996) y las limitaciones lingüísticas (Galbraith, 1999). En el modelo de Bereiter y Scardamalia (1987), *knowledge telling and knowledge transforming*, la función epistémica se da por las restricciones pragmáticas y retóricas para comunicar el conocimiento, un ejemplo de estas restricciones serían los textos argumentativos. En estos tipos de texto, el escritor tiene que convencer al lector, por lo que se plantea un problema que hay que solucionar (cómo explicar el contenido, cómo adaptarlo al lector de forma convincente).

Asimismo, en el modelo de Hayes y Flower (1980) el proceso de planificación, por abstracción, por analogía y por modelaje (Hayes & Nash, 1996), también crea nuevo conocimiento. En el modelo *knowledge-constituting* (Galbraith, 1999) se pone el énfasis en que cuando se escriben las ideas como borrador (*rough drafting*), sin restricciones de organización, ni retóricas, es cuando se generan nuevas ideas. El tener que traducir ideas complejas a palabras escritas genera nuevo conocimiento.

Por otra parte, la construcción de la mente es social y el lenguaje desempeña un papel mediador. Así como la mente se forma dialógicamente, los textos también son **dialógicos**: "La relación con el significado es siempre dialógica" (Bakhtin, 1986, p.121). Cada función en el desarrollo cultural del niño aparece dos veces: primero en el plano social, y más tarde, en el nivel individual; primero entre personas (interpsicológicos), y después en el niño (intrapsicológico) (Vygotsky, 1978). Cada expresión se basa en expresiones anteriores, en la historia del uso de la lengua, y cada expresión hace avanzar la historia (Volosinov, 1986, *apud* Bazerman, 2004). Los textos tienen una variedad de expresiones, adoptan una postura de polifonía o dialogismo (Bakhtin, 1981).

Basándose en estos autores, Kristeva (1980) desarrolló el término de intertextualidad. Cada texto está relacionado con otros textos. El escritor utiliza recursos de otros textos y estos, a su vez, están relacionados con textos anteriores. La escritura es el producto de su contexto social, no es una actividad autónoma, toda escritura es el producto de la colaboración y de mediación social, siempre hay una co-autoría en el texto (Prior, 2006). La intertextualidad no es solo una referencia a otros textos, también indica cómo se posiciona el escritor para hacer su propia declaración (Bazerman, 2004). Algunos estudiosos aplican el concepto de intertextualidad al aprendizaje de la escritura académica (e.g., Porter, 1986; Recchio, 1991). El escritor se identifica con una comunidad discursiva particular, por lo que utilizará los recursos textuales pertenecientes a esa comunidad. Por lo tanto, la **identidad social** (e.g., como investigador o autor académico) se construye discursivamente (Berkenkotter & Huckin, 1995; Ivanic, 1998; Lillis, 2001).

Por otra parte, la escritura sirve para **establecer/mantener redes de influencia**. Las comunidades discursivas mantienen su organización, su poder, su estado y su actividad principalmente a través de las redes de textos (por ejemplo, los artículos publicados) (Bazerman & Paradis, 1991; Myers, 1989).

4.6 Registro

Un *registro* es un conjunto de características léxicas, gramaticales y discursivas que se dan en contextos particulares de uso del lenguaje y que cumplen unos objetivos específicos (Halliday & Hassan, 1989; Biber, 1995; Schleppegrell, 2001; Uccelli, Dobbs & Scott, 2013). Por lo tanto, el análisis del registro analiza la variación lingüística, esto es, las características inherentes al lenguaje que describe la elección sistemática que hace un hablante de la pronunciación, del vocabulario o de las estructuras gramaticales dependiendo de factores no lingüísticos, como factores sociales o contextos situacionales (Biber & Conrad, 2009). Desde una perspectiva de desarrollo, la adquisición del registro consiste en llegar a dominar la amplia gama de opciones expresivas de una lengua y ser capaz de utilizarlas de acuerdo a los diferentes contextos comunicativos (Ravid & Berman, 2009). El proceso educativo formal implica enfrentarse a una gran variedad de registros. A partir de la educación secundaria y, sobre todo, en el nivel universitario los alumnos han de llegar a dominar los registros especializados de las diferentes disciplinas (Biber & Conrad, 2009; Schleppegrell & Colombi, 2002).

Desde la perspectiva de registro se analizan las características lingüísticas que son comunes a unos textos de una variedad y se analiza la situación de uso. Las características siempre son funcionales desde esta perspectiva, es decir, hay características que tienden a darse en un registro específico, porque están determinados por los propósitos y el contexto situacional concreto.

El análisis desde la perspectiva de género es similar al análisis del registro, pero difiere en que se centra en las estructuras convencionales que se utilizan para componer un texto, por ejemplo, el inicio y final de una carta (Biber & Conrad, 2009). Desde la *lingüística sistémico-funcional* el registro refiere al contexto situacional, que es una configuración de campo, tenor y modo. Como ya hemos mencionado en otro apartado, campo sería lo que se lleva a cabo; modo sería el canal de la comunicación y tenor serían las relaciones entre los participantes. Estos tres valores del contexto de la situación tienen su realización lingüística en las metafunciones (Halliday, 1978). La función *ideacional* contiene recursos para la creación de contenido; la función *textual*, recursos para la organización de textos; la función *interpersonal*, recursos para interactuar. Así, por ejemplo, el lenguaje de un género de investigación en biología será parecido a uno de historia porque tienen como objetivo explicar una investigación. Sin embargo, el registro puede ser diferente, ya que será un reflejo del contexto de situación, qué es lo que se explica (*campo*), quién lo escribe (*tenor*) y el rol que tiene el lenguaje en la interacción (*modo*) (Nesi & Gardner, 2012). Así, las diferencias entre registros dependen del nivel de generalidad del análisis. Por ejemplo, registros generales como el hablado o escrito se definirán básicamente por el modo. En cambio, registros más específicos, como la sección de metodología de un artículo de biología, se definirán por las circunstancias de producción, la audiencia a la que va dirigida, etc. (Biber, 1995). El análisis del registro supone identificar las características léxico-gramaticales que se dan en situaciones sociales específicas (Biber & Conrad, 2009; Nesi & Gardner, 2012).

4.6.1 Léxico-gramática

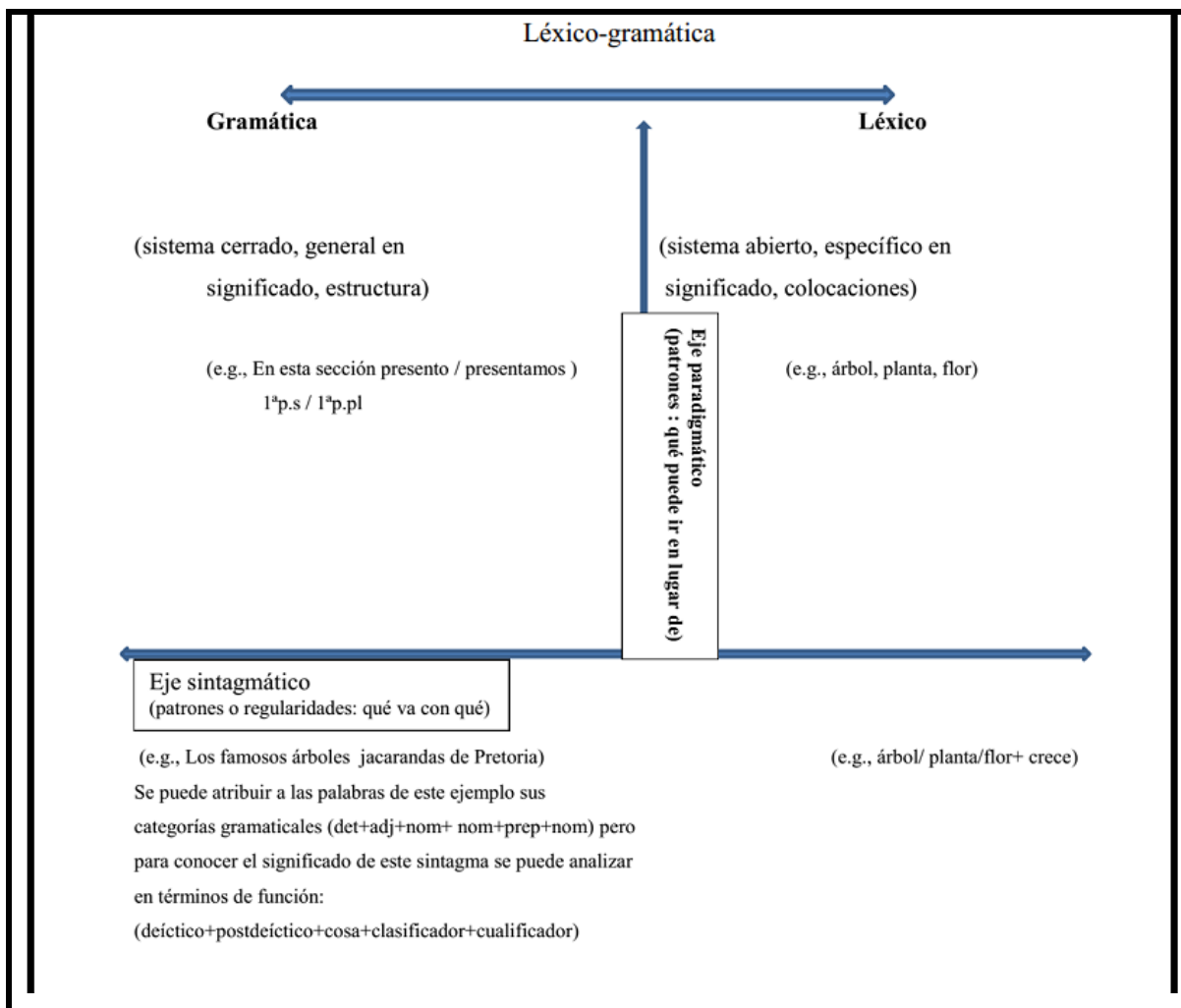
Para describir el concepto teórico de léxico-gramática de la *lingüística sistémico-funcional* sobre el cual está basada esta investigación hay que tener en cuenta algunas consideraciones previas. Podemos diferenciar dos enfoques respecto al lenguaje: el formal y el funcional. Respecto al primer enfoque, el lenguaje se entiende desde una perspectiva formal y sintáctica, sobre todo por la influencia de la obra de Chomsky. Y una aproximación al estudio del lenguaje asume que el lenguaje es esencialmente un fenómeno cognitivo, formado por un sistema complejo y abstracto de reglas.

Así, las teorías formales conciben la gramática como regla y su objetivo es identificar las propiedades de esos sistemas de reglas. Desde esta perspectiva, la relación entre gramática y léxico es modular. La sintaxis, la morfología y el léxico son módulos independientes.

Por otro lado, desde las teorías funcionales se entiende el lenguaje desde una perspectiva funcional y semántica. Concibe el lenguaje en términos de la función social a la que sirve. Se asume que el lenguaje es un recurso sistemático para expresar significado en un contexto. Así las teorías funcionales conciben la gramática como recurso para crear significado y su objetivo es describir los recursos disponibles en el lenguaje para expresar significado.

Desde una perspectiva funcional la relación entre gramática y léxico es continua (ver figura 4.1). Son diferentes perspectivas sobre un fenómeno, el léxico-gramatical (Martínez Lirola, 2007; Matthiessen, 2009). Como vemos en la figura 4.6.1 los significados que están lexicalizados están contruidos como específicos, en sistemas abiertos, en colocaciones y suelen estar muy taxonomizados, por ser de dominios específicos. Los significados gramaticalizados están contruidos como generalidades, en sistemas cerrados, en estructuras. La *gramática sistémico-funcional* se basa en la distinción saussureana de relaciones paradigmáticas y sintagmáticas. En toda descripción del lenguaje se pueden establecer relaciones paradigmáticas, relativas a la sustitución de elementos lingüísticos por otros en un contexto particular; o relaciones sintagmáticas, relativas a la ordenación de los elementos lingüísticos dentro una unidad más grande. La *gramática sistémico-funcional* favorece el interés por las relaciones paradigmáticas (opciones) sobre las sintagmáticas (estructura) (Bednarek, 2010; Martínez Lirola, 2007). En cambio, desde la *lingüística de corpus*, teniendo en cuenta las colocaciones o fraseología, se considera el *principio de elección abierta* como un modelo secundario. Se asume que los dos ejes, paradigmático y sintagmático, tienen el mismo potencial de significado (Sinclair, 2000).

Figura 4.6.1 Léxico-gramática en un continuum

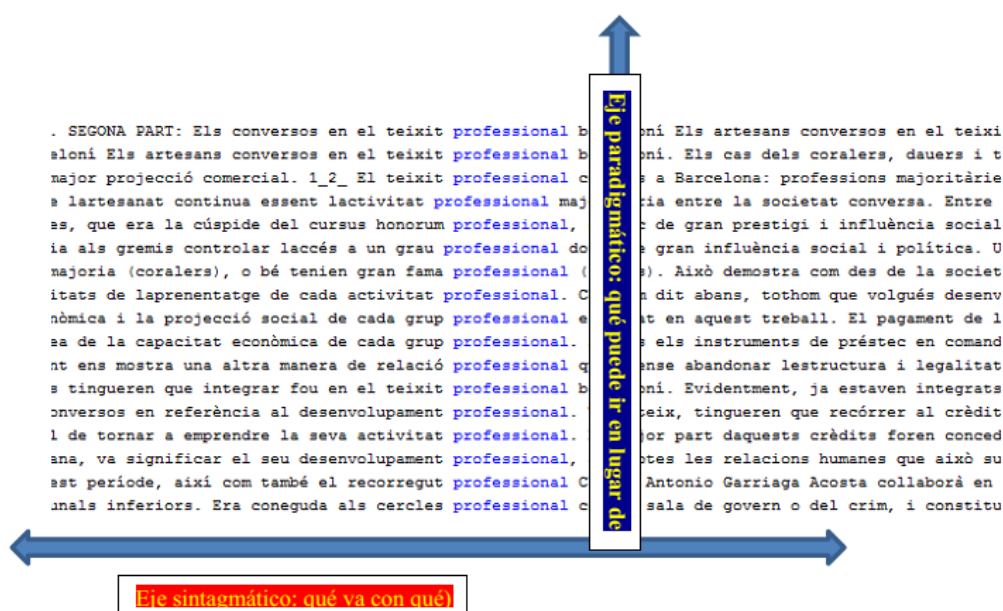


(Adaptado de Halliday 1985 y Matthiessen, 2009)

En esta figura podemos ver los extremos del continuum, el gramatical y el léxico. Asimismo, mostramos ejemplos de la gramática y léxico desde una perspectiva paradigmática y sintagmática. Estos son los extremos, pero ¿qué es lo que se encuentra en medio de estos polos? Lo que los teóricos sistémico-funcionales llaman la léxico-gramática o sistema fraseológico (Halliday, 1985; Matthiessen 2009). Desde un punto de vista sintagmático se encuentran patrones que combinan estructuras y colocaciones. Tal como manifiesta Halliday (1985), estos patrones no se dilucidan fácilmente con la introspección, es gracias al corpus que podemos hacer una descripción de ellos. Asimismo, Sinclair, 2004, *apud* Bednarek, 2010, señala que: "...concordance lines show the essence of Saussurean linguistics in a nutshell"

En la figura 4.6.2 ejemplificamos con los trabajos de historia de final de máster (Corpus TARBUC) esta afirmación. En el eje sintagmático vemos que *profesional* va precedido en muchos casos por los sustantivos: *tejido*, *actividad*, *fama*, *grupo*, *relación*, *recorrido*, etc. En el eje paradigmático vemos como el sustantivo *tejido* puede ser sustituido por *actividad*, *grado*, *fama*, etc. Es decir, que hay una posibilidad de elección.

Figura 4.6.2 Líneas de concordancia del programa AntConc con el eje sintagmático y paradigmático



Por otro lado, para determinar cómo está organizada la complejidad del lenguaje nos tendríamos que fijar en el léxico y en la gramática. Cada uno revela esta complejidad de modo diferente. El léxico se muestra a través de la densidad léxica y la gramática, a través de una mayor intrincación (Halliday, 1985).

4.7 Metadiscurso

Metadiscurso es un término que se utiliza, sobre todo, en estudios de análisis del discurso y lingüística aplicada y está basado en la idea de que los textos escritos son una forma de interacción social. En los textos escritos el escritor y lector han de negociar los significados, el escritor ha de crear un discurso convincente para que sea aceptado por el lector. El concepto de *metadiscurso* refiere a los recursos lingüísticos interpersonales utilizados por el escritor para organizar un discurso o, más concretamente, refiere a la postura del escritor hacia el contenido y al lector (Hyland, 2005; Hyland & Tse, 2004). El término fue utilizado por primera vez por Harris (1959) y desarrollado posteriormente por Williams (1981), Vande Kopple (1985) y Crismore (1989). En la actualidad es reconocido el trabajo de Hyland (2005) sobre *metadiscurso* en escritura académica. Las interacciones entre escritor y lector han sido tema de interés en las perspectivas sistémico-funcionales y socio-constructivistas, pero los recursos o marcadores lingüísticos que sirven a la función interaccional han recibido diferentes nombres, tales como *evidentiality* (Chafe, 1986) *affect* (Ochs & Schieffelin, 1989) *evaluation* (Thomson & Hunston, 2000) *appraisal* (Martin, 2000) *stance* (Biber & Finegan, 1989) y *discourse stance* (Berman, Ragnarsdóttir & Strömqvist, 2002), entre otros, tal como recogen Gray & Biber (2012).

Otro concepto relacionado con el *metadiscurso* o, más concretamente con el posicionamiento, es el de voz de autor, que revela la presencia del autor. Este constructo tiene tres dimensiones: la individual, la social y la dialógica (Tardy, 2012). La dimensión individual se refiere a la representación de uno mismo en el texto, a la posición y opinión del escritor, lo que Ivanic (1998) denomina *self*, la identidad discursiva y relacional del escritor. En este sentido, la voz está relacionada con *autoridad* y *autenticidad*. La *autoridad* se refiere a la capacidad de ser creíble y convincente para la comunidad discursiva. La *autenticidad* es la habilidad del escritor para ser consistente con las prácticas discursivas de la comunidad a la que pertenece el escritor. En definitiva, la voz está relacionada con el posicionamiento y estilo (Nelson & Castelló, 2012).

La dimensión social de la voz está influenciada por el contexto. El escritor selecciona determinados textos y autores para alinear su trabajo con una comunidad específica, elige entre formas particulares de discurso. Por lo tanto, un texto nunca es totalmente original, en él se dan una multiplicidad de voces, una *heteroglosia* en el sentido de Bakhtin (1981), que resalta que no solamente los diferentes autores, que ha consultado el escritor, se manifiestan en el texto, sino también el género del que hace uso el escritor contribuye a esta polifonía (Tardy, 2012). El llegar a desarrollar una voz académica propia es una tarea compleja. Tal como sintetiza Ivanic (1998), "The only way an apprentice member can learn to become a full member is by copying, adapting and synthesizing from the works of other members" (p.3). Hay que entender esta afirmación en su contexto, Ivanic (1998) compara la escritura con las bellas artes, en el Renacimiento los aprendices aprendían copiando, adaptando y sintetizando aspectos de otros artistas hasta crear una obra propia.

¿Qué tienen en común estas dos actividades? Por un lado, son el producto de su contexto social. La escritura, como ya hemos mencionado, no es una actividad autónoma, toda escritura es el producto de una colaboración y mediación social, es decir, hay una coautoría en cada texto. El escritor utiliza recursos de textos previos y estos aparecen en relación con otros textos (intertextualidad). La intertextualidad no es solo una referencia a otros textos, sino que también indica cómo se posiciona el escritor para hacer su propia declaración. El escritor utilizará aquellos recursos textuales que pertenecen a la comunidad discursiva con la que se identifica (Bazerman, 2004; Ivanic, 1998; Prior, 2006). No obstante, para evitar que los estudiantes confundan esta utilización de recursos textuales con el plagio son necesarias estrategias instructivas para que aprendan a desarrollar su propia voz siguiendo los requisitos normativos de la escritura académica (Castelló, Iñiesta, Pardo, Liesa & Martínez-Fernández, 2011b). La dimensión dialógica refiere a la co-construcción de la voz del escritor con el aspecto social, pero también a la percepción que tiene el lector de esa voz (Tardy, 2012).

Como hemos visto, el concepto de voz está relacionado con los recursos interaccionales del metadiscurso, es decir, la postura y afiliación, marcadores que analizamos en nuestro estudio, ya que tal como sugiere Hyland (2004b), "Metadiscourse is particularly important at advanced levels of academic writing as it represents writers' attempts to present and negotiate propositional information in ways that are meaningful and appropriate to a particular disciplinary community"(p.134). Asimismo, Snow & Uccelli (2009) señalan que el uso de la taxonomía de marcadores metadiscursivos propuesta por Hyland (2005) podría resultar relevante para el estudio del lenguaje académico desde una perspectiva de desarrollo. Por consiguiente, hemos utilizado el modelo de Hyland (2005) y Hyland & Tse (2004), que se basa en la distinción que hace Thompson (2001) entre los recursos *interactivos* e *interaccionales*. Los marcadores *interactivos* o *textuales* se utilizan para organizar el discurso. Permiten al escritor gestionar el flujo de la información y de forma explícita guiar al lector en la interpretación que se desea conseguir. Los marcadores que se utilizan para este propósito son, por ejemplo, los conectores textuales, estructuradores de la información, reformuladores, etc.

Los recursos *interaccionales* se utilizan para adoptar un punto de vista sobre el contenido e involucrar al lector para que acepte sus argumentaciones. Un lector, que al pertenecer a la comunidad discursiva específica, tiene sus propias opiniones sobre el tema y sobre las convenciones retóricas que se deberían utilizar para resaltar la importancia y credibilidad del texto (Hyland, 2005). Para ello se utilizan marcadores que tienen una función de posicionamiento o de afiliación.

El *posicionamiento* o *postura* puede definirse como la expresión de la actitud, sentimientos, juicios o compromisos con el contenido proposicional de un mensaje a través de elementos gramaticales y léxicos (Biber & Finegan, 1989). La función de *posicionamiento* se expresa por medio de los marcadores de auto-mención, intensificadores, atenuadores, de actitud.

La *afiliación*, por el contrario, se refiere a características lingüísticas utilizadas por el escritor para involucrar al lector, para guiarlo a la interpretación de un argumento y convertirlo en un participante en el discurso (Hyland, 2008a). Para expresar la función de *afiliación* se utilizan marcadores como los pronombres del lector, instrucciones que dirigen a otras partes del texto, al argumento, al conocimiento compartido, etc. (Hyland, 2005).

Respecto a la *postura*, los marcadores de auto-mención (e.g., *nosotros*, *nuestro*) son recursos lingüísticos útiles de la función interpersonal, ya que ponen de manifiesto la presencia del autor (Martínez, 2005; Mur, 2007). Además, los marcadores de la modalidad epistémica, los intensificadores y atenuadores o mitigadores (e.g., *seguro*, *podría*), que expresan el compromiso o la falta de compromiso del escritor con el contenido de la proposición, han sido objeto de muchas investigaciones, ya que expresar adecuadamente las aserciones es un aspecto destacado en la comunicación escrita del conocimiento científico (Crismore, Markkanen, & Steffensen 1993; Lafuente, 2008; Salager-Meyer, 1994, Vázquez & Giner, 2008). Sin embargo, los intensificadores han sido estudiado menos extensivamente que los atenuadores, aunque los dos conceptos son igualmente interesantes de investigar (Vázquez & Giner, 2008).

Hay un consenso generalizado en que los miembros de la comunidad académica no pueden hacer afirmaciones categóricas acerca de sus resultados (Lafuente, 2008), más bien han de expresar los resultados como posibilidad, es decir, proyectar una imagen de humildad (Swales, 1990; Hyland, 1996). Más allá de esto, el hecho es que el uso de marcadores de certeza tiene que ver con algunas limitaciones establecidas en la investigación científica profesional. Una suposición general es que cuanto más robustos, estadísticamente hablando, son los resultados se requieren menos atenuadores, aunque los resultados en el uso de estos dos marcadores podrían variar, especialmente en determinados artículos científicos, debido a que influyen factores como el patrocinio de las investigaciones o el factor de impacto de la revista (Gross & Chesley, 2012).

No obstante, los atenuadores son esenciales en la escritura académica y aclarar la motivación para su uso es interesante (Markkanen & Schröder, 1997; Lewin, 2005), especialmente en el campo de inglés para fines específicos, ya que incluso los estudiantes avanzados de L2, tienen dificultades en el uso adecuado de los atenuadores (Hyland, 1996). Los atenuadores no son intrínsecos a los textos; sus significados se realizan a través de la interacción escritor-lector y esto está relacionado con el discurso específico de cada comunidad (Markkanen & Schröder, 1997).

Por medio de los atenuadores, los escritores pueden presentar una proposición como una posibilidad o una opinión, en lugar de un argumento factual y esto podría ayudarles a ganar la aceptación del lector, de la comunidad académica y, por lo tanto, repercutir en la calidad de su trabajo (Hyland, 1998). Por otro lado, las funciones de los atenuadores han sido definidas por los lingüistas de diferentes maneras. Lakoff (1973) muestra que la asignación de la verdad y la falsedad de conceptos en lenguaje natural es una cuestión de grado y, por lo tanto, los atenuadores expresan vaguedad. Un punto de vista alternativo considera los atenuadores como un recurso para sustentar una aserción precisa: "the strongest claim a careful researcher can make" (Salager-Meyer, 1994, p. 151). Otros estudiosos, como Myers (1989) o Swales (1990), consideran los atenuadores como una estrategia para expresar cortesía, honestidad y modestia, todo ello para participar en las interacciones sociales de la comunidad científica.

Los escritores de todas las disciplinas, incluyendo las tradicionalmente más exigentes en cuanto a contenido, como la biología o la medicina, utilizan atenuadores (Hyland, 2005; Oliver del Olmo, 2008). Desde la perspectiva del desarrollo cognitivo, el uso de atenuadores implica considerar múltiples perspectivas y esto es un factor que está ligado al desarrollo de un escritor académico competente (Berman, 2004). Investigaciones sobre marcadores metadiscursivos han demostrado que, independientemente de la disciplina, los escritores académicos más avanzados utilizan más atenuadores (Hyland & Milton, 1997; Aull & Lancaster, 2014).

Aun así, el uso de más atenuadores, por sí mismo, no siempre es un indicador de la calidad de la escritura, ya que los estudiantes podrían haber aprendido a utilizar los atenuadores como meras fórmulas léxicas y aplicarla a los textos de forma indiscriminada (Hyland & Milton, 1997). No obstante, hay que tener en cuenta que la mayoría de los manuales de escritura académica en español sugieren que, en la escritura académica, habría que obviar las referencias al escritor o al lector. Además, Mendiluce (2005), después de haber efectuado una extensa revisión de manuales de escritura científica, concluye que, en general, no se tienen en cuenta los marcadores discursivos, como intensificadores o atenuadores. En cambio, a los estudiantes de universidad se les pide cada vez con más frecuencia que adopten una posición explícita sobre las cuestiones sobre las que escriben, con el fin de construir su propia voz (Castelló et al., 2011a). Por consiguiente, para los estudiantes es un desafío hacer frente a las estrategias retóricas propias de las comunidades académicas.

CAPÍTULO 5

5. El análisis del discurso y la lingüística de corpus

Como ya hemos mencionado, nuestro objetivo es caracterizar el registro del trabajo de investigación mediante la función metadiscursiva interaccional y las propiedades léxico-gramaticales. Esto comporta hacer un análisis cualitativo y cuantitativo. Por lo tanto, a continuación expondremos los argumentos que justifican la decisión de adoptar conjuntamente el *análisis del discurso* (AD) y la *lingüística de corpus* (LC) como la metodología más adecuada para este propósito.

Ambas tradiciones tienen en común que se basan en el estudio de la lengua en uso e intentan identificar patrones en datos lingüísticos naturales, es decir, en datos empíricos. En la interpretación de los resultados tienen en cuenta el contexto social y cultural en el que se da el discurso (Calsamiglia & Tusón, 1999; Charles, Pecorari & Hunston, 2009; Conrad, 2002). Generalmente, se pueden considerar los estudios de LC como un tipo de AD, ya que describen el uso de formas lingüísticas en su contexto (Biber, Connor & Upton, 2007).

Cualquier texto se puede analizar desde una perspectiva global o local, la perspectiva global tiene en cuenta el texto en su conjunto, su contenido, estructura, etc., mientras que la perspectiva local tiene como objeto de análisis las estructuras lingüísticas (Calsamiglia & Tusón, 1999). El análisis del discurso se suele definir como el estudio más allá del nivel de oración y del párrafo, cómo estos se combinan para crear significado y cumplir diferentes propósitos, mientras que las técnicas usadas en la LC tienen como finalidad identificar patrones recurrentes en un nivel inferior, es decir, palabras o frases (Charles, Pecorari & Hunston, 2009).

A veces se utiliza el término *pragmática* para el estudio del significado contextualizado de la lengua en uso, mientras que el término *análisis del discurso* se reserva para mostrar como las oraciones se combinan en el *texto* para crear significado y coherencia y así formar los diferentes géneros, por ejemplo, el narrativo, expositivo, etc. Nosotros concebimos *análisis del discurso* como inclusivo de ambos términos, el estudio de la lengua en uso y el estudio de textos (Gee & Handford, 2012, p. 1), ya que tal como sugieren Calsamiglia & Tusón (1999), "todo análisis del discurso es pragmática" (p.11).

Por otro lado, el utilizar la LC y el AD como metodologías complementarias nos puede ofrecer un análisis más rico, de hecho, en la práctica se utilizan cada vez más la combinación de ambas metodologías (Ädel & Reppen, 2008; Biber, Connor & Upton, 2007; Charles, Pecorari & Hunston, 2009; Conrad, 2002). Una descripción detallada de diversos estudios que analizan la escritura académica combinando estas dos metodologías se puede ver en Charles, Pecorari & Hunston, 2009; Conrad, (2002).

5.1 El análisis del discurso

El análisis del discurso es un método cualitativo para analizar textos, explora la conexión entre lenguaje, conocimientos, poder y prácticas sociales (Muncie, 2006), aunque también se ha descrito como una metodología o perspectiva teórica (Phillips & Hardy 2002).). El AD examina patrones del lenguaje a través de textos y considera las relaciones entre lenguaje y sus contextos sociales y culturales, cómo el uso del lenguaje presenta diferentes maneras de ver y entender el mundo, cómo está influenciado por las relaciones sociales, y también cómo las visiones del mundo están construidas por medio del uso del discurso (Paltridge, 2012, p. 2). De esta definición se desprende que el lenguaje es una forma dinámica de práctica social. Esta concepción del lenguaje tiene sus raíces en la obra de Wittgenstein (1953), en su noción de *juegos de lenguaje*, para este autor el significado es la aplicación del signo que hacemos, es decir, el uso que hacemos de él. Los significados son abiertos, es decir, que requieren del contexto para su interpretación.

Austin (1962) propuso que el lenguaje es una forma de acción que permite cambiar cosas y que es una acción socialmente situada. Posteriormente, Foucault (1971) consideró el lenguaje, en concreto, el *discurso* como el eje principal para entender las prácticas sociales (Bhatia, Flowerdew & Jones, 2008; Calsamiglia & Tusón, 1999; Jørgensen & Phillips, 2002). Halliday (1973) es otra figura destacada que considera el lenguaje como un sistema de significados relacionado con el contexto social. De hecho, su *gramática sistémico-funcional* (SFG) ha influido en el análisis del discurso contemporáneo, como, por ejemplo, el análisis del discurso crítico y el análisis del discurso multimodal (Bhatia, Flowerdew & Jones, 2008).

Una primera aproximación al concepto de *discurso* nos lleva a considerar estas tres dimensiones: a) lenguaje en uso, b) comunicación de creencias o ideas, c) interacción en una situación social. Así, un discurso es una forma de lenguaje en uso que se da en un evento comunicativo, para comunicar creencias, ideas o emociones, pero más que simplemente comunicar, las personas interactúan en situaciones sociales específicas. Los estudios sobre discurso tienen como objeto de análisis las modalidades, oral y escrita, y el contexto, la situación social o evento comunicativo en que se produce la conversación o texto (Van Dijk, 1997, p. 2).

5.1.1 Enfoques del análisis del discurso

Se puede abordar el estudio del discurso desde diferentes disciplinas como la lingüística, psicología, las ciencias sociales o la filosofía, entre otras, pero todas tienen en común que intentan comprender la estructura y la función del uso del lenguaje para comunicar significado (Bhatia, 1993; Van Dijk, 1997).

El término *discurso* ha sido objeto de muchas definiciones, ya que como acabamos de ver, su área de estudio es muy amplio y puede abordarse desde muchas perspectivas (Biber, Connor & Upton, 2007). Según la orientación teórica adoptada, en un extremo del continuum nos encontramos con estudios que se centran en características formales y funcionales del uso del lenguaje en los textos, mientras que en el otro extremo predominan las perspectivas socioculturales del uso del lenguaje, desde las cuales se destaca la acción

social de la comunicación (Bhatia, 1993; Paltridge, 2012). En el primer caso se incluyen los trabajos que son una extensión del formalismo gramatical, como la lingüística de texto y los estudios que analizan el registro y género desde la perspectiva sistémica. En cambio, ejemplos del otro extremo son aquellos trabajos que analizan los géneros profesionales y de investigación y, los de tradición etnometodológica (Bhatia, 1993).

Según Schifffrin, Tannen & Hamilton (2001) las definiciones de *análisis del discurso* se pueden agrupar en tres categorías generales: 1) el estudio de la *lengua en uso*, 2) el estudio de la *estructura lingüística más allá de la frase*, y 3) el estudio de las *prácticas sociales asociadas al lenguaje y la comunicación*.

5.1.1.1. El estudio de la lengua en uso

La *lengua en uso* se ha estudiado desde diferentes perspectivas como, por ejemplo, la pragmática o la lingüística funcional. Por lo tanto, es una aproximación más funcional al discurso, cuyos representantes más destacados serían Halliday, Austin, Grice, Widdowson (Jones, 2012). Su objeto de análisis es cómo se utilizan las palabras y las estructuras lingüísticas para expresar significado. Esta categoría, en general, aborda el problema de la variación lingüística en el discurso o texto, es decir, muestra que las elecciones de elementos lingüísticos son sistemáticas y siguen unos principios cuando aparecen en un contexto discursivo. Utilizan, generalmente, un análisis *bottom-up*, se analiza primero el léxico y la forma y, a raíz de ellos, emergen los patrones de los diferentes tipos de discurso. Los estudios de la *lengua en uso* que utilizan corpus estudian la distribución y las funciones de las características lingüísticas superficiales. La mayoría de estudios basados en corpus (*corpus-based*) se adscriben a esta categoría. Por tanto, son estudios cuantitativos que utilizan extensos corpus, analizados con herramientas computacionales (Biber, Connor & Upton, 2007; Upton & Cohen, 2009).

5.1.1.2. El estudio de la estructura lingüística más allá de la frase

La estructura lingüística *más allá de la frase* se ha estudiado desde la perspectiva lingüística, cognitiva y computacional. Es una perspectiva más formal del discurso, que tiene sus raíces en el formalismo ruso (V. Propp), que aplicó las mismas técnicas de análisis gramatical de las oraciones para el estudio de cuentos (Jones, 2012). Tiene como objeto de estudio la organización interna de los textos en un nivel superior de la frase (e.g., párrafos), para ello utilizan un análisis *top-down*. Se describen las funciones discursivas que realizan determinadas características léxico-gramaticales que dan, por ejemplo, coherencia y cohesión al texto. Los estudios de esta categoría suelen ser cualitativos, se analizan pocos textos y de un género determinado, ya que su análisis es laborioso, porque debe hacerse manualmente (Biber, Connor & Upton, 2007; Upton & Cohen, 2009).

Tanto el estudio de la *lengua en uso* como el estudio *más allá de la frase* comparten un enfoque principal en la forma lingüística y cómo las estructuras de la lengua se utilizan para la comunicación (Biber, Connor & Upton, 2007).

Además, Charles, Pecorari & Hunston (2009) consideran el análisis del discurso y los estudios basados en corpus no como metodologías separadas, sino como constituyentes de un continuum. Para estos autores, los análisis *bottom-up* se situarían en el continuum más cerca del estudio basado en corpus, mientras que el enfoque *top-down* estaría más próximo al análisis del discurso.

5.1.1.3. El estudio de las prácticas sociales asociadas al lenguaje y la comunicación

Esta categoría se centra en la construcción social del discurso. Su interés de estudio se centra en cómo las comunidades especializadas usan el texto, pero, generalmente, el análisis de las estructuras lingüísticas y su uso no constituyen su principal objeto de estudio. Más bien se interesa por las acciones de los participantes de un evento comunicativo o las relaciones de poder (Biber, Connor & Upton, 2007; Upton & Cohen, 2009).

5.2 El análisis del discurso especializado

Por otro lado, se sugiere que el discurso especializado debería ser analizado de manera integral para una mejor comprensión de su naturaleza retórica (Pérez-Llantada, 2012). Bhatia (2002b), justamente, propone un modelo multinivel y por lo tanto holístico para analizar el discurso científico. Este modelo integra el discurso como texto, que implica un conocimiento textual; el discurso como género, conocimiento del género; y el discurso como práctica social, conocimiento social.

El discurso como texto refiere al análisis del uso del lenguaje en un nivel superficial, incluye aspectos formales y funcionales, por ejemplo características léxico-gramaticales, cohesión, tema-remata, etc. Es decir, se centra en las propiedades de la construcción del producto textual, más que en la interpretación o uso de ese producto. Así, la pregunta asociada a este nivel es: ¿Qué características léxico-gramaticales son estadísticamente significativas y/o distintivas en cuanto a su función?

El discurso como género incluye el contexto en un sentido amplio, es decir, no solo cómo se ha construido el texto sino también cómo se interpreta o se usa para fines específicos en el contexto social, institucional o profesional. Las preguntas relacionadas con este nivel suelen ser de carácter socio-cognitivo y etnográfico: ¿Por qué las personas construyen el discurso de la manera que lo hacen y qué es lo que lo hace posible?

El discurso como práctica social considera el contexto social. ¿Cómo los géneros mantienen o cambian las estructuras sociales o relaciones profesionales, es decir, cómo se limita el discurso por las prácticas sociales, identidades y estructuras sociales? (Bhatia, 2002b, p. 17). Tal como remarca este autor, los niveles son complementarios y depende del objetivo planteado que el discurso se pueda abordar desde el nivel más contextual, si el interés es profundizar en una comprensión sociolingüística e ir bajando de nivel; o bien, desde un nivel textual, que es, generalmente, el objetivo de la lingüística aplicada y que puede encontrar en los niveles superiores las explicaciones a las características léxico-gramaticales y discursivas.

En este apartado hemos visto que el *análisis del discurso* permite abordar desde diferentes perspectivas la complejidad inherente a cualquier acto de comunicación y, por lo tanto, también del discurso académico escrito. Pero como ya hemos mencionado, una integración del *análisis del discurso* con la *lingüística de corpus* puede contribuir a una descripción más amplia y detallada de los artículos de investigación. Por eso, a continuación explicaremos las características de la lingüística de corpus.

5.3 La lingüística de corpus

En primer lugar, presentamos las características de la *lingüística de corpus* (de ahora en adelante: LC) y la discusión imperante en torno a considerarla una teoría o metodología (apartado 5.3.1). En segundo lugar, repasamos los dos enfoques más distintivos *corpus-driven* versus *corpus-based* (apartado 5.3.2). Seguidamente repasamos las diferentes definiciones de corpus (5.3.3). En el apartado (5.3.4) exponemos la relación de la LC con el marco teórico que adoptamos, es decir, el funcionalismo. Al final de cada apartado explicitamos nuestro posicionamiento, es decir, las decisiones que hemos tomado respecto a las diversas cuestiones que se plantean desde la LC.

Para explicar la lingüística de corpus tomaremos como hilo conductor la conocida analogía que Fillmore (1992) hizo sobre las dos corrientes, la empírica y la racionalista, que caracterizan varios aspectos de la LC, representándolas en dos tipos de lingüistas. Por un lado, el lingüista de corpus estudia hechos concretos; y por el otro, el *armchair linguist*, lingüista de sillón, que piensa sobre los hechos. Concluye diciendo que estos dos tipos de lingüistas tienen que coexistir, que se necesitan mutuamente. Veremos cómo esta dicotomía se refleja al considerar la LC como una teoría o una metodología y al considerar la manera de analizar los datos: *corpus-driven* versus *corpus-based*. Finalmente constataremos que en la actualidad se diluye esta dualidad y se postula por aunar las dos vertientes.

5.2.1 Entre la teoría y la metodología

La metodología utilizada en este estudio se basa en corpus. La LC se refiere a una base metodológica que permite el estudio del lenguaje basado en ejemplos reales de usos del lenguaje (McEnery & Wilson, 1996). De esta definición se desprende que la LC pertenece al ámbito de la lingüística, pero no en el sentido como pudiera pertenecer la sintaxis, la semántica, etc., cuya función es describir algún aspecto del lenguaje, sino que es una metodología que puede utilizarse en la mayoría de ámbitos de la lingüística, por lo tanto pertenece a la lingüística aplicada (McEnery & Wilson, 1996). La lingüística aplicada se preocupa de buscar soluciones a los problemas que tienen que ver con el lenguaje, sus usos y usuarios (Berns & Matsuda, 2006). Es por ello que su dominio de investigación no tiene unos límites definidos, más bien, como nos indica Payrató (1997), es una dimensión particular de la lingüística en que se aúnan disciplinas tradicionales o “inmanentes” y “externas”. Aunque la LC pertenece a la lingüística aplicada se diferencia de otras disciplinas, que se engloban bajo esta dimensión, por su metodología.

El hecho de si la LC es considerada propiamente una teoría (e.g., Leech, 1992; Stubbs, 1993; Tognini-Bonelli, 2001; Teubert, 2005) o una metodología (e.g., Biber, Conrad, Reppen, 1998; McEnery & Wilson, 1996) ha sido objeto de debate (Gries, 2010).⁴

Así, por ejemplo, Tognini-Bonelli (2001) defiende su visión de la LC como una disciplina independiente que llama provisionalmente *corpus-driven linguistics* (CDL) con argumentos y ejemplos. Señala que autores como Halliday (1993) consideran que la LC no solamente se basa en recopilar datos cuantitativos sino que además los lingüistas teorizan sobre estos datos, lo que conlleva un cambio cualitativo en la comprensión que tenemos sobre el lenguaje. Lo que sí parece innegable es que la LC conlleva el desarrollo de teorías sobre el lenguaje que inducen a replantearse ciertos postulados firmemente asentados (Hunston, 2002). Esta contribución al desarrollo de teorías se puede atribuir al estatus teórico que tiene la LC y que le confiere su metodología particular. Si una metodología es el uso de un conjunto de reglas para alcanzar unos objetivos, la metodología de la LC se caracteriza por su capacidad de definir las reglas previamente a su aplicación, hecho que le permite contribuir a otros dominios, por ejemplo, la lexicografía, la enseñanza de lenguas, etc.. (Tognini-Bonelli, 2001). Otros autores en cambio (Mc Enery, Xiao & Tono, 2006), sin negar que la LC sea un dominio de investigación con un estatus teórico propio, sostienen que es una metodología:

While we agree that corpus linguistics is ‘really a domain of research’ and ‘has a new enterprise and a new philosophical approach to linguistic enquiry’, we maintain that corpus linguistics is indeed a methodology rather than an independent branch of linguistics in the same sense as phonetics, syntax, semantics or pragmatics.

(Mc Enery, Xiao & Tono, 2006, p.7)

⁴Para una explicación más detallada véase (Mc Enery, Xiao & Tono, 2006). Para una discusión y crítica detallada véase McEnery & Gabrielatos (2008), y Stubbs (1996) para los principios que rigen el análisis de textos en la tradición británica desde Firth y Halliday hasta Sinclair.

Esta es la perspectiva que adoptamos nosotros, es decir, consideramos la LC como una metodología con amplias posibilidades de aplicación a otras áreas y teorías de la lingüística (Mc Enery, Xiao & Tono, 2006).

5.2.2 El enfoque *corpus-driven* versus *corpus-based*

El hecho de que algunos autores consideren la LC como teoría o como metodología suele coincidir en cómo utilizan el corpus para su estudio (Gries, 2010). Un corpus es, en un sentido general, un conjunto de textos escritos u orales guardados en un formato electrónico (Conrad, 2002). Hay dos extremos en el continuum, el enfoque *corpus-driven* y el *corpus-based*. A continuación exponemos sucintamente algunas características diferenciales.

El *corpus-driven* es una forma de explorar los datos, básicamente inductiva, ya que se generaliza a partir de datos lingüísticos particulares. Esta metodología construye la teoría exclusivamente desde los datos, por eso se postula que la LC es una teoría y está muy asociada a los *neo-Firthians*. Los lingüistas que utilizan el enfoque corpus-driven postulan que no hay necesidad que el corpus sea representativo, más bien sostienen que a medida que aumenta el corpus este es representativo por acumulación. Es por eso que consideran que el tamaño del corpus ha de ser grande. Consideran que no se debe anotar el corpus, ya que analizan los datos sin teoría previa establecida. Una anotación impondría sus restricciones teóricas a los datos. Tampoco hace una distinción entre léxico, sintaxis, pragmática, semántica y discurso, ya que provee solamente un nivel de descripción del lenguaje, la unidad de sentido.

El *corpus-based* suele ser de carácter deductivo, ya que usa los datos para explorar una teoría o hipótesis y posteriormente validarla o refutarla según las evidencias. Desde esta perspectiva se sostiene que la LC es una metodología (McEnery & Hardie, 2011; Tognini-Bonelli, 2001).

Hemos expuesto los rasgos que diferencian a estos dos enfoques, obviamente en la práctica estos límites quedan desdibujados en el continuum. Ya que como nos recuerda Stubbs (1996) refiriéndose a los lingüistas del enfoque *corpus-driven*, "The linguist always approaches data with hypotheses and hunches, however vague"(p.47).

Por otra parte, Biber, Conrad, Reppen (1998) describen así las ventajas de adoptar una aproximación basada en corpus:

- Al analizar los patrones actuales del uso del lenguaje en textos naturales se genera evidencia empírica.
- Permite una adecuada y fiable representación del discurso, ya que utiliza una muestra amplia y representativa de textos naturales, *corpus*, como base de su análisis.
- Utiliza técnicas automáticas computacionales e interactivas para el análisis, que permiten la replicabilidad de los procesos.
- Depende tanto de análisis cuantitativos y cualitativos. Va más allá del simple conteo de características lingüísticas, es importante incluir una interpretación cualitativa funcional para explicar los datos.

A pesar de estas ventajas, el enfoque basado en corpus también ha suscitado críticas. Por ejemplo, Widdowson (1998) considera que los datos de un corpus no dan cuenta de los factores contextuales, ya que hay una separación entre texto y contexto comunicativo. Este mismo autor plantea la cuestión de cómo se pueden presentar los datos del corpus como ejemplos para los estudiantes, es decir, como herramienta pedagógica para adaptarlos a su entorno sociocultural más próximo. Desde la lingüística aplicada a la educación Flowerdew (2005, 2008) y Tribble (2002), en respuesta a estas críticas, proponen combinar la lingüística basada en corpus con una perspectiva teórica de *género*. Ésta considera las funciones sociales del género como un conocimiento esencial, que deberían adquirir los alumnos para desarrollar sus habilidades de escritura en una disciplina académica específica (Bazerman, 1988; Berkenkotter & Huckin, 1995; Miller, 1984). Una explicación detallada de las teorías de género la presentamos en el apartado 4.3.

Así, la *nueva retórica* tiene en cuenta la dimensión etnográfica en la interpretación de los datos (Flowerdew, 2005). La etnografía es el estudio del contexto sociocultural, percepciones y significados en un sistema cultural (Whitehead, 2005). Tal como refiere Flowerdew (2005), algunos estudios han combinado el estudio basado en corpus con una perspectiva etnográfica al utilizar informantes o al considerar la influencia de otros textos, la intertextualidad. Por ejemplo, Gledhill (1996) en artículos de investigación médica preguntó a especialistas qué textos eran para ellos los más representativos de la disciplina; Hyland (1998) consultó a especialistas sobre el uso de los mitigadores/atenuadores en biología. Milton (1998) comparó los exámenes escritos en inglés por estudiantes de Hong Kong con los que recibieron las máximas puntuaciones por los examinadores de Cambridge; McEnery & Kifle (2002) investigaron el uso de marcadores epistémicos en textos argumentativos escritos por estudiantes no nativos de inglés con los escritos por nativos ingleses. La diferencia en los resultados la atribuyeron a los libros de textos utilizados por los no nativos de inglés. Estos son algunos ejemplos en los que se ha demostrado la importancia de considerar los factores contextuales e intertextuales a la hora de interpretar los datos. Desde este punto de vista, Flowerdew (2005) sostiene que este es justamente el valor de trabajar con un corpus pequeño especializado, se considera pequeño un corpus de 250.000 palabras, donde el compilador es también el analista y conoce, por tanto, el contexto sociocultural en el que se han creado los textos.

Para finalizar este apartado queremos resaltar que nuestro trabajo ha adoptado esta perspectiva metodológica. A la par que una metodología basada en corpus sustentamos una aproximación desde la teoría de género, es decir, un enfoque etnográfico, en el que el contexto y la intertextualidad se tienen en cuenta a la hora de interpretar los resultados. La persona que suscribe esta tesis ha actuado como compilador y analista, aun así no conoce con exactitud el contexto sociocultural más próximo a los estudiantes como, por ejemplo, los antecedentes lingüísticos, si los estudiantes presentan dificultades lectoras o escritoras. Obviamente, esto es una limitación.

Como acabamos de explicar, hay divergencias en la manera de analizar los datos de un corpus por parte de los lingüistas, pero si en algo hay un consenso generalizado es en considerar la LC como una aproximación empírica a la descripción del uso del lenguaje, en tanto que se basa en datos lingüísticos auténticos. El corpus, definido como una colección de textos orales o escritos, auténticos y representativos de un lenguaje, compilados con un propósito específico (e.g., Sinclair, 1991; Biber et al., 1998; Hunston, 2002), permite una aproximación empírica a los datos, ya que son datos objetivos observados en el lenguaje tal y como se usa. Esta corriente empírica entra en contraposición con el tipo de estudio del lenguaje realizado tradicionalmente con carácter racionalista, es decir, basado en la introspección (Leech, 1991).

5.4 Definición de corpus

En el apartado anterior ya hemos apuntado una definición de corpus, no obstante a continuación nos detendremos en las características derivadas de las definiciones, ya que las diferentes definiciones resaltan determinados aspectos.

Así, por ejemplo, un corpus es “una colección de textos que se asumen representativos de una lengua o dialecto para ser usado en el análisis lingüístico” (Francis, 1992, p. 17, *apud* Tognini-Bonelli, 2001, p. 53). Leech (1992, p.116), figura destacada en el resurgir de la LC, comenta que “el corpus electrónico no es simplemente una colección hecha al azar de material textual, más bien, se ha compilado con un propósito determinado y se considera representativo de una lengua”. Sinclair (1991), así como Mc Enery & Wilson (2001), resaltan la importancia de la representatividad. Para Atkins, Clear & Osler, (1992) el corpus tiene que construirse con unos criterios de diseño explícitos. Estos también son importantes en la definición que da el grupo EAGLES (*Expert Advisory Group on Language Engineering Standards*) creado en los años 90 en la Unión Europea, con el objetivo de establecer normas de codificación y anotación para facilitar la consulta y la reutilización de los recursos lingüísticos.

Actualmente sus recomendaciones son aceptadas como estándares y definen el corpus como “una colección de piezas de lenguaje seleccionadas y ordenadas con criterios lingüísticos explícitos para ser usada como una muestra del lenguaje” (EAGLES, 1996).

En base a todas estas definiciones, en la lingüística moderna se considera (Flowerdew, 2012; McEnery & Wilson, 2001; Villayandre, 2008) que el corpus debe tener las siguientes características: representatividad, tamaño, formato electrónico, autenticidad de los datos y criterios de selección.

La representatividad significa que la muestra seleccionada debe representar la variedad de la lengua en estudio, es decir, debe ser una muestra representativa. En efecto, ya Chomsky (1965) criticó, con acierto, el posible sesgo de los datos en un corpus. Para solucionar este problema se deberían crear corpus, aunque a pequeña escala, el máximo de representativos de la variedad (e.g., autor determinado, período de tiempo, género, etc.), para que tomados en su conjunto nos den un promedio de toda la lengua.

El tamaño finito se refiere a que los corpus suelen ser finitos. Los corpus generales o de gran escala pueden tener hasta 100 millones de palabras, mientras que los corpus pequeños y más especializados tienen entre 50 mil a 250 mil palabras. También hay corpus abiertos, llamados monitor, como el COBUILD, por ejemplo. Algunos lingüistas como Sinclair (1991) han defendido la necesidad de compilar corpus grandes para que puedan reflejar la variedad de la lengua, no obstante, actualmente se da más importancia a los criterios de diseño. El tamaño apropiado dependerá del propósito de la investigación.

El formato electrónico tiene que ver con que los textos están en formato electrónico, lo que permite una búsqueda rápida, una recuperación de determinadas palabras en su contexto inmediato (concordancia), así como la computación y clasificación de los datos lingüísticos o enriquecimiento de los textos con información adicional (anotación).

La autenticidad de los datos significa que los textos han de ser muestras reales de uso de la lengua. Los criterios de selección exigen que los textos hayan sido escogidos con unos criterios previos lingüísticos y/o extralingüísticos según el objetivo planteado en la investigación.

5.5 La lingüística de corpus y el funcionalismo

En lingüística se suele hacer la dicotomía entre los modelos funcionalistas en oposición a los formalistas. Aún a riesgo de caer en una perspectiva simplista, esta oposición puede ser útil (Gonzálvez-García & Butler, 2006). El funcionalismo, en un sentido amplio, engloba las perspectivas cognitivas, tipológicas y *usage-based*, porque considera el lenguaje no como un sistema abstracto independiente del contexto, sino como un sistema que se usa para expresar sentido en el que interviene el contexto en el que ocurre y la cognición. Evidentemente hay rasgos bien diferentes en cada una de estas perspectivas, pero también comparten muchos aspectos. De hecho, Gonzálvez-García & Butler (2006) en su análisis extenso de las convergencias y divergencias de los modelos funcionalistas y cognitivistas concluyen que, en línea con Nuyts (2005) que no encontró diferencias absolutas, estas dos perspectivas se podrían englobar bajo una categoría superordinada llamada *functionally-oriented linguistics*. Otros autores como Culicover & Jackendoff (2005) también abogan por integrar ideas del formalismo, funcionalismo, cognitivismo y constructivismo; asimismo, Gries (2012) postula que la LC es compatible con la perspectiva cognitiva y psicolingüística y que se pueden beneficiar mutuamente.

El principio fundamental del funcionalismo es que el lenguaje es un medio de comunicación en su contexto social y cognitivo, por lo que no es independiente de factores socioculturales y cognitivos, sino que más bien está formado por ellos (Butler & Gonzálvez-García, 2014). Por lo tanto, podemos decir que, desde una perspectiva funcionalista, se analizan las formas lingüísticas pero explicadas en referencia a las funciones que desempeñan, considerando su contexto comunicativo y, en la mayoría de los casos, investigando las propiedades cognitivas que subyacen al uso del lenguaje. En cambio, bajo una perspectiva formalista las formas lingüísticas se estudian aisladas de su contexto (McEnery & Hardie, 2011).

En un sentido amplio del término, el *funcionalismo* también centra su atención en la flexibilidad del lenguaje derivada del contexto comunicativo; subraya la importancia de estudiar el lenguaje más allá de la frase, es decir, incluye, asimismo, los patrones discursivos; y estudia el uso basándose en datos reales (Butler & González-García, 2014). De hecho, la metodología basada en corpus es una herramienta valorada desde una perspectiva funcionalista, ya que como hemos señalado anteriormente, un corpus es un conjunto de textos, que permite contextualizar el análisis del lenguaje (Meyer, 2002). Concretamente, desde la *lingüística sistémico-funcional* Halliday (2006), *apud* Holtz (2011), resalta la afinidad entre la teoría sistémica y la LC: "Corpus studies underpin the general principle of functional variation in language; they make it possible to quantify the lexicogrammatical differences among different registers and to interpret this kind of variation as a redistribution of probabilities" (p. 294).

Como ya comentamos en el apartado 4.6.1, desde esta teoría se postula que la gramática y el léxico son los polos de un continuum, llamado léxico-gramatical, es decir, que gramática y vocabulario no están en estratos diferenciados (Halliday & Matthiessen, 1985). En cambio, desde el formalismo representado por Chomsky se subraya la autonomía de la sintaxis respecto de la semántica.

Desde la lingüística cognitiva, representada por lingüistas como, por ejemplo, Langacker, Lakoff, Talmy o Taylor, se tiene una perspectiva del lenguaje diametralmente opuesta a la de Chomsky, así el lenguaje no se considera como una facultad autónoma cognitiva, es decir, consideran el conocimiento lingüístico ligado a la cognición general y su adquisición depende de la exposición a instancias del lenguaje. Desde esta perspectiva se asume que las estructuras gramaticales no son autónomas de la semántica (Climent & Mateu, 2009; Croft & Cruse, 2004). De hecho, la semántica es un tema central de estudio en la lingüística cognitiva, pero no desde un punto de vista lógico ni formal, sino conceptual, fruto de la experiencia con el entorno (Climent & Mateu, 2009).

En su comparativa González-García & Butler (2006)⁵ muestran que los modelos que hacen mayor uso de datos lingüísticos auténticos, es decir, de corpus, lo hacen desde una perspectiva funcionalista: la *lingüística sistémico-funcional* (e.g., Halliday & Matthiessen), la gramática emergente (e.g., Givón, Bybee, Thomson) y desde una perspectiva cognitiva y/o constructivista: la gramática de construcciones (e.g., Fillmore, Goldberg). En general, se observa una tendencia cada vez mayor, tanto en la lingüística funcional como cognitiva, de utilizar datos naturales y auténticos para la formulación de las teorías del lenguaje.

Los argumentos aducidos en este apartado corroboran la suposición de que el análisis del discurso y la lingüística basada en corpus es la metodología más adecuada para caracterizar el artículo de investigación mediante las propiedades léxico-gramaticales y la función metadiscursiva interaccional desde un marco teórico funcionalista.

⁵Para una revisión detallada, véase Butler & González-García, (2014) y González-García & Butler (2006).

CAPÍTULO 6

6. Método

En este capítulo presentamos el diseño de los dos estudios de la investigación, cuyo objetivo principal es caracterizar el desarrollo del registro lingüístico y de la función metadiscursiva interaccional en los trabajos de investigación. En el apartado 6.1 se detallan las características del corpus. En el apartado 6.2 se presentan las razones para seleccionar los textos que conforman el corpus. Se describe, en el apartado 6.3, el procedimiento de recogida de textos. En el apartado 6.4 se explican las categorías de análisis y finalmente el procesamiento, anotación y herramientas de análisis (6.5).

6.1 Descripción del corpus

El corpus TARBUC (*Treballs Acadèmics de Recerca de Batxillerat i Universitat en Català*) consta de un total de 110 trabajos de investigación, 57 trabajos de investigación de bachillerato y 53 trabajos de final de máster universitario, con un total de 1.588.461 palabras. La tabla 6.1 muestra el número de textos por nivel académico y por disciplinas, año de realización de los trabajos, la longitud media de los textos por palabras y el número total de palabras. También se puede ver la descripción del sub-corpus de artículos publicados en inglés escritos por los ocho sujetos que hicieron el trabajo final de máster, en revistas especializadas.

Tabla 6.1 Descripción del corpus

	<u>Bachillerato</u>	<u>Universidad</u>	<u>Publicaciones</u>
Año	2003 - 2012	2006 - 2012	2009 - 2012
Número de textos	57	53	8
Longitud media de los textos por palabras	9.599,368	19.647,113	3.936,5
Número total de palabras	547.164	1.041.297	31.492

	<u>Biología</u>	<u>Historia</u>
Número de textos	54	56
Longitud media de los textos por palabras	7.790,926	20.852, 696
Número total de palabras	420.710	1.167.751

6.2 Razones para seleccionar los textos

Las razones para seleccionar los textos que conforman el corpus fueron dos: la primera, basada en las características comunes que tienen los trabajos de investigación de bachillerato y universitarios; y la segunda, en las características divergentes de las disciplinas de biología e historia.

Los trabajos de investigación de bachillerato y los trabajos de final de máster universitarios tienen en común que una de sus finalidades es la iniciación a la investigación. Como ya hemos señalado, en Cataluña, los estudiantes de la escuela secundaria, a los 17-18 años de edad, deben completar un trabajo de investigación como requisito académico de su plan de estudios. Por otra parte, en el nivel universitario, los estudiantes deben presentar un trabajo de final de máster cuando finalizan sus estudios de máster.

En ambas etapas, el trabajo, que es el resultado de una investigación personal guiada, debe ser presentado en forma escrita y defendido oralmente. Esto posibilita que los estudiantes no solo desarrollen las competencias investigadoras en general, sino también las habilidades expresivas, de forma parecida a lo que se hace en las comunidades de discurso científico.

Es decir, desde una etapa educativa temprana los estudiantes tienen que hacer frente a las habilidades de investigación y también al lenguaje especializado y la retórica de su disciplina.⁶

Otra característica que comparten estos trabajos es que en general, tienen la estructura propia de los artículos de investigación, esto es, introducción, método, resultados, y discusión o conclusión (IMRD). Como estos trabajos pueden ser precedentes e ilustran el continuum hacia una escritura de investigación profesional, decidimos incorporar también artículos profesionales de revistas de impacto científico. Estos nos servirían como modelos de escritura disciplinar y como extremo del continuum. Una vez recopilado el corpus del máster de inmunología y habiendo visto que abarcaban las fechas de 2006-2012, nos preguntamos si alguno de los alumnos de máster habría seguido su carrera profesional en esta disciplina. Efectivamente, ocho estudiantes habían publicado en revistas de impacto inglesas, después de haber finalizado sus estudios de máster. Vale la pena mencionar que, sobre todo en ciencias, las publicaciones de calidad se hacen en lengua inglesa. Nos pareció interesante incluir también esas publicaciones, ya que pueden ilustrar la transición de los mismos sujetos entre el mundo universitario y el profesional; o entre una lengua y otra. Este fue el motivo principal para crear el sub-corpus, con las publicaciones en inglés, para el estudio longitudinal, es decir, el estudio 2. En el anexo 2 se pueden consultar los títulos de estas publicaciones.

Como ya hemos señalado la segunda razón para seleccionar los textos se basa en sus características divergentes. La revisión bibliográfica nos indica que cada comunidad disciplinar tiene sus formas discursivas propias, maneras diferentes de utilizar los recursos lingüísticos para transmitir el conocimiento (e.g., Tolchinsky et al., 2013).

⁶La normativa que recoge la finalidad y objetivos de los trabajos de Bachillerato y de Universidad se puede consultar en http://xtec.gencat.cat/web/.content/alfresco/d/d/workspace/SpacesStore/0089/dc2144da-fbf0-4605-848c-89dcfdb37998/treball_recerca.pdf y http://www.giga.ub.edu/acad/gdoc/fitxers/pdf/normes_TFM.pdf

Si cada disciplina tiene su manera particular de comunicar, se espera que esas diferencias sean más evidentes entre las humanidades y las ciencias. Por este motivo optamos por las disciplinas de biología e historia. Estas disciplinas fueron elegidas porque representan diferentes metodologías o tradiciones de investigación (Conrad, 2001; Cortés, 2004). Los estudios de biología se basan en una metodología experimental cuantitativa, mientras que en historia se basan principalmente en una metodología cualitativa, etnográfica, que describe un fenómeno social y cultural determinado, y utiliza datos como las entrevistas, documentos, etc., para explicar los fenómenos (Estrada, 2013).

Sin embargo, el campo de la biología es muy amplio, así que nos decidimos para nuestro estudio por la inmunología, básicamente por dos motivos. Por una parte, es una de las áreas más influyentes de la biomedicina.⁷ Muchas enfermedades en humanos y animales son causadas por alteraciones del sistema inmune. Además sus avances, por ejemplo en vacunas, inciden directamente en la salud a nivel mundial. Por otra parte, debido a su importancia, esta rama de las ciencias biomédicas forma parte de la asignatura de biología en bachillerato.⁸ Respecto a la disciplina de historia, nos inclinamos por seleccionar los textos del Màster d'Estudis Històrics, ya que por su carácter general los trabajos de bachillerato tendrían una similitud en cuanto a temática y metodología utilizada.

6.3 Procedimiento de recogida de textos

Para el procedimiento de recogida de datos, en primer lugar, contactamos con el coordinador del Màster en Immunologia Avançada (UB-UAB) y el coordinador del Màster d'Estudis Històrics (UB), respectivamente. Después de pedir las autorizaciones (véase anexo 1) a los diferentes tutores pudimos compilar los textos que conforman el corpus de universidad.

⁷Para más información consultar: British Society of Immunology
<https://www.immunology.org/page.aspx?pid=1272>

⁸Desde el Institut de Ciències de l'Educació (ICE) se organizan cursos de inmunología para el profesorado de secundaria i bachillerato <http://www.ub.edu/ice/cursos/immunologia>.

Para obtener los textos de bachillerato contactamos con el instituto Àusias March de Barcelona, cuyos tutores nos proporcionaron algunos trabajos. Como no eran suficientes contactamos con el instituto Sant Just Desvern, del área metropolitana de Barcelona, y allí nos proporcionaron los textos que faltaban para completar el corpus. Finalmente, los textos de las revistas indexadas en inglés para el estudio longitudinal los obtuvimos a través de PubMed. Este es un motor de búsqueda de libre acceso en la base de datos de MEDLINE, donde se pueden consultar citas y artículos completos de biomedicina.

6. 4 Categorías de análisis

6. 4. 1 Características léxico-gramaticales

Para examinar las posibles diferencias en el registro académico de los trabajos de investigación analizamos los índices de una selección de **características léxico-gramaticales**: longitud del texto, legibilidad, diversidad y densidad léxica, y dos medidas de complejidad sintáctica. A continuación explicaremos con detalle cada una de ellas, pero antes debemos entender las implicaciones que tiene el término *léxico-gramatical*.

El término *léxico-gramatical* está en línea con la perspectiva de la *lingüística sistémico-funcional* que hemos adoptado en esta investigación. Como ya se ha explicado en el apartado 2.6.1 desde esta perspectiva se opera con el término léxico-gramatical, en vez de con un conjunto de módulos separados como la sintaxis, morfología y lexicón (Halliday, 2009). Así, el léxico y la gramática se entienden como extremos de un continuum y también como diferentes perspectivas que se adoptan para observar el mismo fenómeno, el léxico-gramatical (Halliday, 1985; Halliday & Matthiessen, 1985).

Por otro lado, la unidad de análisis para la observación de las características léxico-gramaticales, según Halliday (2005), es la cláusula. Esta es el centro de la acción en la gramática. La cláusula es "the place, or the locus, where fundamental choices of meaning were made" (Halliday, 2005, p. XV). Las cláusulas crean el texto, que es una entidad semántica, mientras que la cláusula es una entidad léxico-gramatical. Así, tal como plantea Halliday (2009) el problema de estudiar la complejidad gramatical radica en determinar el final de una cláusula compleja, sobretodo en la modalidad oral. No así en la modalidad escrita, donde la ortografía indica el final por medio del punto: "A clause complex is equated with a written sentence (which is the written surrogate of the clause complex to start with)" (Halliday 2009, p. 76). Por lo tanto, nos basamos en este argumento para decidir que la **unidad de análisis** de este estudio sería la oración. Una oración se refiere a la unidad ortográfica que se inserta entre puntos, signos de exclamación o de interrogación, y se inicia con mayúsculas. Aunque la mayoría de los estudios publicados anteriormente han utilizado la cláusula como una unidad de segmentación de texto, consideramos la oración como una unidad de análisis viable de la lengua escrita (Berman & Ravid, 2009), ya que tiene una relación sintáctica significativa y completa (Szmrecsanyi, 2004).

Por lo tanto, la **longitud del texto** se calculó mediante el número de oraciones por texto y el número de palabras por texto. Solo hemos tenido en cuenta el texto verbal que va desde la introducción a la conclusión para contar oraciones; por lo tanto, las notas al pie, cuadros, figuras, expresiones matemáticas no han sido consideradas. *Palabra* se define en este estudio como ortográfica, es decir, una secuencia escrita entre dos espacios en blanco (Jackson, 2002). Palabras que contienen un apóstrofe (') o un guión (-), como los artículos (e.g., *la*) que se unen a un sustantivo mediante un apóstrofe (e.g., *anàlisi*), *l'anàlisi* (el análisis) o palabras con pronombres clíticos, como *haver-hi* (hay), se separaron por un espacio antes de hacer el recuento de palabras.

La diversidad léxica se calculó por medio de la *subsample variety*, esto es, la media de número de *types* extraída del sub-muestreo aleatorio de una determinada longitud (Xanthos, 2014). Utilizamos este cálculo ya que se ha demostrado que otra medida habitualmente utilizada, como el ratio *type/token*, es muy sensible a la longitud del texto (McCarthy y Jarvis, 2010) y la variedad de sub-muestreo es un cálculo robusto para medir la diversidad léxica (Xanthos, 2014).

Para obtener **la densidad léxica** calculamos la frecuencia de palabras léxicas o de contenido, tales como, nombres, verbos, adjetivos y adverbios [N + V + Adj (+ Adv)] que aparecen en un texto como ratio del número total de oraciones de ese texto (e.g., Uccelli, Dobbs & Scott, 2013). En cuanto a los adverbios, los intensificadores, delimitadores, conectores y adverbios deícticos no se consideran palabras léxicas, mientras que aquellos con significado referencial sí. Como la herramienta de análisis utilizada no distingue entre los adverbios léxicos y gramaticales, optamos por calcular la densidad léxica con adverbios y sin ellos para ver si había alguna diferencia significativa en los resultados.

Para determinar la **complejidad sintáctica** utilizamos dos medidas. La primera, **complejidad sintáctica (1)**, se calculó como la media de número de palabras por oración. La longitud de la oración se considera un índice de complejidad sintáctica (e.g., Uccelli, Dobbs & Scott, 2013). Otra medida de la complejidad sintáctica, **complejidad sintáctica (2)**, se calculó utilizando el número de palabras antes del primer verbo finito (e.g., McNamara, Crossley & McCarthy, 2010).

A continuación presentamos unos ejemplos que dan cuenta de las medidas de densidad léxica y la primera medida de complejidad sintáctica, calculada como número de palabras por oración.

En estos dos ejemplos de ciencias, (1) de bachillerato y (2) de universidad, hay un cierto paralelismo en cuanto a estructura. No obstante, en cuanto a densidad léxica en el ejemplo (1), de bachillerato, vemos que antes de la cláusula de relativo hay dos nombres o sustantivos (NN) y un adjetivo (JJ) modificador del nombre. En cambio, en el ejemplo (2), de universidad, hay cinco nombres (NN) que amplían la información del nombre que los precede, es decir, hay una ampliación del grupo nominal y tres adjetivos modificadores antes de la cláusula de relativo.

1. La diabetis és una malaltia crònica en [la que el nivell de glucosa de la
 NN NN JJ IN RC
 sang es troba alterat].

Bachillerato: 19 palabras

2. L'asma és una malaltia respiratòria crònica, de caràcter al·lèrgic en la
 NN NN JJ JJ IN NN JJ IN
 majoria de casos, [que té una alta prevalença en la nostra societat Europea
 NN IN NN RC
 actual].

Universidad: 26 palabras

(Corpus Tarbuc)

Por lo tanto, vemos que en la oración de universidad hay una mayor densidad léxica. Asimismo, esto se traduce en una mayor complejidad sintáctica, ya que en el ejemplo (1) de bachillerato hay diecinueve palabras, mientras que en el ejemplo (2) de universidad hay 26 palabras.

Para ilustrar la segunda medida de complejidad sintáctica, número de palabras antes del primer verbo conjugado o finito, presentamos estos dos ejemplos de universidad (3) de biología e (4) de historia. En el ejemplo (3) de biología vemos que hay 24 palabras antes del primer verbo finito y en el ejemplo (4) de historia hay 25 palabras.

3. Per tant, després d'un determinat estímul, el comportament diferencial dels factors de transcripció STAT3 i STAT5 respecte a STAT1, AP-1, NF-κB i NFAT **altera** l'equilibri entre ells, fet que determina la resposta a futurs estímuls i la generació de tolerancia.

4. Per la seva banda a Valàquia i Moldàvia des de finals del segle XIII i al llarg de la primera meitat del segle XIV es **van formar** dos nuclis incipients de poder, que generaren tot un seguit d'estructures de control social que van permetre l'aparició i articulació dels dos principats.

(*Corpus Tarbuc*)

Todas estas características, pero especialmente la nominalización, hace que el registro académico sea complejo, lo que dificulta su **legibilidad** (Pinker, 2014; Ventola, 1996). Por este motivo también hemos calculado la legibilidad de los textos mediante una herramienta de análisis automático, que explicaremos con más detalle en el apartado 5.3.

6.4.2 Características metadiscursivas interaccionales

Para analizar las características metadiscursivas interaccionales utilizamos la taxonomía presentada por Hyland (2005), después de considerar las presentadas por Salager-Meyer (1994), Vartalla (2001) y Mur (2007). Mientras que las de los otros autores se centran básicamente en los intensificadores y atenuadores o mitigadores, la taxonomía de Hyland (2005) hace una clara distinción entre las dimensiones actitudinales (postura) y de relación con el lector (afiliación). Además se ha utilizado en numerosos estudios previos en el género de investigación, en diferentes disciplinas y lenguas. Así que consideramos útil esta clasificación para explorar cómo se realizan las interacciones y persuasiones en nuestros textos académicos, ya que tal como señalan Snow & Uccelli (2009) esta taxonomía es la más adecuada para el estudio del lenguaje académico desde una perspectiva de desarrollo. Por consiguiente, siguiendo la clasificación de Hyland (2005), analizamos los textos para determinar las funciones de interacción, es decir, **la postura** (*stance*) y **la afiliación** (*engagement*) que habitualmente se realizan por medio de las siguientes categorías gramaticales: verbos léxicos en modo condicional, verbos modales, perífrasis verbales, adjetivos, adverbios y pronombres personales.

Como ya hemos señalado en el apartado 2.7, la **postura** refiere a una dimensión actitudinal. Es la voz textual del escritor y a través de ella se muestra como presenta y transmite sus juicios, opiniones y compromisos. Se realiza principalmente a través de cuatro recursos: los intensificadores, los atenuadores, los marcadores de actitud y los marcadores de auto-mención. Los *intensificadores* señalan la certeza del escritor y su participación en una propuesta (e.g., *seguro, demuestran, sin duda*). Los *atenuadores o mitigadores* indican la renuncia del escritor para expresar un compromiso total con una propuesta (e.g., *es posible, podría, tal vez*).

Los marcadores de *actitud* indican la actitud afectiva del escritor a una proposición, mediante la transmisión de sorpresa, importancia, etc. (e.g., *importante, de manera sorprendente, curiosamente*). Los marcadores de *auto-mención* hacen referencia explícitamente al escritor, a la identidad del autor, por medio de pronombres personales y artículos posesivos (e.g., *nosotros, nuestro*).

A continuación presentamos ejemplos extraídos del corpus TARBUC para ilustrar algunos marcadores de postura, que hemos clasificado en: (5) atenuadores, (6) marcadores de auto-mención e intensificadores y (7) actitud.

5. **Sembla** que activin STAT1 ...
6. Cosa que **ens porta a afirmar** que ...
7. És **important** destacar que si l'anàlisi ...

La **afiliación** se refiere a una dimensión de asociación con el lector, donde el escritor incluye a los lectores como participantes del discurso anticipando sus posibles objeciones y guiándolos hacia las interpretaciones del contenido. La afiliación se realiza mediante los siguientes recursos: los pronombres del lector, las directivas y los conocimientos compartidos. Los *pronombres del lector* son marcadores que indican la inclusión de los lectores como miembros de la disciplina, el escritor intenta guiarlos a través de un argumento como si estuvieran participando en el discurso y tuvieran objetivos compartidos (e.g., *nosotros inclusivo*).

Las *directivas* son marcadores que instruyen al lector a mirar a otra parte del texto (e.g., véase la figura 1), la forma de llevar a cabo algún tipo de acción (e.g., abrir la válvula) o a interpretar un argumento (e.g., se ha de pensar). Los *conocimientos compartidos* son marcadores que dan a los lectores un papel activo en la construcción del argumento, es decir, les permite reconocer el conocimiento como familiar (e.g., es bien conocido, obviamente).

Los siguientes ejemplos representan algunos marcadores de afiliación: (8) pronombres del lector, (9) directivas, y (10) apelaciones al conocimiento compartido.

8. La ressenya científica **ens** indica que la resposta...

9. En la figura 6 **veiem** els nivells...

10. **Se sap que** malgrat...

Creemos que el uso de esta clasificación es una forma viable de análisis de corpus, ya que permite la comparación de la frecuencia de uso de marcadores en diferentes idiomas, como el catalán y el inglés. Aunque las formas lingüísticas son diferentes en cada idioma, la función interaccional que cumplen es básicamente la misma. En la página siguiente mostramos un cuadro resumen (6.4.2.1) de los diferentes marcadores, su función, su forma y algunos ejemplos extraídos del corpus.

6.4.2.1. Cuadro resumen de las características metadiscursivas interaccionales

Características metadiscursivas interaccionales

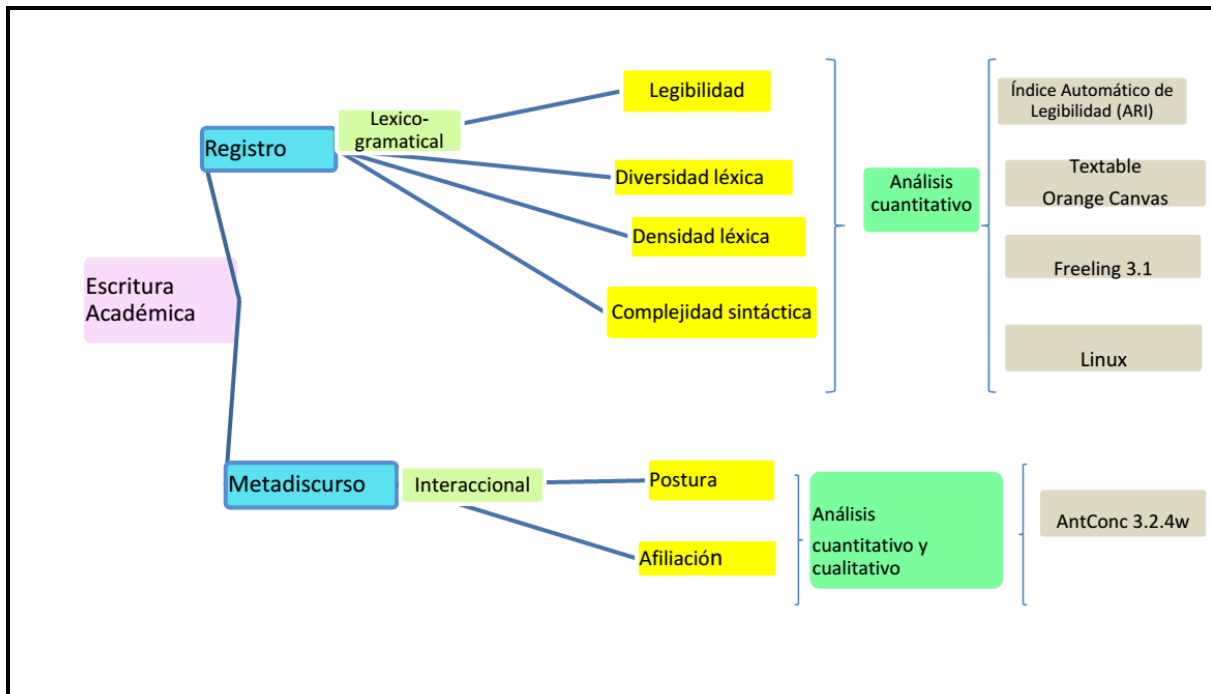
Función		Forma	Ejemplos
Postura			
Intensificadores	El escritor expresa certeza respecto a la afirmación y demuestra así el grado de involucración con el tema. En cierta manera restringe la alternancia de dialogo por utilizar un modo más taxativo.	La forma suele ser: verbos que indican dirección (suele ser auxiliar + infinitivo afirmativo) o introducido por adverbios de negación o adverbios de manera (como+vaux+inf)	demonstrat, segurament degut, és evident, es dedueix, confirmar, constancia cosa que ens porta a afirmar, per la qual cosa no podem dir, no hem obtingut un perfil immunològic suficientment clar, com era d'esperar, les deduccions ens condueixen a pensar ,etc.
Atenuadores	Los atenuadores se caracterizan en su función en que el escritor no se compromete del todo con la proposición. Expresa su distanciamiento con la proposición por medio de un razonamiento plausible más que por la certeza. Así se evita ser demasiado categorial.	La forma se caracteriza por el uso de verbos semánticos de apariencia, de probabilidad o de adverbios que indican duda o desconocimiento (potser, qui sap...), modo condicional.	Podria ser, semblen indicar, pot ser degut, permetria, podrien, indicaria, correspondria, pot ser atribuible, etc.
Marcadores de actitud	El escritor expresa su estado afectivo para con el tema. Se refleja su actitud de sorpresa, importancia, etc.	Suelen ser verbos afectivos, adverbios, adjetivos.	Important, interessant, rellevant, volem, agradaria, útil, sorprèn, etc.

	Función	Forma	Ejemplos
Afiliación			
Pronombres del lector	Incluye al lector por medio del “nosotros” de inclusión. Con la intención de hacer partícipe al lector de la disciplina en cuestión.	Se presenta con verbos conjugados en 1ª persona plural, verbos de percepción, pronombres personales.	Quan mirem, observem, podem veure, nostre, nosaltres, trobem, etc.
Directivas textuales	El escritor dirige al lector a otra parte del texto por medio de actos directivos: ver figura, tabla, etc.	Pueden ser imperativos o verbos de percepción.	S'observa, es descriu, es mostra, il·lustra, com ja s'ha dit en, en la figura, taula, etc.
Directivas cognitivas o al argumento	El escritor instruye al lector cómo interpretar o posicionarse respecto de un argumento. Hace notar al lector que tiene que fijarse p.ej. nótese, hay que considerar	Perifrasi verbal (caldre + inf, haver de + inf) Pensar, meditar, reflexionar algo con atención y cuidado	Cal considerar que, s'ha de pensar, cal destacar, cosa que indica, cal tenir en compte, etc.
Conocimiento compartido	El escritor se dirige al lector para incluirlo en la disciplina, por medio de referencias al conocimiento compartido. Para hacer notar que la proposición, concepto o teoría es conocida en el ámbito.	Utiliza verbos que indican conocimiento: conèixer, saber, etc.	Es coneguda com, se sap que, es coneix, etc.

6.5 Procesamiento y herramientas de análisis

Este apartado describe el procesamiento, anotación del corpus y herramientas de análisis utilizadas para abordar esta investigación. El gráfico 6.1 permite tener una visión de conjunto de este apartado.

Gráfico 6.1 Visión de conjunto de las unidades y herramientas de análisis



A la izquierda de la imagen vemos que analizamos la escritura académica a través del registro y el metadiscursos, mediante unas dimensiones. A la derecha detallamos el tipo de análisis, cuantitativo y/o cualitativo, y las herramientas de análisis. En las líneas que siguen explicamos todo el proceso detalladamente.

6.5.1 Procesamiento

Tal como se puede ver en el gráfico 6.1, examinamos, por un lado, el registro a través de ciertas características léxico-gramaticales por medio de un análisis cuantitativo. Por otro lado, determinamos la función metadiscursiva interaccional a través de un análisis cuantitativo y cualitativo. Es decir, primero examinamos todo el corpus para elaborar una lista con los marcadores interaccionales que aparecían en los textos. Posteriormente determinamos, teniendo en cuenta el contexto, su función. Este proceso es imprescindible, ya que como se muestra en (11a) y (11b), la misma forma lingüística (*mostra*) puede funcionar como un término de referencia (11a) o como un marcador de interacción (11b), en que la forma lingüística tiene la función de dirigir al lector a mirar la figura, es un marcador directivo.

11a Els microlitros de **mostra** que fóssin necessaris per a tenir 200mg

11b Tal i com es **mostra** a la figura 3.

Por otra parte, la misma forma lingüística (*indica*) puede dar cuenta de dos diferentes funciones. En (12a) funciona como un marcador directivo para el lector, mientras que en (12b) funciona como un marcador intensificador, para señalar certeza del escritor.

12a Tal com s'**indica** en la taula.

12b **Fet que indica** que s'ha dut amb èxit el procés.

Para determinar la **fiabilidad** de las decisiones consultamos con dos evaluadores independientes, uno licenciado en biología y lingüística y otro licenciado en lingüística. Comprobaron el 30% de las formas lingüísticas y decidieron su función específica de acuerdo a ese contexto. La comparación entre nuestra codificación y la de los evaluadores revelaron una fiabilidad de 0.86 (Kappa), lo que indica un alto grado de acuerdo. Para resolver las dudas o desacuerdos discutimos los casos para llegar a un consenso.

Esto respecto al análisis manual o cualitativo. Para los análisis automáticos o cuantitativos se pueden ver las herramientas utilizadas en el lado derecho del gráfico 1. Para los análisis cuantitativos de las características léxico-gramaticales se utilizaron el Índice Automático de Legibilidad (ARI) para ver el grado de legibilidad. El programa *Texttable* de Orange Canvas se utilizó para determinar la diversidad léxica. El programa *Freeling* 3.1 y *Linux* para calcular la densidad y complejidad sintáctica.

Antes de efectuar la anotación de cualquier corpus, se ha de preparar para su posterior procesamiento informático. Digitalizamos todos los trabajos, ya que los de bachillerato, por ejemplo, no lo estaban. Posteriormente editamos los textos, es decir, eliminamos gráficos, figuras, notas al pie, y en el caso de los textos de historia las transcripciones literales, fragmentos de otros autores, documentos o entrevistas. Esto es una convención para presentar los datos en historia (Estrada, 2013) y consideramos que, como no las habían redactado los propios estudiantes, distorsionarían los resultados. Finalmente se crearon dos versiones diferentes del corpus: una versión de texto sin formato; y una segunda versión de texto sin formato enriquecida con información adicional, que en este caso fue etiquetada morfológicamente con el software de acceso libre *FreeLing* 3.1 (Padró & Stanilovsky, 2012). Con ello se obtiene un output como el que se muestra en el ejemplo 6.5.1.

```
aquelles aquell DD0FP0
situacions situació NCFP000
en en SPS00
que que PR0CN000
s' es P0300000
altera alterar VMIP3S0
la el DA0FS0
correcta correcte AQ0FS0
homeòstasi homeòstasi NCMS000
d' de SPS00
algun algun DI0MS0
òrgan òrgan NCMS000
degut deure VMP00SM
a a SPS00
una un DI0FS0
infecció infecció NCFS000
```

Ejemplo 6.5.1. Resultado de anotación del corpus. Etiquetado morfológico con FreeLing 3.1

Como se ilustra en el ejemplo, el resultado o output contiene tres columnas: la primera es la palabra o forma (e.g., *situacions*), la segunda es el lema (e.g., *situació*), y la última es la parte del discurso (*Part of Speech: POS*) que es la categoría gramatical en el contexto (nombre común femenino plural NCFP000). Este etiquetado está basado en el grupo EAGLES <http://nlp.lsi.upc.edu/freeling/doc/tagsets/tagset-ca.html>. Una vez obtenidas las categorías gramaticales se pueden aplicar otros comandos de Linux para realizar los análisis y cálculos. En el apartado 6.5.2 explicamos con detalle las herramientas de análisis que utilizamos para esta investigación.

La primera versión, sin formato, la utilizamos para analizar:

- a) la longitud del texto por medio de número de oraciones por texto
- b) la cantidad de palabras por texto
- c) el grado de legibilidad
- d) la primera medida de complejidad sintáctica, es decir, cantidad de palabras por oración
- e) la diversidad léxica
- f) los marcadores metadiscursivos interaccionales de la postura (*stance*) y la afiliación (*engagement*)

La segunda versión, la versión enriquecida morfológicamente, la usamos para analizar:

- g) la segunda medida de complejidad sintáctica, es decir, el número de palabras antes del verbo principal
- h) la densidad léxica con adverbios y sin adverbios

6.5.2 Herramientas de análisis

Todas las herramientas de análisis que hemos empleado son de código abierto; esto facilita la réplica de esta investigación. Algunos de los **datos léxico-gramaticales son los** siguientes: la longitud por medio de número de oraciones por texto (a), la cantidad de palabras por texto (b), la primera medida de complejidad sintáctica (d), la segunda medida de complejidad sintáctica (g) y la densidad léxica (h). Estos se analizaron mediante un *shell script* para el sistema operativo Linux. *Shell* es un intérprete y ejecutor de comandos del sistema. Un *script* es un conjunto de comandos que se almacenan en un archivo. El *Shell* puede leer este archivo y ejecutar los comandos como si se escribiesen en el teclado. Linux es un software libre y de código abierto. La ventaja de utilizar linux resulta evidente para el análisis lingüístico de un corpus, ya que, por un lado, permite diseñar el script, ad hoc, es decir, el investigador elabora el comando según el análisis de características lingüísticas que quiera hacer. Por otro, garantiza la fiabilidad y precisión de los resultados, ya que se obtienen de forma automática. El ejemplo 6.5.2 muestra un fragmento del *shell script* para la medida de complejidad sintáctica (d) y la densidad léxica (h) que se aplicó sobre la segunda versión de los textos, enriquecidos con el analizador morfológico FreeLing.

```
{
    if($3 ~ /^nc|^NC/ || $3 ~ /^np|^NP/ || $3 ~ /^aq|^AQ/
|| $3 ~ /^ao|^AO/)
    {
        num_cont++
    }
    if($3 ~ /^rg|^RG/)
    {
        num_cont_adv++
    }
    if(!verb_found)
    {
        if($3 ~ /^vm|^VM/)
        {
            verb_found = 1
            num_cont++
        }
        else
        {
            sentence[num_sentences]++
            if(verbose)
            {
                sentence_str = sentence_str " " $1
            }
        }
    }
}
```

Ejemplo 6.5.2. Fragmento del *shell script* para Linux que se diseñó para este estudio.

El *shell script* permite escribir comandos más cortos en la terminal, para este caso concreto bastaba con escribir lo siguiente: *awc nombre archivo.txt verb*, y esto genera un output como el que se ilustra a continuación.

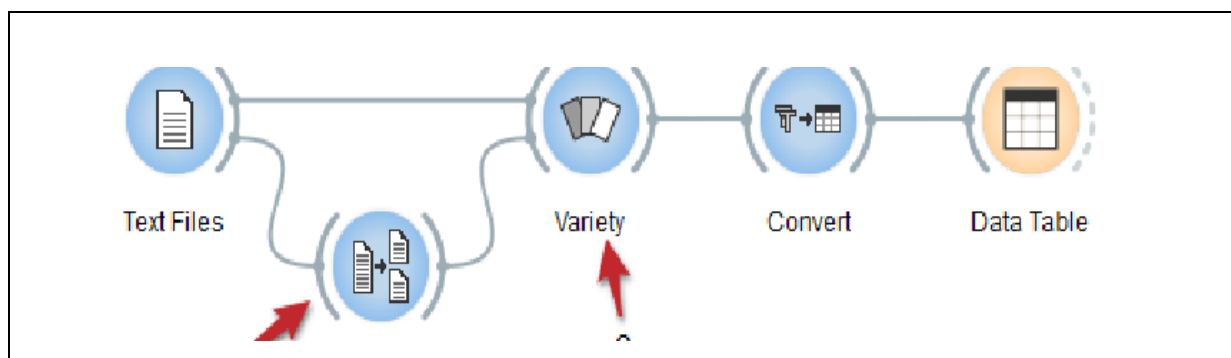
```
[File Report]

* Lines: 15300
* Sentences: 398
* Total WBV: 1906
* AVG WBV per Sentece : 4.78894
* MAX: 26
* MIN: 0
* Lexical Density: 15.8643 W/Sentence
* Lexical Density + Adverbs: 17.1432 W/Sentence
```

Ejemplo 6.5.3 Resultado después de aplicar el *shell script*.

En el ejemplo 6.5.3 se puede ver el output después de aplicar el *shell script*. El resultado muestra entre otros datos el total de palabras antes de un verbo principal (*total WBV*), la media por oración (*AVG WBV per sentence*), densidad léxica sin y con adverbios. Estos resultados se transcriben en la matriz de datos para poder realizar los análisis estadísticos posteriores.

Otros datos, como la diversidad léxica (e), fueron analizados con un software de libre acceso, *Textable* (Xanthos, 2014), que es una extensión de Orange Canvas (un entorno de programación visual para la minería de datos) que funciona en Windows y MacOSX. *Textable* proporciona toda la infraestructura para la apertura de los archivos, la segmentación en palabras, etc., por medio de una interfaz gráfica nada complicada, como se ilustra en el ejemplo 6.5.4.



Ejemplo 6.5.4 Interfaz del programa *Textable*, que se utilizó para calcular la diversidad léxica.

Textable permite evaluar la diversidad léxica utilizando la variedad media (es decir, el número medio de palabras distintas o *types*) en sub-muestras, de un tamaño dado, seleccionadas al azar en cada texto.

Los dos parámetros principales de este enfoque de la medición de la diversidad léxica son número de sub-muestras y tamaño de la sub-muestra. El número de sub-muestras es principalmente una cuestión de coste computacional: por lo general, cuanto mayor sea el número de sub-muestras, más largo será el tiempo de ejecución, y más precisa será la medición. Se suele poner un valor de 1000; a menos valor, el cálculo toma demasiado tiempo. El ajustar el tamaño de la sub-muestra es un proceso más delicado. Por un lado, no puede exceder la longitud (en número de palabras o *tokens*) del texto más pequeño en el conjunto de textos que se desea comparar, ya que de lo contrario no sería posible extraer una sub-muestra de este tamaño del texto más pequeño. Por otro lado, si el tamaño de la sub-muestra es demasiado pequeño, la medida será incapaz de discriminar adecuadamente los textos que se comparan. Por lo tanto, el tamaño de la sub-muestra es algo que se necesita ajustar empíricamente (Xanthos, 2014).

El grado de legibilidad (c) lo calculamos por medio de la herramienta de análisis *Índice Automático de Legibilidad* (ARI) disponible online en readability-score.com. En cuanto al grado de legibilidad (facilidad de lectura), todas las medidas tradicionales (e.g., Flesh Reading Ease y Flesh-Kincaid) basan sus cálculos en aspectos superficiales del texto como son aspectos semánticos (estructura de las palabras) y sintácticos (estructura de las oraciones) (Chall & Dale, 1995). Muchas de estas medidas tradicionales utilizan algoritmos que calculan el promedio de sílabas por palabra para determinar la dificultad de las palabras y la dificultad de las oraciones. Es decir, que solo darían resultados exactos para textos en lengua inglesa. El *Índice Automático de Legibilidad* (ARI) no se basa en sílabas sino en caracteres, por lo que debería producir un resultado razonablemente más preciso en otras lenguas. La fórmula que utiliza para generar el índice es la siguiente:

$$4.71 \left(\frac{\text{characters}}{\text{words}} \right) + 0.5 \left(\frac{\text{words}}{\text{sentences}} \right) - 21.43$$

El índice que genera es una puntuación numérica que indica el nivel o grado de legibilidad para una población dada (Collins-Thomson, 2014). Tal como se explica en readability-score.com para su interpretación se relaciona el resultado con el nivel de escolaridad del sistema americano. Es decir, si el resultado está en 12 es el nivel de lectura que habría de haber conseguido una persona al finalizar la educación secundaria (17-18 años). Un texto para ser leído por el público en general debería tener un grado de 8 (equivalente a 13-14 años).

Para analizar los **marcadores metadiscursivos interaccionales de la postura** (*stance*) y la **afiliación** (*engagement*) (f) se utilizó el programa de análisis de texto, de acceso abierto, *AntConc 3.2.4w*, utilizando un procedimiento de dos pasos. En primer lugar, se creó una lista cerrada de las formas lingüísticas que representan las funciones de postura y afiliación utilizando la función de *Word List* del programa, tal como se muestra en el ejemplo 6.5.5

Rank	Freq	Word	Lemma Word Form(s)
986	4	illuminades	
987	4	illuminats	
988	4	im	
989	4	imatge	
990	4	impregnats	
991	4	impressionista	
992	4	improvisacions	
993	4	inconscientment	
994	4	incògnites	
995	4	indica	
996	4	infinít	
997	4	influençiat	
998	4	inici	
999	4	iniciis	
1000	4	innovacions	
1001	4	intencions	

Ejemplo 6.5.5 Captura de imagen del programa *AntConc 3.2.4w* del apartado *Word list*

Como vemos en el ejemplo, mediante esta aplicación se pueden ver todas las palabras del texto y su frecuencia de aparición. Pero con ello solo tendríamos información sobre la forma lingüística y frecuencia, pero no sobre la función que cumple en su contexto inmediato. En segundo lugar, para determinar la función se llevó a cabo un análisis más cualitativo. Para ello, se utilizó la aplicación de concordancia para determinar el contexto de las formas lingüísticas y decidir su función específica de acuerdo a ese contexto. El ejemplo 6.5.6 muestra esta aplicación en la que podemos ver el contexto más inmediato. Si se pulsara sobre la forma marcada en azul, aparecería el contexto más amplio, es decir, la palabra en todo el texto.

Corpus Files	Concordance	Concordance Plot	File View	N-grams	Collocates	Word List	Keyword List
1.txt labs.txt	Hit	KWIC				File	
	1	rematisme el 1913 5. Com ja indica el t��tol de l'��pera, el t				1.txt	
	2	��lisi formal de l'obra ens indica que ��s una obra de formes				1.txt	
	3	rematisme el 1913 5. Com ja indica el t��tol de l'��pera, el tem				labs.txt	
	4	n��lisi formal de l'obra ens indica que ��s una obra de formes n				labs.txt	

Ejemplo 6.5.6 Captura de imagen de la aplicaci  n de concordancia.

Hasta aqu   hemos visto las herramientas de an  lisis de las categor  as l  xico-gramaticales y metadiscursivas interaccionales. Pasemos, ahora, a la presentaci  n de los resultados.

CAPÍTULO 7

7. Resultados del estudio 1

En este apartado presentamos los resultados del estudio 1, en el que comparamos transversalmente un corpus de trabajos de investigación de bachillerato y de final de máster universitario en las disciplinas de biología e historia, escritos en catalán. En el apartado 7.1, mostramos los resultados para el primer objetivo, es decir, estudiar el registro académico en los trabajos de investigación, por medio de un análisis cuantitativo de las características léxico-gramaticales. En el apartado 7.2 presentamos los resultados para el segundo objetivo que nos habíamos propuesto. Esto es, analizar cuantitativamente y cualitativamente los marcadores que tienen una función metadiscursiva interaccional en los trabajos de investigación.

7.1 Resultados del registro académico

Uno de los objetivos que nos habíamos planteado era analizar el registro académico en los trabajos de investigación. Por lo tanto, examinar si había variación del registro según el nivel educativo y según la disciplina. Para ello llevamos a cabo un análisis cuantitativo de las siguientes características léxico-gramaticales: (i) la diversidad léxica, (ii) la densidad léxica (iii) la complejidad sintáctica y (iv) la legibilidad de los textos.

Los análisis preliminares de los datos, mediante el programa *Statistics Package for Social Sciences* (SPSS), revelaron que muchos de los supuestos para pruebas paramétricas no se cumplían, como la normalidad de la distribución y la homogeneidad de la varianza. Debido a esto, aplicamos una prueba no-paramétrica, Mann-Whitney U, para evaluar el efecto del nivel académico y de disciplina en las características léxico-gramaticales.

En primer lugar, presentamos los resultados descriptivos de las características léxico-gramaticales según nivel y disciplina. En segundo lugar, mostramos los resultados inferenciales por nivel, entre bachillerato y universidad. Finalmente, presentamos los resultados inferenciales en las disciplinas de biología y en historia.

7.1.1 Resultados descriptivos de las características léxico-gramaticales

La tabla 7.1 ilustra los resultados descriptivos entre los niveles de bachillerato y universidad. Es decir, las medias, la desviación estándar, medianas y los rangos promedios. Vale la pena mencionar que en pruebas no paramétricas las medianas y los rangos promedios muestran la información de una forma más precisa, ya que las medias y la desviación estándar pueden no ajustarse a una interpretación tradicional, ya que los datos no están normalmente distribuidos (Field, 2000).

Los resultados descriptivos refieren a la longitud de los textos contabilizada en dos medidas: números de oraciones y en número de palabras. La diversidad léxica se refiere al número medio de palabras distintas o *types*. La complejidad sintáctica (1): calculada como el número de palabras por oración. La complejidad sintáctica (2) mide el número de palabras antes del primer verbo finito. La densidad léxica sin adverbios y con adverbios, calcula la frecuencia de palabras léxicas (sustantivos, adjetivos, verbos y con/sin adverbios). La legibilidad se ha de interpretar de la siguiente manera: cuanto más alto el valor mayor dificultad de legibilidad.

Tabla 7.1 Resultados descriptivos de los trabajos de investigación por nivel, bachillerato y universidad

Bachillerato (n=57)					Universidad (n=53)				
Características léxico-gramaticales									
	M	SD	Mdn	Rango promedio	M	SD	Mdn	Rango promedio	
Longitud (núm. de oraciones)	424.7	243.6	404	50.47	657.5	505.0	536	60.91	
Longitud (núm. de palabras)	9599.3	43211.0	9724	46.70	19647.1	14702.1	15375	64.96	
Diversidad léxica	75.7	1.4	75.9	61.04	75.2	1.4	75.1	49.54	
Complejidad sintáctica (1)	25.9	4.7	23.9	37.95	32.7	6.1	29.7	74.38	
Complejidad sintáctica (2)	4.4	.59	04.5	39.59	5.2	.68	05.1	72.61	
Densidad léxica con adverbios	13.3	2.5	13.50	36.04	16.8	2.5	16.7	76.43	
Densidad léxica sin adverbios	12.4	2.2	12.50	35.26	15.9	2.3	15.8	77.26	
Legibilidad	10.8	2.4	10.9	48.39	12.5	5.0	13.0	63.15	

Tabla7.2 Resultados descriptivos de los trabajos de investigación por disciplina, biología e historia

Biología (n=54)					Historia (n=56)			
Características léxico-gramaticales								
	M	SD	Mdn	Rango promedio	M	SD	Mdn	Rango promedio
Longitud (núm. de oraciones)	323.7	235.9	263.5	35.5	742.4	432.9	626	74.7
Longitud (núm. de palabras)	7790.9	3385.2	6863.5	35.5	20852.6	13318.3	15371	74.7
Diversidad léxica	76.1	1.2	76.3	69.4	74.8	1.3	74.8	42.0
Complejidad sintáctica (1)	28.7	6.3	25.8	55.9	29.6	6.5	25.8	55.1
Complejidad sintáctica (2)	4.8	.79	04.7	55.5	4.8	.70	04.7	55.4
Densidad léxica con adverbios	14.4	2.6	14.40	51.8	15.6	3.3	14.9	59.0
Densidad léxica sin adverbios	13.7	2.5	13.70	53.3	14.4	3.0	13.9	57.6
Legibilidad	9.3	3.0	09.9	36.7	13.8	3.5	13.1	73.6

7.1.2 Resultados inferenciales de las características léxico-gramaticales por nivel

Los trabajos de investigación de bachillerato y de universidad nos parecían diferentes en cuanto al registro, expresado a través de sus características léxico-gramaticales. Por eso, el estudio 1 tenía la finalidad de examinar si hay variación del registro según el nivel académico. Partimos de la asunción que hay diferencias en el registro en función de la edad y el nivel de alfabetización (e.g., Bar-Ilan & Berman, 2007; Ravid & Tolchinsky, 2002). Por lo tanto, esperamos encontrar diferencias entre niveles académicos. Mediante la estadística inferencial responderemos a la siguiente hipótesis.

H₀: Los trabajos de investigación de universidad no presentan un mayor nivel de desarrollo de las características léxico-gramaticales que los trabajos de bachillerato.

H₁: Los trabajos de investigación de universidad presentan un mayor nivel de desarrollo de las características léxico-gramaticales que los trabajos de bachillerato.

En la tabla 7.3 mostramos de forma resumida los resultados de las características léxico-gramaticales, según el nivel académico, obtenidos mediante la prueba Mann-Whitney U. Podemos ver los valores de las medias, la puntuación de Z y el tamaño del efecto. Para estimar el tamaño del efecto, hemos utilizado la siguiente ecuación $r = \frac{Z}{\sqrt{N}}$ en la que Z es la puntuación-z y N es el número de observaciones totales que se hicieron (Field, 2000).

Tabla 7.3 Resultados léxico-gramaticales según el nivel académico

	Bachillerato	Universidad		
	Media	Media	Z	r (tamaño del efecto)
Longitud (núm. de oraciones)	424.7	657.5	-1.71 ns	-.16
Longitud (núm. de palabras)	9599.3	19647.1	-3.00 **	-.29
Diversidad léxica	75.7	75.2	-1.89 ns	-.18
Complejidad sintáctica (1)	25.9	32.7	-5.9 **	-.23
Complejidad sintáctica (2)	4.4	5.2	-5.43**	-.52
Densidad léxica con adverbios	13.3	16.8	-6.63**	-.64
Densidad léxica sin	12.4	15.9	-6.90**	-.66

La prueba no-paramétrica, Mann-Whitney U, reveló que los trabajos de investigación de bachillerato y los de universidad no difieren significativamente en longitud, calculada en número de oraciones, $U = 1224$, ns , $r = -.16$. En cambio, los trabajos de universidad son significativamente más largos, en número de palabras, que los de bachillerato $U = 1009$, $p = .003$, $r = -.29$. En cuanto a la diversidad léxica, no hay diferencias significativas entre los dos niveles $U = 1194.5$, ns , $r = -.18$. Hay mayor complejidad sintáctica en los trabajos universitarios comparados con los de bachillerato. Esta diferencia es significativa tanto para la primera medida de complejidad sintáctica, es decir, la media de palabras por oración, $U = 510$, $p = .001$, $r = -.23$; como para la segunda medida, número de palabras antes del primer verbo finito, $U = 603.5$, $p = .001$, $r = -.52$. Asimismo, los trabajos universitarios presentan significativamente mayor densidad léxica que los trabajos de bachillerato. Densidad calculada con adverbios $U = 401$, $p = .001$, $r = -.64$ y sin adverbios $U = 357$, $p = .001$, $r = -.66$. En cuanto a la legibilidad, los trabajos universitarios son más difíciles de leer que los de bachillerato, $U = 1105$, $p = .015$, $r = -.23$.

Por otro lado, la prueba de correlación de Kendall Tau demostró que tanto en el nivel de bachillerato (ver tabla 7.4) como en el nivel de universidad (ver tabla 7.5), la legibilidad tiene una correlación alta con el número de palabras, la primera medida de complejidad sintáctica, esto es, la media de palabras por oración, y con la densidad léxica. La densidad léxica a su vez está correlacionada con la primera medida de complejidad sintáctica.

Tabla 7.4 Correlaciones de las características léxico-gramaticales en bachillerato

	1	2	3	4	5	6	7	8
1: Legibilidad	-							
2: Num. de oraciones	.069	-						
3: Num. palabras	.259**	.764**	-					
4: DivLex	-.114	-.061	-.013	-				
5: (SC1) Wps	.550**	-.012	.249*	.025	-			
6: (SC2) wbv	.065	-.136	-.055	-.027	.225*	-		
7: DL con adv.	.614**	-.051	.187*	.034	.808**	.223*	-	
8: DL sin adv	.585**	-.064	.170	.045	.792**	.239*	.957**	-

* $p < .05$. ** $p < .01$.

Tabla 7.5 Correlaciones de las características léxico-gramaticales en universidad

	1	2	3	4	5	6	7	8
1: Legibilidad	-							
2: Num. de oraciones	.420**	-						
3: Num. de palabras	.540**	.835**						
4: DivLex	-.16	-.146	-.174					
5: (SC1) Wps	.247**	-.125	.019	-.069				
6: (SC2) Wbv	.061	-.083	-.060	-.244*	.152			
7: DL con adv	.356**	-.018	.131	-.064	.586**	.099		
8: DL sin adv	.329**	-.058	.093	-.056	.633**	.153	.905**	-

* $p < .05$. ** $p < .01$.

Para determinar qué variables discriminan entre dos niveles, es decir, qué variables son las mejores predictoras para determinar que un estudiante pertenezca al nivel de bachillerato o universidad, realizamos un análisis discriminante. Este es similar al análisis de regresión logística y permite, asimismo, determinar el número mínimo de dimensiones necesarias para alcanzar la mejor clasificación posible. No obstante, como cualquier prueba estadística ha de cumplir unos supuestos previos, como por ejemplo que las matrices de covarianzas sean iguales (Torrado & Berlanga, 2007). En nuestro caso, y después de hacer la prueba de M. de Box, vimos que no se cumplía este supuesto para nuestros datos, por lo que optamos por el método *bootstrapping*, que es una técnica de remuestreo de datos. Permite el análisis aunque no se cumplan los supuestos de normalidad de la población. En el *bootstrapping* la muestra de datos es tratada como población en la que se escogen sub-muestras de forma repetida para el análisis e inferencias estadísticas (Field, 2000).

En la tabla 7.6 mostramos los resultados del análisis discriminante mediante la técnica *bootstrapping*. Los resultados muestran que la densidad léxica es la variable que mejor clasifica a los sujetos en un nivel u otro, ya que su autovalor o *eigenvalue* es el más alto. Le siguen las dos medidas de complejidad sintáctica y las dos medidas de longitud del texto. Todas estas variables son significativas por lo que es el modelo que mejor explica la predicción de que un estudiante sea de bachillerato o de universidad.

Tabla 7.6 Resultados del análisis discriminante mediante *bootstrapping*

	Eigenvalue	Canonical correlation	Effect size r	Chi square	p value
Densidad Léxica	.57	.60	.77	$\chi^2 (1)= 48.9$	.000
Complejidad Sintáctica 1	.39	.53	.72	$\chi^2 (1)= 35.7$	.000
Complejidad Sintáctica 2	.36	.51	.71	$\chi^2 (1)= 33.7$	.000
Longitud: Número de palabras	.22	.42	.64	$\chi^2 (1)= 21.8$	.000
Longitud: Número de oraciones	.09	.28	.52	$\chi^2 (1)= 9.2$	.002

7.1.3 Resultados inferenciales de las características léxico-gramaticales según la disciplina

Otro de los objetivos planteados para el estudio 1, era examinar si había variación del registro, expresado por medio de las características léxico-gramaticales, según la disciplina. Partimos del supuesto que el registro varía según la disciplina (e.g., Biber & Conrad, 2001), este rasgo parece ser más evidente entre biología e historia (Halliday & Martin, 1993). En consecuencia, prevemos encontrar diferencias de registro entre las disciplinas de biología e historia. Por medio de la estadística inferencial responderemos a la siguiente hipótesis.

H₀: Las características léxico-gramaticales en los trabajos de investigación no son significativamente diferentes en las disciplinas de biología e historia.

H₁: Las características léxico-gramaticales en los trabajos de investigación son significativamente diferentes en las disciplinas de biología e historia.

En la tabla 7.7 mostramos de forma resumida los resultados de las características léxico-gramaticales, según las disciplinas, obtenidos mediante la prueba Mann-Whitney U. Podemos ver los valores de las medias, la puntuación de Z y el tamaño del efecto.

Tabla 7.7 Resultados de las características léxico-gramaticales según las disciplinas

	Biología	Historia		
	Media	Media	Z	r (tamaño del efecto)
Longitud (núm. de oraciones)	323.7	742.4	-6.43 **	-.62
Longitud (núm. de palabras)	7790.9	20852.6	-6.45**	-.62
Diversidad léxica	76.1	74.8	-4.51 **	-.43
Complejidad sintáctica (1)	28.7	29.6	-0.13 ns	-.01
Complejidad sintáctica (2)	4.8	4.8	-0.01 ns	-.01
Densidad léxica con adverbios	14.4	15.6	-1.17 ns	-.11
Densidad léxica sin adverbios	13.7	14.4	-0.70 ns	-.07
Legibilidad	9.3	13.8	-6.07**	-.58

La prueba no-paramétrica, Mann-Whitney U, reveló que los trabajos de investigación de historia son más largos que los de biología. Difieren significativamente en longitud, calculada en número de oraciones, $U = 436$, $p = .001$, $r = -.62$ y en número de palabras $U = 432$, $p = .001$, $r = -.62$. En cambio, los trabajos de biología tienen más diversidad léxica que los de historia, esta diferencia es significativa $U = 758$, $p = .001$, $r = -.43$. No hay diferencias significativas entre las dos disciplinas ni en la primera medida de complejidad sintáctica, $U = 1490$, ns , $r = -.01$; ni en la segunda medida, $U = 1509$, ns , $r = -.01$.

Tampoco presentan diferencias significativas en cuanto a densidad léxica, con adverbios $U = 1315$, ns , $r = -.11$, y sin adverbios, $U = 1394$, ns , $r = -.07$. En cuanto a la legibilidad, los trabajos de historia son más difíciles de leer que los de biología, $U = 497$, $p = .001$, $r = -.58$.

7.2 Resultados de los marcadores metadiscursivos interaccionales

Otro de los objetivos que nos habíamos planteado era analizar los recursos retóricos que aparecen en los trabajos de investigación de los estudiantes. En concreto, los que tienen una función metadiscursiva interaccional, por medio de: (i) los marcadores lingüísticos que señalan la voz textual del escritor para involucrarse con su texto (*postura*), y (ii) los que utiliza el escritor para involucrar al lector (*afiliación*). Por lo tanto, se trata de identificar los marcadores metadiscursivos interaccionales en los trabajos de investigación, su distribución por nivel académico y por disciplina. Para ello llevamos a cabo un análisis cuantitativo y cualitativo de los siguientes marcadores interaccionales. Respecto a la *postura* analizamos los intensificadores, los atenuadores, los marcadores de actitud y marcadores de auto-mención. De los marcadores que refieren a la *afiliación* analizamos los pronombres del lector, las directivas, hacia el texto y hacia el argumento; y los conocimientos compartidos. En primer lugar, presentamos los resultados descriptivos. En segundo lugar, los resultados inferenciales por nivel y disciplina.

7.2.1 Resultados descriptivos de los marcadores interaccionales

La tabla 7.8 ilustra los resultados descriptivos entre los niveles de bachillerato y universidad y la tabla 7.9 muestra los resultados descriptivos según las disciplinas, biología e historia. Es decir, las medias, la desviación estándar, medianas y los rangos promedios de los marcadores calculados como porcentaje, esto es, frecuencia de aparición en un texto multiplicado por cien dividido por el número total de palabras de cada texto.

Tabla 7.8 Resultados descriptivos de los marcadores interaccionales según los niveles

Bachillerato (<i>n</i> = 57)					Universidad (<i>n</i> = 53)			
<i>Postura</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>Rango promedio</i>	<i>Mean</i>	<i>Mdn</i>	<i>SD</i>	<i>Rango promedio</i>
Intensificadores	0.07	0.06	0.06	53.82	0.09	0.05	0.09	57.30
Atenuadores	0.05	0.03	0.05	49.29	0.08	0.07	0.08	62.18
Actitud	0.18	0.06	0.78	51.80	0.08	0.07	0.05	59.48
Auto- mención	0.30	0.14	0.52	60.60	0.13	0.07	0.13	50.02
<i>Afiliación</i>								
Pronombres lector	0.08	0.04	0.07	59.26	0.08	0.04	0.13	51.45
Directivas al texto	0.05	0.02	0.07	55.52	0.06	0.02	0.07	55.48
Directivas al argumento	0.007	0.00	0.01	49.55	0.16	0.01	1.09	61.90
Conocimiento compartido	0.03	0.02	0.04	64.28	0.01	0.01	0.02	46.06

Tabla 7.9 Resultados descriptivos de los marcadores interaccionales según las disciplinas

Biología (<i>n</i> = 54)					Historia (<i>n</i> = 56)			
<i>Postura</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>Rango promedio</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>Rango promedio</i>
Intensificadores	0.12	0.09	0.09	70.72	0.05	0.03	0.05	40.82
Atenuadores	0.09	0.08	0.07	65.24	0.05	0.03	0.05	46.11
Actitud	0.08	0.06	0.66	54.32	0.18	0.07	0.79	56.63
Auto- mención	0.36	0.23	0.50	74.78	0.08	0.04	0.13	36.91
<i>Afiliación</i>								
Pronombres lector	0.08	0.03	0.13	50.95	0.08	0.05	0.07	59.88
Directivas al texto	0.10	0.09	0.08	73.88	0.01	0.08	0.03	37.78
Directivas al argumento	0.01	0.01	0.03	64.81	0.14	0.00	1.06	46.53
Conocimiento compartido	0.03	0.02	0.03	61.81	0.02	0.01	0.04	49.41

7.2.2 Resultados inferenciales de los marcadores interaccionales por niveles

Para cumplir con el objetivo de analizar la frecuencia de aparición de los marcadores metadiscursivos interaccionales, es decir, su distribución por nivel académico, realizamos una prueba no-paramétrica, Mann-Whitney U.

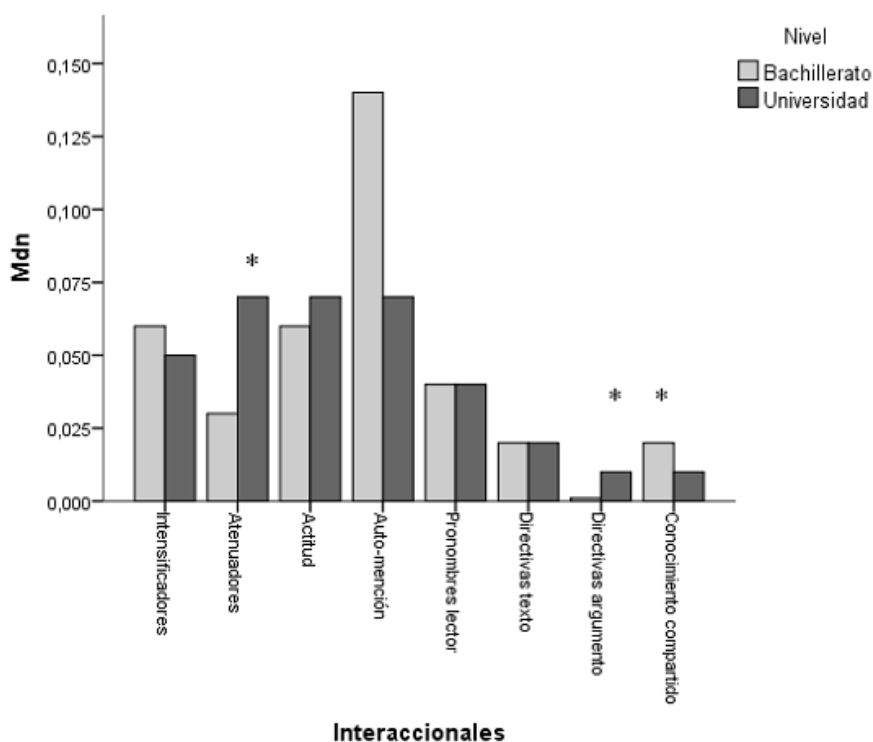
Suponemos que habrá un patrón de desarrollo en cuanto a la frecuencia de aparición de los marcadores interaccionales entre niveles educativos (Aull & Lancaster, 2014), en nuestro caso, entre bachillerato y universidad, ya que partimos del supuesto que hay un desarrollo en la expresión de la postura influenciado por factores como el cognitivo (Berman, 2007). Por consiguiente, mediante la estadística inferencial responderemos a la siguiente hipótesis.

H₀: Los trabajos de investigación universitarios no muestran un mayor nivel de desarrollo de los marcadores interaccionales que los trabajos de bachillerato.

H₁: Los trabajos de investigación universitarios muestran un mayor nivel de desarrollo de los marcadores interaccionales que los trabajos de bachillerato.

En el gráfico 7.1 se pueden ver los resultados inferenciales de los marcadores interaccionales por niveles

7.1 Resultados inferenciales de los marcadores interaccionales por niveles



Tal como muestra el gráfico 7.1, en relación a los marcadores de *postura*, la prueba no-paramétrica, Mann-Whitney U, reveló que los trabajos de investigación de universidad presentan una mayor frecuencia de aparición de intensificadores que en bachillerato, no obstante esta diferencia no es significativa, $U = 1415$, ns , $r = -0.05$. En cambio, en los trabajos universitarios hay un mayor porcentaje de atenuadores y esta diferencia es significativa, $U = 1156$, $p = .03$, $r = -0.20$. Para las variables de actitud, $U = 1299$, ns , $r = -0.12$.y auto-mención, $U = 1220$, ns , $r = -0.16$., las diferencias no son significativas. Respecto a los marcadores de *afiliación*, los resultados muestran un mayor porcentaje de marcadores directivos al argumento en universidad comparado con bachillerato y esta diferencia es significativa, $U = 1171$, $p = .03$, $r = -0.26$.

En cambio, en bachillerato hay un mayor porcentaje de marcadores de conocimiento compartido y esta diferencia es significativa, $U = 1110$, $p = .003$, $r = -0.29$. Para los otros marcadores, pronombres del lector $U = 1296$, ns , $r = -0.12$ y directivas al texto $U = 1509$, ns , $r = -0.0$, no hay diferencias significativas estadísticas. La prueba de correlación de Kendall Tau reveló que en el nivel de universidad los atenuadores están correlacionados con los marcadores directivos al argumento $\tau = .34$, $p < .001$.

7.2.3 Resultados inferenciales de los marcadores interaccionales por disciplinas

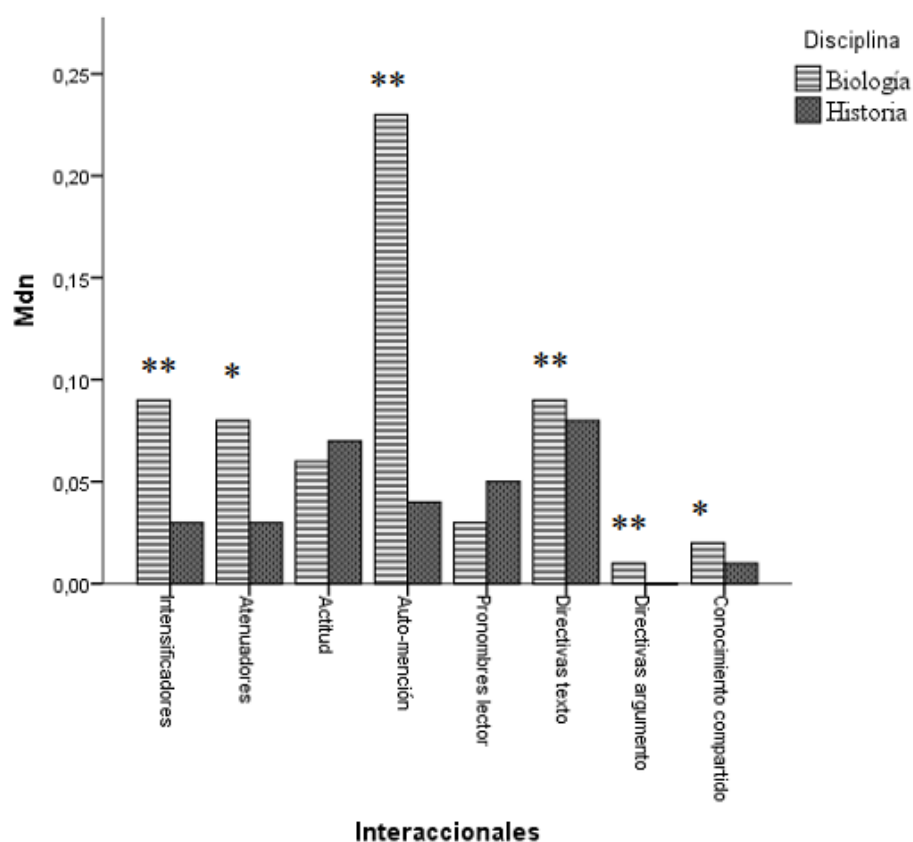
Para cumplir con el objetivo de examinar los marcadores metadiscursivos interaccionales entre las dos disciplinas, biología e historia, y determinar si hay variaciones entre estas, pasamos la prueba estadística Mann-Whitney U. Dado que en cada disciplina el escritor utiliza los marcadores de la postura y la afiliación de forma diferente (Hyland, 2005), esperamos encontrar diferencias significativas entre biología e historia. Para comprobarlo formulamos las siguientes

Hipótesis:

H₀: La frecuencia de los marcadores interaccionales en los trabajos de investigación no es significativamente diferente entre biología e historia.

H₁: La frecuencia de los marcadores interaccionales en los trabajos de investigación es significativamente diferente entre biología e historia.

Gráfico 7.2 Resultados inferenciales de los marcadores interaccionales según disciplinas



Los resultados muestran (ver gráfico 7.2) que algunas diferencias entre biología e historia son significativas. Biología tiene un porcentaje mayor en intensificadores $U = 690$, $p = .001$, $r = -0.47$; en atenuadores $U = 986$, $p = .002$, $r = -0.30$; en auto-mención $U = 471$, $p = .001$, $r = -0.60$; en directivas a otras partes del texto $U = 519$, $p = .001$, $r = -0.57$; en directivas al argumento $U = 1009$, $p = .001$, $r = -0.30$; y al conocimiento compartido $U = 1171$, $p = .04$, $r = -0.19$. El resto de los marcadores, como los marcadores de actitud y los pronombres del lector, no muestran diferencias significativas.

CAPÍTULO 8

8. Resumen de los resultados del estudio 1

Antes de pasar a discutir los resultados obtenidos en el estudio 1, presentamos un resumen de los mismos, aunque solamente los que refieren a diferencias significativas. En el estudio 1, en primer lugar, nos habíamos planteado examinar si había variación del **registro** en los trabajos de investigación, analizada por medio de las características léxico-gramaticales, según el nivel académico y según la disciplina. Los resultados obtenidos para el nivel académico refutan la hipótesis cero (H_0). Así, adoptamos la hipótesis alternativa (H_1), es decir:

H_1 : Los trabajos de investigación de universidad presentan un mayor nivel de desarrollo de las características léxico-gramaticales que los trabajos de bachillerato.

Los resultados muestran que los trabajos de universidad tienen mayor longitud en número de palabras que los de bachillerato. No hay diferencias en cuanto a la diversidad léxica, pero sí en la complejidad sintáctica. Los trabajos de universidad son más complejos, según las dos medidas de complejidad sintáctica y son más densos, es decir, tienen mayor densidad léxica. Esto conlleva una mayor dificultad de legibilidad. Además, los resultados discriminantes muestran que la densidad léxica es la que mejor predice la clasificación en un nivel u otro.

En cambio, los resultados según la disciplina no refutan la hipótesis cero, por lo tanto sugerimos que:

H_0 : Las características léxico-gramaticales en los trabajos de investigación no son significativamente diferentes en las disciplinas de biología e historia.

Los resultados muestran que la longitud es mayor en la disciplina de historia y son más difíciles de leer. Por el contrario, hay mayor diversidad léxica en la disciplina de biología. No obstante, en cuanto a la complejidad sintáctica y la densidad léxica no hay diferencias significativas.

En segundo lugar, en el estudio 1 también examinamos el **metadiscurso**, mediante los marcadores interaccionales, según el nivel académico y la disciplina. Los resultados obtenidos para el nivel académico refutan la hipótesis cero (H_0). En consecuencia, adoptamos la hipótesis alternativa (H_1):

H_1 : Los trabajos de investigación universitarios muestran un mayor nivel de desarrollo de los marcadores interaccionales que los trabajos de bachillerato.

Los resultados indican que en los trabajos de investigación de universidad hay un mayor porcentaje de atenuadores y de directivas al argumento que en los de bachillerato. Además estas dos variables en el nivel universitario están muy correlacionadas. En cambio, los trabajos de bachillerato presentan más marcadores de conocimiento compartido que los de universidad.

Respecto al análisis según las disciplinas, adoptamos también la hipótesis alternativa. Esta es:

H_1 : La frecuencia de los marcadores interaccionales en los trabajos de investigación es significativamente diferente entre biología e historia.

Los resultados muestran que en biología hay un mayor porcentaje de intensificadores, atenuadores, marcadores de auto-mención, directivas al texto, al argumento y conocimiento compartido, comparado con la disciplina de historia.

CAPÍTULO 9

9. Discusión

En una primera aproximación, analizamos un producto de la escritura académica de estudiantes de bachillerato y final de máster universitario mediante las siguientes variables: la longitud, la diversidad léxica, la complejidad sintáctica y la densidad léxica para trazar el desarrollo del **registro académico** en los trabajos de investigación. Nuestros datos amplían el panorama de la reciente investigación orientada al desarrollo del registro académico (e.g., Strömqvist et al. 2002; Woerfel & Yilmaz, 2011), proporcionando, por primera vez para el catalán y en trabajos de investigación datos cuantificados. Como nota general, hemos resaltado que en nuestro estudio la mayoría de las características léxico-gramaticales han puesto de relieve cambios significativos que se producen en el desarrollo de la escritura de los trabajos de investigación a través de los niveles académicos, la escuela secundaria y la universidad.

Encontramos diferencias en la longitud de los textos, en universidad los textos son más largos en número de palabras. Este resultado puede ser explicado por la característica de los trabajos de la disciplina de historia, en que la media de palabras por texto es superior a biología. Creemos que esto responde a las diferentes metodologías de cada disciplina que influye en la forma de escribir.

Nuestro valor de diversidad léxica ($M = 75-76$) es similar a los valores reportados en trabajos previos que han calculado este valor con el algoritmo VocD, por ejemplo, en narraciones escritas en inglés (70-80) (Mc Namara, Crossley & Mc Carthy 2010; Uccelli, Dobbs & Scott, 2013) y textos expositivos de la escuela secundaria (80-90) (Berman & Verhoeven, 2002).

Una alta diversidad léxica es un rasgo distintivo de la escritura académica y es considerado como una característica de los escritores competentes (Malvern, Richards, Chipere & Duran, 2004; Snow & Uccelli, 2009; Uccelli, Dobbs & Scott, 2013). Al igual que en el estudio de Johanson (2008), en el que la diversidad léxica no difirió en textos narrativos y expositivos escrita por jóvenes de 17 años y adultos, nosotros tampoco hemos encontrado diferencias entre los niveles académicos. Por lo tanto, nuestros resultados podrían sugerir que no hay cambios de desarrollo en relación con la diversidad léxica entre la escuela secundaria y la universidad, aunque desde una perspectiva retórica contrastiva se ha demostrado que para las culturas orientadas al contenido, como el español y el catalán, la diversidad léxica es un signo de estilo formal en el registro académico (Cuenca, 2003), por lo que se esperaría un aumento de la diversidad léxica con un mayor desarrollo de la alfabetización académica.

Así, esta ausencia de diferencias entre los niveles también podría explicarse por el efecto de la utilización de un género específico, en este caso el género de investigación, donde la repetición de palabras es generalmente la norma, ya que la repetición léxica es una característica distintiva del género del artículo de investigación (Conrad, 1996; Mirzapour & Ahmadi, 2011). Cuando los científicos comunican sus hallazgos utilizan repetidamente las mismas palabras para asegurarse de que se están refiriendo a los mismos conceptos (Biber & Finegan, 1994). Esta explicación es plausible, ya que diversos autores (e.g., Berman & Verhoeven, 2002) han demostrado que los valores de diversidad léxica difieren a través de diferentes géneros. Decidir si la ausencia de diferencias en la diversidad léxica en nuestros datos se debe a las características del desarrollo o es un reflejo de la adopción por parte de los estudiantes de la convención del género de investigación requiere un mayor análisis.

Por otro lado, nuestro índice de complejidad sintáctica y densidad léxica muestra un patrón claro de desarrollo. Los textos de ambos niveles académicos (es decir, de bachillerato y universitarios) difieren en las dos medidas de complejidad sintáctica y la densidad léxica en la dirección esperada. A diferencia de otros estudios en los que la complejidad sintáctica se mantiene estable, por ejemplo, en el primer y cuarto año universitario en textos argumentativos escritos en español (Neff & Prieto, 1994), nosotros encontramos un claro aumento entre la escuela secundaria y el nivel universitario. Nuestros resultados concuerdan con los resultados de estudios anteriores que mostraron que la complejidad sintáctica aumenta en función de la edad y la alfabetización, sobre todo a partir de la escuela secundaria (Berman & Ravid, 2009). Además, otros autores, por ejemplo, Mc Namara et al. (2010) y Uccelli et al. (2013), encontraron estas características predictoras de la calidad de la escritura en ensayos.

La complejidad sintáctica es inseparable de la densidad de contenido (Berman & Ravid, 2009) y aumenta con el desarrollo (Christie & Derewianka, 2008; Halliday, 2004; Veel, 1997; Halliday & Martin, 1994). Nuestro resultado de la densidad léxica es consistente con estudios previos en diferentes idiomas (Baratta, 2010; Colombi, 2002; Cutillas, Tolchinsky, Rosado & Perera, 2014) que mostraron un aumento en la densidad léxica a medida que aumentaban las habilidades de escritura. Ya hemos comentado que la densidad léxica también es un rasgo típico de los artículos de investigación, por ejemplo, la biomédica (Dubois, 1982; Guillén Galve, 1998; León & Divasson, 2006). A pesar de esto, vemos que los resultados del estudio 2 (apartado 9.2), estudio longitudinal de trabajos de investigación universitarios en catalán y publicaciones en inglés, los trabajos universitarios muestran una densidad léxica significativamente mayor que las publicaciones. Esta diferencia podría deberse a diferencias de lenguas (Johansson, 2008) o al hecho de que los estudiantes relacionan los textos densos con un registro formal, y por lo tanto, aspiran a emular este efecto retórico (Baratta, 2010).

En definitiva, si tenemos en cuenta que los resultados de las características léxico-gramaticales como la complejidad sintáctica y la densidad léxica no varían entre disciplinas, podemos confirmar que estas son indicadoras del desarrollo del registro académico en estos trabajos de investigación.

Por otro lado, nuestros resultados muestran que las características léxico-gramaticales, la complejidad sintáctica y la densidad léxica, en los trabajos de investigación no difieren entre las disciplinas de biología e historia. Estos resultados se pueden interpretar a la luz de la *lingüística sistémico-funcional*, es decir, las posibles diferencias en un registro reflejarían las diferencias en la configuración de los parámetros del campo, tenor y modo del discurso (Biber, 1988; Halliday & Hasan, 1989; Holtz, 2011). Los indicadores, en nuestro caso la complejidad sintáctica y la densidad léxica, que reflejan los parámetros de campo, tenor y modo, no varían entre disciplinas. Por lo tanto, podemos inferir que no hay diferencias entre estos parámetros en las dos disciplinas. Además, nuestros resultados sugieren que los universitarios, tanto de biología e historia, han alcanzado el registro que se supone propio del nivel de máster universitario.

En una segunda aproximación, nos centramos en el **metadiscurso**, la distribución de los marcadores de la postura y de afiliación a través de los trabajos de investigación de **bachillerato y de universidad**. Estudios previos (e.g., Cortes, 2004) sugieren que la simple exposición, a través de la lectura académica, a los marcadores que cumplen una determinada función no da lugar a la adquisición de estas por parte de los estudiantes. Aun así, partimos de la suposición de que los estudiantes de nivel universitario tienen una mayor alfabetización académica que los de bachillerato, y por tanto, esto se vería reflejado en el nivel de desarrollo de los marcadores interaccionales.

Dos conclusiones principales surgen de la distribución de la postura y la afiliación a través de los niveles académicos: los trabajos de investigación universitarios contienen más atenuadores y más directivas al argumento que los de bachillerato, que, por el contrario, contienen más marcadores de conocimiento compartido.

En primer lugar, el uso de atenuadores se desarrolla como una función a medida que hay un progreso en la alfabetización académica. Los escritores más avanzados, en este caso, los estudiantes universitarios, utilizan más atenuadores que los estudiantes de secundaria. Sin embargo, nuestros resultados han demostrado que los estudiantes de secundaria también son capaces de utilizar atenuadores mediante el uso de los verbos modales con un sentido epistémico (e.g., *podría*, *debería*, etc.), que se relacionan con estados hipotéticos o la consideración de múltiples perspectivas. Nuestros resultados, por lo tanto, concuerdan con estudios previos sobre el desarrollo de la postura (Berman, 2007, Uccelli, Dobbs & Scott, 2013). Otras investigaciones sobre metadiscurso, en inglés, también ponen de manifiesto que, sea cual sea la disciplina, los escritores académicos que están más avanzados utilizan más atenuadores (e.g., Aull & Lancaster, 2014; Hyland & Milton, 1996). En consecuencia, nuestro hallazgo sugiere que esto también es extensible a la lengua catalana.

Por otra parte, los estudiantes universitarios utilizan significativamente más directivas al argumento o actos cognitivos, en términos de Hyland (2005), que son aquellos que guían a los lectores por una línea de razonamiento (e.g., *hecho que lleva a pensar*, *hay que considerar*), que los estudiantes de secundaria. Vale la pena mencionar que los atenuadores en los trabajos universitarios están altamente correlacionados con las directivas al argumento. Esto significa que a mayor uso de atenuadores, mayor uso de directivas al argumento. Veamos esto en un ejemplo en la que aparece una directiva cognitiva o al argumento (en negrita) y un atenuador (subrayado) en la misma oración.

Les mostres amb Ht no presenten gaires diferències entre la freqüència de cèl·lules INKT i nTregs reclutades dins el mateix teixit, **el que fa pensar** que la majoria de cèl·lules infiltrants podrien ser cèl·lules efectores [...]

Los estudiantes de bachillerato, en cambio, utilizan significativamente más marcadores al conocimiento compartido (e.g., *es bien conocido*, *se sabe que*, etc.), donde se apela a los lectores a reconocer algo como una aceptación general.

Se sabe que los escritores avanzados tienen más conciencia del lector, pero también que en una etapa profesional, tal como la etapa *knowledge-crafting*, el escritor tiene en cuenta la representación de él mismo, del texto y del lector al redactar un texto coherente y efectivo (Bereiter & Scardamalia, 1987; Kellogg, 2008). Se ha demostrado que desde una edad temprana existe la conciencia del lector, pero esto no es lo mismo que tener en cuenta cómo un lector imaginario podría interpretar el texto. Para ello, el escritor tiene que tener una representación completa del lector (Kellogg, 2008). Aunque se trata de una característica de la etapa profesional o *knowledge-crafting*, nuestros resultados muestran que los estudiantes universitarios son capaces de ello, ya que mediante el empleo de las directivas a los argumentos, sutilmente llevan al lector a una interpretación particular, de forma parecida a como lo hacen los profesionales.

A pesar de la precaución necesaria que hay que tener a la hora de interpretar los datos, sobre todo, debido al tamaño de la muestra ($N=110$), nuestros resultados, considerados en su conjunto, sugieren que no solo los factores lingüísticos, sino también los factores socio-cognitivos influyen en la configuración del desarrollo de un escritor académico competente (Berman, 2004). Según esta autora, la competencia implica tres tipos de conocimiento: el lingüístico, dominio de las características léxicas y gramaticales; el cognitivo, la habilidad para integrar las formas para cumplir con las funciones del discurso; y el social, la capacidad de adaptarse a diferentes contextos comunicativos. Desde este punto de vista, nuestros resultados muestran que estos trabajos universitarios reflejan la interacción entre la facultad cognitiva, es decir, la consideración de múltiples perspectivas y las convenciones culturales del contexto comunicativo, esto es, la sensibilidad a las necesidades discursivas. En resumen, los estudiantes universitarios movilizan su capacidad para elegir entre todas las opciones retóricas disponibles y poder guiar al lector a través de múltiples perspectivas hacia su propio argumento.

En segundo lugar, el marcador más utilizado es el de auto-mención. Nuestros resultados confirman que el género de investigación contiene un discurso altamente retórico, sustituyendo así cualquier percepción de que sea objetivo e impersonal (e.g., Bazerman, 1988; Hunston, 1994; Swales, 2004).

A través de la auto-mención los escritores proyectan una voz del autor que, cuando se utiliza de forma adecuada, es decir, como lo hace la comunidad específica, da forma a la proyección personal. Esto les permite construir una identidad, haciendo hincapié en su contribución individual, lo que les ayuda a construir su autoridad profesional y su credibilidad (Hyland, 2004a). El uso de los marcadores de auto-mención difiere según el género, la disciplina y la cultura.

Referente al género, por ejemplo, tenemos que tener en cuenta que los trabajos de los estudiantes, tales como los de investigación, se diferencian de los artículos de investigación en diversos aspectos: el público al que va dirigido, el propósito comunicativo, las habilidades requeridas y el conocimiento. Mientras que los trabajos de investigación tienen como objetivo demostrar que las habilidades requeridas han sido adquiridas, los artículos de investigación tienen como objetivo informar de nuevos experimentos y persuadir al lector, para que se dé cuenta de la validez y la importancia de los nuevos hallazgos (Gardner & Nesi, 2013). A pesar de estas diferencias y como ya hemos explicado en el apartado 1.2, hay algunas similitudes entre los dos géneros, es decir, la etapa final del proceso de investigación es un producto escrito en el que se comunican nuevos conocimientos.

En cuanto a las diferencias culturales, por ejemplo, estudios previos, que analizaban el uso de los pronombres personales en artículos de investigación en gestión empresarial publicados en inglés y en español, han demostrado que la frecuencia de auto-mención es mayor en artículos de investigación en inglés que en español. Estos resultados podrían explicarse en parte por los antecedentes culturales del autor (Mur, 2007). Otros estudios han reportado un exceso o uso indebido de los marcadores de auto-mención al comparar artículos de investigación de biología escritos por hablantes nativos y no nativos de inglés (Martínez, 2005).

Determinar si este es el caso en nuestro corpus, va más allá del alcance de este estudio. A pesar de esta pregunta sin respuesta, nuestros hallazgos son interesantes, ya que ponen de relieve el hecho de que los estudiantes de secundaria ya reconocen el estatus especial del

trabajo de investigación como un género con una voz del autor. Esto es especialmente significativo si tenemos en cuenta que en la escritura académica se suele caracterizar el posicionamiento discursivo como distanciado y objetivo (e.g., Berman & Ravid, 2009; Snow & Uccelli, 2009). Por ejemplo, el texto expositivo se caracteriza generalmente por una falta de implicación personal. Por lo tanto, nuestros resultados relativos a la utilización de marcadores de auto-mención también en las publicaciones, estudio 2, corroboran la importancia que tienen estos marcadores en los artículos de investigación, ya que como discutiremos en el estudio 2, hay una diferencia significativa entre la frecuencia con la que se utilizan en las publicaciones y en el nivel universitario.

Nuestra conclusión en cuanto al porcentaje más alto de los marcadores de postura en relación con los de afiliación debe interpretarse a la luz de la perspectiva del metadiscurso. Desde esta perspectiva se sostiene que un texto no solo se compone teniendo en consideración las necesidades de los lectores, sino que es una construcción conjunta entre el lector y el autor, mediante la función interaccional. Esto, a su vez, significa que los escritores animan a los lectores a participar en su diálogo, incitándolos a aceptar o rechazar los argumentos presentados.

El logro de conseguir este compromiso con el lector supone un reto en la mayoría de textos argumentativos y persuasivos, incluidos los trabajos de investigación, que se basan justamente en el uso de marcadores, tales como intensificadores, atenuadores y los marcadores de actitud para realizar esta función dialógica (Thompson, 2001). Lo que nuestro estudio demuestra es que los escritores, incluso en la escuela secundaria, comentan sobre el contenido de sus textos e invitan a los lectores a un diálogo, lo que revela que son capaces de desplegar las formas lingüísticas que son propias del género de investigación.

Respecto al uso de los **marcadores interaccionales por disciplinas**, biología e historia, los resultados muestran que la expresión de postura y afiliación es una característica de la escritura académica en todas las disciplinas (Hyland, 2005), incluso en los textos escritos por aprendices. Si bien Hyland (2005) encontró un mayor uso de atenuadores e intensificadores en las disciplinas discursivas más humanísticas como la filosofía o la sociología que en biología, Cortes (2004) reporta lo contrario. En su corpus los atenuadores

se dan sobre todo en biología, este resultado concuerda con el nuestro. El mayor uso de marcadores de auto-mención, de intensificadores y atenuadores en nuestro corpus de biología respecto al de historia puede deberse a que en general los textos de biología deben de ser más cortos, por requerimiento de las editoriales, y eso provoca que se deban utilizar expresiones concretas, como las que hemos analizado, para crear una impresión de autoridad y credibilidad. En cambio, en historia se permite una mayor verbosidad. Otra explicación plausible sería que los estudiantes de biología están familiarizados con el discurso de su disciplina, que suele ser muy estandarizado (Hyland, 2005). En historia la voz de autor se intercala con transcripciones literales de fragmentos o entrevistas (Estrada, 2013), y hemos de recordar que éstas las hemos eliminado antes de analizar los textos, para no sesgar los resultados. Por lo tanto, nuestros resultados respecto a las disciplinas han de interpretarse con ciertas precauciones. No obstante, cada disciplina tiene una forma particular de utilizar los marcadores interaccionales (Hyland, 2005). Consideramos que un análisis de otros géneros en estas dos disciplinas podría aportar más información al respecto.

CAPÍTULO 10

10. Resultados del estudio 2

En este apartado presentamos los resultados del estudio 2. En él, analizamos longitudinalmente un corpus de trabajos de investigación de final del máster universitario, en catalán, y de publicaciones en inglés en la disciplina de biología y escritos por los mismos sujetos. Igual que en el estudio 1 examinamos el registro y el metadiscurso. El registro se analiza por medio de un análisis cuantitativo de las siguientes características léxico-gramaticales: (i) la diversidad léxica, (ii) la densidad léxica (iii) la complejidad sintáctica. Respecto al metadiscurso analizamos los marcadores interaccionales. De los marcadores que señalan la *postura* examinamos los intensificadores, los atenuadores, los marcadores de actitud y marcadores de auto-mención. De los marcadores que refieren a la *afiliación* analizamos los pronombres del lector, las directivas, hacia el texto y hacia el argumento; y los conocimientos compartidos.

Los análisis preliminares de los datos mostraron que los supuestos para pruebas paramétricas, como la normalidad de la distribución y la homogeneidad de la varianza, no se cumplían. Debido a esto, aplicamos la prueba de los rangos con signo de Wilcoxon para estimar el efecto del nivel académico en las características léxico-gramaticales y en los marcadores de interacción. Para calcular el tamaño del efecto, utilizamos la siguiente ecuación: $r = Z / \sqrt{N}$, en la que Z es la puntuación Z y N es el número de observaciones totales (Field, 2009).

10.1. Resultados descriptivos de las características léxico-gramaticales y los marcadores metadiscursivos interaccionales

La tabla 10.1 ilustra los resultados descriptivos de las características léxico-gramaticales entre los niveles de universidad y publicaciones. Asimismo, la tabla 10.2 muestra los resultados descriptivos de los marcadores metadiscursivos interaccionales, es decir, las medias, la desviación estándar y las medianas.

Tabla 10.1 Resultados descriptivos de las características léxico-gramaticales

Universidad (<i>n</i> = 8)				Publicaciones (<i>n</i> = 8)		
<i>Características léxico-gramaticales</i>						
	<i>M</i>	<i>SD</i>	<i>Mdn</i>	<i>M</i>	<i>SD</i>	<i>Mdn</i>
Longitud (núm. de oraciones)	204.8	46.3	202	156.1	45.8	165
Longitud (núm. de palabras)	6083	895.0	6275	3936	1165	4219
Diversidad léxica	75.8	1.8	76	78.3	1.1	78
Complejidad sintáctica (1)	30.3	4.3	30.5	25.2	2.8	25.0
Complejidad sintáctica (2)	5.3	.51	5.0	10.7	1.8	10.5
Densidad léxica con adverbios	16.3	2.1	17.0	13.8	1.4	14.0
Densidad léxica sin adverbios	15.8	1.9	16.5	13.0	1.3	13.0
Longitud (núm. de oraciones)	204.8	46.3	202	156.1	45.8	165

Tabla 10.2 Resultados descriptivos de los marcadores interaccionales

	Universidad (<i>n</i> = 8)			Publicaciones (<i>n</i> = 8)		
<i>Postura</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>
Intensificadores	0.12	0.06	0.12	0.23	0.24	0.05
Atenuadores	0.06	0.07	0.05	0.06	0.05	0.07
Actitud	0.04	0.05	0.02	0.12	0.14	0.07
Auto- mención	0.19	0.21	0.11	0.36	0.39	0.16
<i>Afiliación</i>						
Pronombres lector	0.12	0.03	0.19	0.007	0.00	0.02
Directivas al texto	0.14	0.13	0.06	0.16	0.13	0.08
Directivas al argumento	0.01	0.007	0.01	0.02	0.01	0.03
Conocimiento compartido	0.007	0.00	0.01	0.005	0.00	0.01

10.2 Resultados inferenciales de las características léxico-gramaticales

Uno de los objetivos de este segundo estudio era examinar las características léxico-gramaticales entre los trabajos de investigación de los estudiantes universitarios y los artículos publicados en inglés por ellos, y determinar si hay variaciones entre estos trabajos. Esperamos encontrar diferencias entre trabajos y publicaciones, porque estamos comparando no solamente dos tipologías de lenguas diferentes, sino también las dos culturas de escritura, la orientada a la forma y la orientada al contenido (e.g., Cuenca, 2003).

Aunque los autores sean los mismos sujetos, no sabemos si transfieren su estilo de escritura, determinada culturalmente, cuando escriben en inglés. Por lo tanto, mediante la estadística inferencial responderemos a la siguiente hipótesis.

H₀: Las características léxico-gramaticales de los artículos publicados no son significativamente diferentes a los trabajos de investigación universitarios.

H₁: Las características léxico-gramaticales de los artículos publicados son significativamente diferentes a los trabajos de investigación universitarios.

En la tabla 10.3 vemos los resultados.

Tabla 10.3 Resultados características léxico-gramaticales entre universidad y publicaciones

	Universidad	Publicaciones		
	Media	Media	Z	r (tamaño del efecto)
Longitud (núm. de oraciones)	204.8	156.1	-1.68 ns	-.42
Longitud (núm. de palabras)	6083	3936	-2.38 **	-.58
Diversidad léxica	75.8	78.3	-2.41**	-.60
Complejidad sintáctica (1)	30.3	25.2	-2.38 **	-.60
Complejidad sintáctica (2)	5.3	10.7	-2.55**	-.64
Densidad léxica con adverbios	16.3	13.8	-2.38**	-.60
Densidad léxica sin adverbios	15.8	13.0	-2.55**	-.64

Los resultados de las características léxico-gramaticales muestran una diferencia significativa en cuanto a número de palabras. Los trabajos universitarios ($Mdn = 6275$) tienen más palabras que las publicaciones ($Mdn = 4219$) $Z = -2.38$, $p = 0.01$, $r = -.58$. En cambio, los trabajos universitarios ($Mdn = 202$) y las publicaciones ($Mdn = 165$) no difieren significativamente en el número de oraciones, $Z = -1.68$, ns , $r = 0.42$. Las publicaciones ($Mdn = 78.00$) muestran significativamente mayor diversidad léxica que los trabajos universitarios ($Mdn = 76.00$), $Z = -2.41$, $p = 0.01$, $r = -.60$. La complejidad sintáctica, calculada como media de número de palabras en una oración, de los trabajos universitarios son más complejos que las publicaciones ($Mdn = 30.5$ y $Mdn = 25.0$, respectivamente). Esta diferencia es estadísticamente significativa $Z = -2.38$, $p = 0.01$, $r = -.60$. Sin embargo, para la segunda medida de la complejidad sintáctica, número de palabras antes del primer verbo finito en una oración, observamos lo contrario: las publicaciones ($Mdn = 10.50$) son más complejas que los trabajos universitarios ($Mdn = 5.0$), $Z = -2.55$, $p = 0.01$, $r = -.64$. Los textos universitarios ($Mdn = 17.00$) muestran significativamente mayor densidad léxica, con adverbios, que las publicaciones ($Mdn = 14.00$), $Z = -2.38$, $p = 0.01$, $r = -.60$.

10.3 Resultados inferenciales de los marcadores metadiscursivos interaccionales

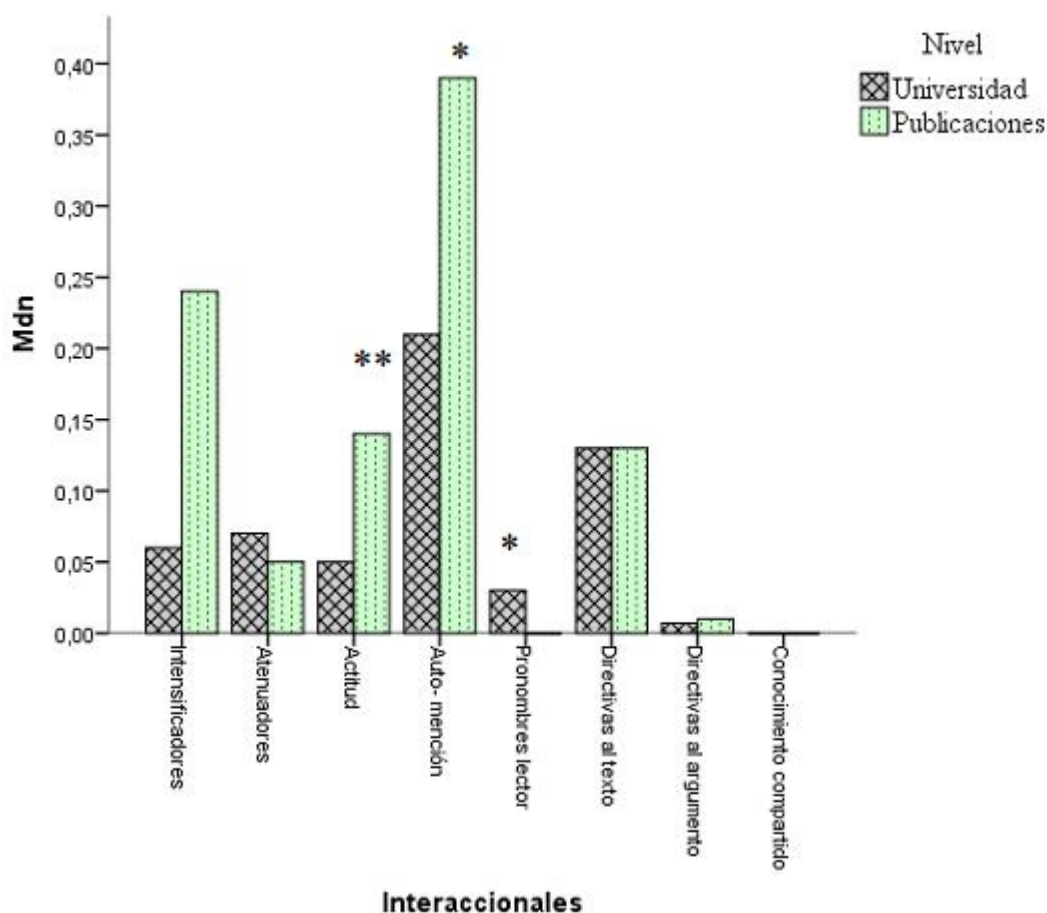
Otro de los objetivos de este segundo estudio era examinar los marcadores interaccionales de los trabajos de investigación universitaria y de las publicaciones. Suponemos encontrar diferencias ya que los géneros que analizamos son diferentes (Hyland, 2005). H_1 : La frecuencia de los marcadores interaccionales en los trabajos de investigación es significativamente diferente entre biología e historia. Para comprobarlo formulamos las siguientes hipótesis.

H_0 : La frecuencia de los marcadores interaccionales no es significativamente diferente entre los trabajos de investigación universitaria y las publicaciones.

H_1 : La frecuencia de los marcadores interaccionales es significativamente diferente entre los trabajos de investigación universitaria y las publicaciones.

En el gráfico 10.1 se ilustran los resultados de los marcadores interaccionales en universidad y en las publicaciones.

Gráfico 10.1 Resultados marcadores interaccionales en universidad y publicaciones



Los resultados muestran que hay un mayor uso de los marcadores de postura ($Mdn = 0.82$, $Mdn = 0.41$) en las publicaciones y la universidad, respectivamente, que los marcadores de afiliación ($Mdn = 0.17$, $Mdn = 0.20$). La diferencia entre los marcadores de postura es significativa $Z = -2.52$, $p = 0.01$, $r = -.63$, mientras que la diferencia en el uso de marcadores de afiliación no es significativa $Z = -.84$. Respecto a los marcadores de la postura y de afiliación (véase el gráfico 10.1) los resultados muestran que las publicaciones ($Mdn = 0.14$) tienen significativamente más marcadores de actitud que los textos universitarios ($Mdn = 0.05$), $Z = -2.36$, $p = 0.01$, $r = -.59$.

Además las publicaciones ($Mdn = 0.39$) muestran significativamente mayor porcentaje de marcadores de auto-mención, que los textos universitarios ($Mdn = 0.21$), $Z = -2.10$, $p = 0.03$, $r = -.53$. Los textos universitarios tienen significativamente más porcentaje de pronombres del lector ($Mdn = 0.03$) que las publicaciones ($Mdn = 0.003$) $Z = -2.20$, $p = 0.02$, $r = -.55$. Para los demás marcadores, tales como los intensificadores, atenuadores, directivos al texto, al argumento y marcadores de conocimiento compartido, no hay diferencias significativas entre los trabajos universitarios y las publicaciones.

CAPÍTULO 11

11. Resumen de los resultados del estudio 2

A continuación presentamos un resumen de los resultados que dan cuenta de las diferencias significativas. En el estudio 2, en primer lugar, nos habíamos planteado examinar si había variación del **registro** en los trabajos de investigación universitarios y las publicaciones, analizada esta por medio de las características léxico-gramaticales, según el nivel académico. Los resultados obtenidos para el nivel académico refutan la hipótesis cero (H_0). Así, adoptamos la hipótesis alternativa (H_1), es decir:

H_1 : Las características léxico-gramaticales de los artículos publicados son significativamente diferentes a los trabajos de investigación universitarios.

Los resultados muestran que los trabajos de universidad tienen mayor longitud en número de palabras que las publicaciones. Por el contrario las publicaciones tienen mayor diversidad léxica. Los trabajos de universidad son más complejos, según la primera medida de complejidad sintáctica y son más densos, es decir, tienen mayor densidad léxica que las publicaciones. En cambio, las publicaciones tienen mayor complejidad sintáctica, medida esta como media de palabras antes del primer verbo finito.

En segundo lugar, en el estudio 2 también examinamos el **metadiscurso**, mediante los marcadores interaccionales, según el nivel académico. Los resultados obtenidos para el nivel académico refutan la hipótesis cero (H_0). En consecuencia, adoptamos la hipótesis alternativa (H_1):

H_1 : La frecuencia de los marcadores interaccionales es significativamente diferente entre los trabajos de investigación universitarios y las publicaciones.

Los resultados revelan que hay un mayor uso de los marcadores de postura en las publicaciones respecto a los trabajos de universidad. Asimismo, las publicaciones tienen más marcadores de actitud y más marcadores de auto-mención, que los textos universitarios. Por el contrario, los trabajos universitarios tienen más porcentaje de pronombres del lector que las publicaciones. Para los demás marcadores, tales como los intensificadores, atenuadores, directivos al texto, al argumento y marcadores de conocimiento compartido, no hay diferencias significativas entre los trabajos universitarios y las publicaciones.

CAPÍTULO 12

12. Discusión

En este estudio analizamos una variedad de características léxico-gramaticales y discursivas, que se supone que son propias del registro académico y en concreto del artículo de investigación. Si bien ha habido estudios previos que han analizado o bien el registro o el metadiscurso en artículos de investigación, nuestro análisis es innovador ya que examina la transición de ocho sujetos que escribieron sus textos en la universidad en catalán y posteriormente publicaron sus trabajos como profesionales en lengua inglesa. Nuestro análisis se centró, por una parte, en la diversidad léxica, la complejidad sintáctica y la densidad léxica; y por otra, en los marcadores de interacción, como una ventana que nos permitiera observar el desarrollo de las habilidades de escritura en una comunidad discursiva específica.

Encontramos diferencias en la longitud de los textos, en cuanto a número de palabras, pero este resultado podría explicarse debido a las instrucciones de la tarea. Aunque los textos universitarios tienen más palabras, se pedía a los estudiantes que limitaran sus escritos a 30 páginas. Además las normas de publicación de cada revista también imponen restricciones a la longitud de los artículos. Como ya hemos mencionado en el estudio 1, la diversidad léxica se considera que es una característica de los escritores competentes (Malvern et al., 2004) y esto parece ser consistente con nuestros hallazgos, ya que las publicaciones mostraron mayor diversidad léxica en comparación con los trabajos universitarios, sobre todo, si se tiene en cuenta que en las culturas orientadas a la forma, como el inglés, tienden a la repetición léxica (Cuenca, 2003).

El menor uso de la diversidad léxica en los textos universitarios, sin embargo, también podría explicarse por las características del registro. Como ya hemos mencionado los registros informativos, tales como el registro académico, se caracterizan por el uso de los mismos términos técnicos (Biber, 1995). A pesar de esto, hay otros recursos para crear la diversidad en los textos, por ejemplo, sustituciones de adjetivos o derivar adverbios de adjetivos (Hinkel, 2004). Parece que los estudiantes no hacen uso de estas técnicas.

Por otra parte, nuestro resultado en cuanto a la media de 25 palabras por oración en las publicaciones es consistente con la media de 24-28 palabras por oración reportada para el inglés en textos científicos (Bazerman, 1988; Gross, Harmon, & Reidy, 2002). Nuestro resultado, la media de 30 palabras por oración, en los trabajos universitarios, es similar al promedio reportado para la ciencia escrita en italiano, otra lengua romance (Scarpa, 2007). Para la segunda medida de complejidad sintáctica, la media de número de palabras antes del primer verbo finito en una oración, las diferencias observadas podrían explicarse por la tipología lingüística. El español y el catalán permiten cambiar el orden de los constituyentes de una oración, mientras que el inglés es más estricto en cuanto al orden de los elementos lingüísticos (Vázquez, Fernández, & Martí, 2000). Además, los textos universitarios muestran significativamente mayor densidad léxica que las publicaciones. Esta diferencia podría deberse a diferencias de idioma (Johansson, 2008) o podría ser que los estudiantes relacionan los textos densos al registro formal (Baratta, 2010).

Sugerimos que nuestros resultados podrían ser interpretados como que los estudiantes no equilibran el uso de la densidad léxica. Los escritores deberían de utilizar la densidad léxica de forma compensada, ya que un uso excesivo de nominalizaciones, por ejemplo, da la sensación de un texto que carece de fluidez (Pinker, 2014; Ventola, 1996).

Los siguientes ejemplos ilustran la densidad léxica y la complejidad sintáctica (medida como número de palabras en una oración). Como indicamos unas líneas más abajo, el ejemplo (b) es del apartado de introducción de un trabajo de final de máster universitario, y el ejemplo (c) de la parte introductoria de una publicación. En el (b) hay cinco sustantivos (NN), tres adjetivos (JJ) y tres frases preposicionales (IN) que funcionan como modificadores del sustantivo antes de la coma, mientras que en la publicación (c) solo aparecen tres sustantivos, tres adjetivos y un sintagma preposicional para expresar más o menos la misma idea.

b) L'asma és una malaltia respiratòria crònica, de caràcter al·lèrgic en la
 NN NN JJ JJ IN NN JJ IN
 majoria de casos,
 NN IN NN
 que té una alta prevalença en la nostra societat Europea actual.

c) Allergic asthma is a common inflammatory disease of the airway,
 JJ NN JJ JJ NN IN NN
 and long-term therapy is aimed at counteracting episodes of
 bronchospasm and reducing allergic inflammation.

Si analizamos estos ejemplos a la luz de las etapas de desarrollo de la complejidad en textos académicos de Biber, Gray & Poonpon (2013), podríamos decir que el ejemplo de los estudiantes (b) es más complejo que el de la publicación (c). El ejemplo (b) correspondería al estadio 5: amplia incrustación de frases en el núcleo del sustantivo, múltiples frases preposicionales que funcionan como post-modificadores, con niveles de incrustación. En cambio el ejemplo (c) correspondería al estadio 4: incrustación de frases en el núcleo del sustantivo, adjetivos atributivos y sustantivos como pre-modificadores. En este sentido, podríamos argumentar que los estudiantes no se ajustan a los requisitos de las convenciones específicas de su disciplina.

No obstante, hay que tomar esta argumentación con cautela, ya que las diferentes tipologías lingüísticas imponen sus restricciones al hacer comparaciones. Lo que si podemos afirmar es que la densidad léxica en la redacción científica se logra generalmente a través de las nominalizaciones (Guillén Galve, 1998; Halliday & Martin, 1993). En consecuencia, una investigación futura de las nominalizaciones sería necesaria para obtener una imagen más detallada de los trabajos de investigación de los estudiantes.

Un segundo objetivo del estudio 2 fue examinar la distribución de los marcadores de la postura y de afiliación en los trabajos de investigación y en los artículos publicados. Nuestro hallazgo, a partir de la diferencia entre el total de los marcadores de postura en los trabajos de universidad y las publicaciones, sugiere que llegar a dominar el discurso retórico de la comunidad específica es un proceso largo. Hay factores lingüísticos, cognitivos y sociales que interactúan en la transición de un escritor nativo a un escritor experto (Berman, 2004). Sin embargo, no es solo una cuestión de desarrollo, es también una cuestión de saber y de adecuación a la escritura académica. De hecho, el análisis que hizo Hyland (2005) de 240 artículos de investigación, en ocho disciplinas distintas, mostró que en todas las disciplinas los marcadores de postura son más comunes que los de afiliación. En nuestro corpus los estudiantes universitarios hacen un uso significativamente mayor de los pronombres del lector, que son marcadores de afiliación, en comparación con las publicaciones, que muestran una mayor proporción de marcadores de auto-mención.

Consideramos que la falta de una voz del autor, en los textos de los estudiantes, es una muestra de las diferencias fundamentales respecto a la finalidad de los trabajos de final de máster y las publicaciones. Ya hemos mencionado que se trata de géneros diferenciados y los estudiantes tienen que demostrar la adquisición de las habilidades necesarias (Gardner & Nesi, 2013).

No estamos sugiriendo que los estudiantes no han podido cumplir con las exigencias de las convenciones retóricas de la comunidad específica. El uso que hacen los estudiantes de los marcadores de afiliación podría corresponder a un compromiso con la audiencia a la que va dirigida, es decir, los estudiantes tienen que alinear su discurso, su conocimiento sobre el tema, con las personas que los han de evaluar y para ello incluyen al lector, en este caso el profesor.

Nuestros resultados muestran que no hay diferencias significativas en las proporciones de intensificadores y atenuadores entre los niveles, aunque el porcentaje de intensificadores en las publicaciones es casi cuatro veces superior. Para determinar si los intensificadores y los atenuadores han sido incorrectamente utilizados o no se ajustan a la sección correspondiente en los trabajos universitarios, se necesita un análisis adicional y más profundo.

En contra de lo esperado, ya que teníamos en cuenta los resultados del estudio de Hyland (2005), en que los atenuadores son más frecuentes que los intensificadores en las publicaciones, en nuestro corpus los intensificadores son más frecuentes que los atenuadores. Este resultado se podría explicar, en parte, teniendo en cuenta el alto nivel de competitividad en las ciencias biomédicas, en el que influyen factores como el patrocinio de las investigaciones o el factor de impacto de la revista. Asimismo, se supone que los resultados más robustos requieren menos atenuadores (Gross & Chesley, 2012).

Además, Hyland (1998) encontró en su análisis de artículos de biología que dentro del posicionamiento, los marcadores de actitud eran los menos frecuentes. En cambio, nuestros resultados muestran una diferencia significativa en las publicaciones respecto a los marcadores de actitud.

Los datos muestran que después de la auto-mención y los intensificadores, los marcadores de actitud son más frecuentes, incluso más que los atenuadores. Por ejemplo, en una publicación de nuestro corpus, encontramos un marcador de actitud expresando una valoración cinco veces, como en el ejemplo siguiente.

Interestingly, TREX2 expression is largely reduced in the mice lacking IKK α , which is required to maintain skin homeostasis and prevent skin cancer. (Sub-corpus TARBUC)

Los marcadores de actitud más comunes que encontramos, y esto concuerda con los hallazgos de Hyland (1998), se refieren a cuestiones que el escritor resalta como importantes o interesantes, tal como mostramos en el siguiente ejemplo:

Importantly, and in contrast to previous trials, our patients were HIV-1-infected individuals who were demonstrably immune suppressed, having failed to make serological and CD4 cell immune responses to vaccination with at least one of the three selected antigens (hepatitis A, hepatitis B and tetanus toxoid). (Sub-corpus TARBUC)

El uso de marcadores de actitud se basa en los valores de cada disciplina e incluso en los diferentes antecedentes culturales, tales como el inglés y el español que son superados por unas convenciones disciplinarias comunes (Mur, 2007). Como nuestra muestra de publicaciones revela un alto porcentaje de marcadores de actitud, el uso de estos marcadores podría ser un indicador del dominio de las convenciones propias del discurso de la comunidad disciplinar. Por lo tanto, sugerimos que la diferencia observada entre las publicaciones y los trabajos universitarios en el uso de estos marcadores puede ser debida a la falta de competencia de la retórica disciplinaria.

Nuestros resultados en cuanto a los resultados de auto-mención corroboran la afirmación de que el artículo de investigación es un género con un discurso altamente retórico (Bazerman, 1988; Hunston, 1994; Swales, 2004). Mediante la auto-mención los escritores proyectan una voz propia. Esta proyección personal permite al escritor construir su identidad, porque resalta su contribución individual (Hyland, 2004a). Nuestros resultados del alto porcentaje de auto-mención junto con un menor uso de atenuadores podrían ser interpretados como una manera de proyectar la autoridad profesional. No obstante, también hay que tener en cuenta otros factores como el cultural (Mur, 2007) o el dominio que el autor tenga del inglés, la lengua de difusión por excelencia (Martínez, 2005). Una explicación adicional, a las diferencias observadas, podría estar relacionada con la prescripción de manuales de estilo académico, que sugieren que en la escritura académica no deberían aparecer referencias al escritor o lector.

Nuestros resultados deben ser vistos a la luz de algunas limitaciones. Es un análisis basado en un tamaño de muestra pequeño, y por otra parte, se requiere precaución al comparar las características léxico-gramaticales en dos lenguas diferentes. Consideramos que sería necesario ampliar la investigación sobre estas características, para determinar en qué medida las diferencias entre los textos catalanes e ingleses se deben a la tipología lingüística.

CAPÍTULO 13

13. Conclusiones

En este estudio hemos examinado algunas características léxico-gramaticales, como la diversidad léxica, la complejidad sintáctica y la densidad léxica; así como los marcadores metadiscursivos de interacción, es decir, la postura y la afiliación, como reflejo de los cambios evolutivos de la escritura de investigación, que se dan desde la escuela secundaria hasta la publicación.

Los estudiantes adquieren nuevas formas de comunicación cuando hacen la transición de un contexto a otro, es decir, de la escuela secundaria a la universidad y de la universidad al mundo profesional. El escribir un trabajo de investigación es una tarea compleja que requiere la elección adecuada y la gestión de recursos lingüísticos para adaptarse al contexto del discurso. Los estudiantes han de adaptar el propio discurso al de la comunidad de práctica, a las convenciones propias de la comunidad específica. No obstante, los estudiantes tienen un conocimiento experiencial limitado en la escritura de géneros de investigación, aunque en niveles académicos avanzados es crucial para ellos llegar a dominar estos géneros si quieren ser miembros de una comunidad disciplinaria. Por lo tanto, nuestro principal objetivo era llenar el vacío existente en la investigación del registro y el metadiscurso en trabajos de investigación de estudiantes, escritos en lengua catalana.

Que tengamos noticia nosotros, este estudio es el primero en examinar el registro académico y la dimensión retórica del género de investigación de acuerdo con las características siguientes: por un lado, textos en catalán y escritos por autores cuyas edades estaban comprendidas entre la adolescencia y la edad adulta y, por otro, textos que marcan la transición del contexto universitario al profesional, del catalán al inglés.

Hemos proporcionado datos cuantificados para ampliar un poco más el panorama, el cual nos permite conocer mejor el desarrollo del lenguaje escrito y averiguar qué recursos lingüísticos y retóricos utilizan los estudiantes en el trabajo de investigación, es decir, cuando participan en las actividades de una comunidad de práctica. No obstante, necesitamos más investigaciones para confirmar si las diferencias observadas son propias del desarrollo, los indicios de una tendencia general en los trabajos de investigación escritos por estudiantes o simplemente un rasgo de nuestra muestra seleccionada.

A parte de algunas **limitaciones** de nuestro estudio, que ya hemos comentado en las discusiones, un aspecto que necesitaría de una mayor investigación sería examinar los marcadores interaccionales utilizados no como frecuencia total, sino en cada una de las diferentes secciones (IMRD) de los trabajos de investigación. Esto permitiría un mejor conocimiento del uso de estos, sobre todo de cara a la implementación de estrategias formativas futuras. Tampoco disponemos de ninguna información sobre las características de la alfabetización de los autores y, por supuesto, esta pequeña muestra no es representativa de todos los estudiantes.

Algunas **futuras líneas de investigación** que nos hemos propuesto, basadas en este corpus, serían examinar con más detalle la complejidad sintáctica, es decir, analizar las extensiones del grupo nominal, así como el desarrollo del uso de las nominalizaciones. Asimismo, sería relevante proporcionar un cuadro evolutivo de cómo se desarrollan las líneas de argumentación entre niveles académicos, esto es, cómo se presentan las aserciones, los argumentos, contraargumentos, evidencias, datos, etc. Estos aspectos estarían relacionados con el escritor, pero también sería sumamente interesante incluir al lector, ya que la escritura académica es una tríada entre escritor, texto y lector. En este sentido, y siguiendo la propuesta de Lewin (2005), que formuló la sugerente pregunta "¿quién decide cuándo un atenuador es un atenuador, el lingüista, el autor o el lector" (Lewin, 2005, p. 164), se necesitaría obtener una mayor comprensión de cómo los lectores reciben las intenciones de los escritores, esto es, profundizar en las percepciones del lector.

Por otro lado, nuestro estudio destaca por sus **aplicaciones en el ámbito educativo**. En primer lugar, hemos compilado un corpus de trabajos académicos de investigación de bachillerato y máster universitario en las disciplinas de biología e historia en catalán, además del sub-corpus de artículos publicados en inglés por los mismos estudiantes. Este corpus, llamado TARBUC (*Treballs Acadèmics de Recerca de Batxillerat i Universitat en Català*), está formado por textos producidos en situaciones reales, es decir, muestras de lengua en uso de estudiantes que están en el proceso de aprendizaje de la escritura académica y está disponible para la comunidad investigadora. Existen muchos corpus compilados que corresponden a muestras de lengua normativa, pero actualmente uno de los desafíos más interesantes es crear corpus de aprendices (Tolchinsky, 2014; Tolchinsky & Martí, 2013).

En segundo lugar, hemos aportado, con el análisis de nuestro corpus de aprendices, un marco evolutivo de la escritura académica, en concreto del registro y del metadiscurso, en el género de investigación desde bachillerato hasta la publicación. Estos resultados pueden ofrecer orientaciones para intervenciones didácticas con el propósito de mejorar la escritura de los estudiantes, tanto en L1 como en L2, orientaciones que ya hemos empezado a publicar (Pujol Dahme, 2014; 2015). Además, como todo corpus, se puede utilizar de una forma más directa, es decir, el profesorado puede presentar ejemplos de estos datos reales para discutirlos posteriormente con el alumnado (Carter & Mc Carthy, 1995).

En definitiva, nuestros resultados corroboran la aserción que el desarrollo de la alfabetización académica requiere de una adaptabilidad o flexibilidad retórica para seleccionar adecuadamente una combinación de formas y funciones lingüísticas y adaptarlas a diferentes registros y géneros. La capacidad de desplegar el registro académico y las convenciones retóricas, de forma adecuada a una comunidad discursiva específica, es en sí mismo un proceso prolongado.

Por último, con esta investigación hemos proporcionado una visión del desafío que supone no solo la transición de un contexto a otro, sino también de una lengua a otra. El panorama general que se desprende de nuestro estudio es un reflejo de la complejidad en la larga trayectoria hacia la maestría. En este proceso están estrechamente vinculados el repertorio lingüístico, la madurez cognitiva, las restricciones que imponen los géneros y la retórica de cada comunidad disciplinaria. Las implicaciones de nuestros resultados han sido discutidas teniendo en cuenta factores lingüísticos, cognitivos y socioculturales. Por consiguiente, esperamos haber proporcionado un respaldo más para integrar estas perspectivas en la caracterización de las etapas evolutivas del lenguaje escrito. Concretamente, en el género de investigación, cuyo dominio es fundamental para lograr el éxito académico.

Bibliografía

- Ädel, A., & Reppen, R. (Eds.) (2008). *Corpora and discourse: The challenges of different settings*. Amsterdam: John Benjamins.
- Alamargot, D., & Andriessen, J. (2002). The "power" of text production activity in collaborative modelling: Nine recommendations to make a computer supported situation work. En P. Brna, M. Baker, K. Stenning, & A. Tiberghien (Eds.), *The role of communication in learning to model* (pp. 275-299). Mahwah, NJ: Lawrence Erlbaum Associates.
- Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Anthony, L. (2011). Antconc (Version 3.2.4). Tokyo, Japan: Laurence Anthony, University of Waseda.
- Arendt, H. (1958). *The human condition* (2nd ed.). The University of Chicago Press.
- Atienza, E. (1992). Propuesta de evaluación del texto escrito en enseñanza secundaria. *Tesis doctoral*. Universidad de Barcelona.
- Atkins, S., Clear, J., & Osler, N. (1992). Corpus design criteria. *Literary and Linguistic Computing*, 7(1), 1-16.
- Aull, L. L., & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus based-comparison. *Written Communication*, 31(2), 151 – 183.
- Austin, J. (1962). *How to do things with words: the William James lectures delivered at Harvard university in 1955*. Oxford, MA: Clarendon Press.
- Bakhtin, M. (1981). *The dialogic imagination*. Austin: University of Texas.
- Bakhtin, M. (1986). The problem of speech genres. En C. Emerson, & M. Holquist (Eds.), *Speech genres and other late essays* (V. McGee, Trad., pp. 60-101). Austin, TX: University of Texas Press.
- Bamford, J., & Bondi, M. (2005). Introduction. En *Dialogue within discourse communities* (pp. VII-XXIV). Tübingen: Max Niemeyer.
- Baratta, M. (2010). Nominalization development across an undergraduate academic. *Journal of Pragmatics*, 42(4), 1017-1036.

- Bar-Ilan, L., & Berman, R. (2007). Developing register differentiation: The Latinate-Germanic divide in English. *Linguistics*, 45(1), 1-35.
- Barton, D., Hamilton, M., & Ivanic, R. (Eds.). (2000). *Situated literacies: Reading and writing in context*. London: Routledge.
- Bawarshi, A. S., & Reiff, M. J. (2010). *Genre: An introduction to history, theory, research, and pedagogy*. West Lafayette, Indiana: Parlor Press.
- Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison, WI: The University of Wisconsin Press.
- Bazerman, C. (1994). Systems of genres and the enactment of social intentions. En A. Freedman, & P. Medway (Eds.), *Genre and the new rhetoric* (pp. 83-96). London, UK: Taylor & Francis.
- Bazerman, C. (2004). Intertextualities: Volosinov, Bakhtin, Literary Theory, and Literacy Studies. En A. Ball, & S. Warshauer Freedman (Eds.), *Bakhtinian perspectives on language, literacy, and learning* (pp. 53-65). Cambridge: Cambridge University Press.
- Bazerman, C. (2005). Communication in the scientific community. En S. Restivo, *Science, technology, and society: An encyclopedia* (pp. 55-61). New York, NY: Oxford University Press.
- Bazerman, C. (2008). *Handbook of research on writing: History, society*. New York, NY: Lawrence Erlbaum Associates.
- Bazerman, C., & Paradis, J. (Eds.). (1991). *Textual Dynamics of the professions: Historical and contemporary studies of writing in professional communities*. WAC Clearinghouse Landmark Publications in Writing Studies. Madison, WI: University of Wisconsin Press.
- Bazerman, C., & Prior, P. (2005). Participating in emergent socio-literate worlds: Genre, disciplinarity, interdisciplinarity. En J. Green, & R. Beach (Eds.), *Multidisciplinary perspectives on literacy research* (pp. 133-178). NCTE.
- Beaugrande, R. (1980). *Text, discourse, and process: toward a multidisciplinary science of texts*. Norwood, NJ: ALEX.
- 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, NJ: Ablex Publishing Corporation.

- Beaugrande, R., & Dressler, W. (1981). *Einführung in die Textlinguistik*. Tübingen: Max Niemeyer.
- Bednarek, M. (2010). Corpus linguistics and systemic functional linguistics: Interpersonal meaning, identity and bonding in popular culture. En M. Bednarek, & J. Martin (Eds.), *New discourse on language: Functional perspectives on multimodality, identity and affiliation* (pp. 412-460). London, UK: Continuum.
- Beers, S., & Nagy, W. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing: An Interdisciplinary Journal*, 22, 185-200.
- Bennett, K. (201). Academic discourse in Portugal: A whole different ballgame? *Journal of English for Academic Purposes*, 9, 21-32.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.
- Berkenkotter, C., & Huckin, T. N. (1993). Rethinking genre from a sociocognitive perspective. *Written Communication*, 10(4), 475-509.
- Berkenkotter, C., & Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale, NJ, England: Erlbaum Associates.
- Berkenkotter, C., Huckin, T., & Ackerman, J. (1991). Social context and socially constructed tests: The initiation of a graduate student into a writing research community. En C. Bazerman, & J. Paradis (Eds.), *Textual dynamics of the professions: Historical and contemporary studies of writing in professional communities* (pp. 191-215). Madison, Wisconsin: The University of Wisconsin Press.
- Berman, R. (2011). Language development and literacy. En R. Levesque (Ed.), *Encyclopedia of adolescence* (pp. 1548-1557). New York, NY: Springer.
- Berman, R. (2004). Between emergence and mastery: The long developmental route of language acquisition. En R. Berman (Ed.), *Language development across childhood and adolescence* (pp. 9-34). Amsterdam: Benjamins.
- Berman, R. (2007). Developing linguistic knowledge and language use across adolescence. En E. Hoff, & M. Schatz, *Blackwell handbook of language development* (pp. 347-367). Malden, MA: Blackwell Publishing.

- Berman, R., & Katzenberger, I. (2004). Form and function in introducing narrative and expository texts: A developmental perspective. *Discourse Processes*, 39(1), 57-94.
- Berman, R., & Ravid, D. (2009). Becoming a literate language user: Oral and written text construction across adolescence. En D. R. Olson, & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 92-111). Cambridge: Cambridge University Press.
- Berman, R., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language and Literacy*, 5(1), 1-43.
- Berman, R., Ragnarsdóttir, H., & Strömqvist, S. (2002). Discourse stance. *Written Languages and Literacy*, 5(2), 255-289.
- Berns, M., & Matsuda, P. (2006). Applied linguistics: Overview and history. En K. Brown (Ed.), *The encyclopedia of language and linguistics* (pp. 394-405). Oxford, UK: Elsevier.
- Bhatia, V. (1993). *Analysing genre: Language use in professional settings*. London: Longman.
- Bhatia, V. (2002a). Applied genre analysis: Analytical advances and pedagogical procedures. En A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 279-284). Mahwah, NJ: Erlbaum.
- Bhatia, V. (2002b). Applied genre analysis: A multi-perspective model. *Ibérica. Revista de la Asociación Europea de Lenguas para Fines Específicos*(4), 3-19.
- Bhatia, V. (2004). *Worlds of written discourse-A genre-based view*. London, UK: Continuum.
- Bhatia, V., Flowerdew, J., & Jones, R. (Eds.). (2008). *Advances in discourse studies*. New York, NY: Routledge.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. (1993). Using register-diversified corpora for general language studies. *Computational Linguistics*, 219-242.
- Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.

- Biber, D. (2006). *University Language: A Corpus-based study of spoken and written registers*. Amsterdam, Netherlands: John Benjamins.
- Biber, D. (2009). Are there linguistic consequences of literacy? Comparing the potentials of language use in speech and writing. En D. Olson, & N. Torrance (Eds.), *Cambridge handbook of literacy* (pp. 75-91). Cambridge: Cambridge University Press.
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. New York, NY: Cambridge University Press.
- Biber, D., & Finegan, E. (1989). Styles of stance in English: lexical and grammatical marking of evidentiality and affect. *Text*, 93-124.
- Biber, D., & Finegan, E. (1994). *Sociolinguistic perspectives on register*. New York: Oxford University Press.
- Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9(1), 2-20.
- Biber, D., Connor, U., & Upton, T. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. Amsterdam, Netherlands: John Benjamins Publishing.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge, UK: Cambridge University Press.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5-35.
- Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology*. New York, NY: Cambridge University Press.
- Blue, G. (2003). *Developing academic literacy*. Bern: Peter Lang.
- Bonilla, S. (1997). Estudio preliminar. Traducción de la obra de Beaugrande R. y Dressler W. En *Introducción a la lingüística del texto* (pp. 7-25). Barcelona: Ariel.
- Bourdieu, P. (1991). *Language & symbolic power*. Cambridge, MA: Harvard University Press.
- Breetveld, I., van den Bergh, H., & Rijlaarsdam, G. (1994). Relations between writing processes and text quality: when and how? *Cognition and Instruction*. 12, 103-123.

- Bruce, I. (2008). *Academic writing and genre: A systematic analysis*. London, UK: Continuum.
- Butler, C., & González-García, F. (2014). *Exploring functional-cognitive space. Studies in Language Companion Series, 157*. Amsterdam, Netherlands: John Benjamins.
- Calsamiglia, H., & Tusón, A. (1999). *Las cosas del decir. Manual de análisis del discurso*. Barcelona: Ariel Lingüística.
- Candlin, C., & Hyland, K. (1999). *Writing: Texts, processes and practices*. London, UK: Longman.
- Cantero, F., & Mendoza, A. (2003). Conceptos básicos de la lengua y la literatura. En A. Mendoza, & E. Briz (Eds.), *Didáctica de la lengua y la literatura para primaria* (pp. 33-78). Madrid: Prentice Hall.
- Cañedo-Arguelles, M. (2013). Ecología. En L. Tolchinsky (Ed.), *La escritura académica a través de las disciplinas* (pp. 127-148). Barcelona: Octaedro.
- Carlino, P. (2003). Alfabetización académica: Un cambio necesario, algunas alternativas posibles. *Educere*, 6(20), 409-420.
- Carlino, P. (2005). *Escribir, leer y aprender en la universidad. Una introducción a la alfabetización académica*. Buenos Aires, Argentina: Fondo de cultura económica.
- Carlino, P. (2013). Alfabetización académica diez años después. *Revisat mexicana de investigación educativa*, 18(57), 355-381.
- Carter, R., & McCarthy, M. (1995). Grammar; the spoken language. *Applied Linguistics*, 16(2), 141-158.
- Cassany, D., & Castellà, J. (2010). Aproximación a la literacidad crítica. *Perspectiva, Florianópolis*, 28(2), 353-374.
- Castelló, M. (2014). Los retos actuales de la alfabetización académica: estado de la cuestión y últimas investigaciones. En I. Ballano, & I. Muñoz (Eds.), *Escribir en el contexto académico* (pp. 153-173). Publicaciones de la Universidad de Deusto, Serie Mayor.
- Castelló, M., & Iñiesta, A. (2012). Texts as artifacts-in-activity. Developing authorial identity and academic voice in writing academic research papers. En M. Castelló, & C. Donahue (Eds.), *University writing: Selves and texts in academic societies. Studies in writing* (Vol. 24, pp. 179-200). Bingley, UK: Emerald.

- Castelló, M., Bañales, G., & Vega, A. (2010). Enfoques en la investigación de la regulación de escritura académica: Estado de la cuestión. *Electronic Journal of Research in Educational Psychology*, 8, 1253-1282.
- Castelló, M., Corcelles, M., Iñiesta, A., Bañales, G., & Vega, N. (2011a). La voz del autor en la escritura académica: Una propuesta para su análisis. *Signos*, 44(76), 105-117.
- Castelló, M., Iñiesta, A., Pardo, M., Liesa, E., & Martínez-Fernández, R. (2011b). Tutoring the end-of-studies disseratation: helping psychology students find their academic voice when revising academic texts. *Higher Education*, 63(1), 97-115.
- Chafe, W. (1986). Evidentiality in English conversation and academic writing. En *Evidentiality: the linguistic coding of epistemology* (pp. 261-330). Norwood, NJ: Ablex Publisjing Corporation.
- Chafe, W., & Danielewicz, J. (1987). Properties of spoken and written language. En R. Horowitz, & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 83-113). San Diego, California: Academic Press.
- Charles, M., Pecorari, D., & Hunston, S. (Eds.). (2009). *Academic writing: At the interface of corpus and discourse*. London, UK: Continuum.
- Chitez, M., & Kruse, O. (2012). Writing cultures and genres in European higher education. En M. Castelló, & C. Donahue (Eds.), *University writing: Selves and texts in academic societies*. *Studies in writing* (Vol. 24, pp. 151-175). Bingley, UK: Emerald.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT.
- Christenbury, L., Bomer, R., & Smagorinsky, P. (2009). Introduction. En L. Christenbury, R. Bomer, & P. Smagorinsky (Eds.), *Handbook of adolescent literacy research* (pp. 3-14). New York: The Guilford Press.
- Christie, F. (s.f.). The chaniging face of literacy. En F. Christie (Ed.), *Literacy for a changing world: A fresh look at the basis* (pp. 1-25). Hawthorne, AU: The Australian Council for Educational Research.
- Christie, F., & Derewianka, B. (2008). *School discourse: Learning to write across the years of schooling*. London, U.K.: Continuum.
- Ciapuscio, G. (2005). La noción de género en la lingüística sistémico funcional y en la lingüística textual. *Signos*, 38(57), 31-48.

- Climent, S., & Mateu, J. (2009). *Teories lingüístiques contemporànies. L'estructura de les llengües*. Barcelona: UOC.
- Clyne, M. (1994). *Inter-cultural communication at work. Cultural values in Discourse*. Cambridge: Cambridge University Press.
- Coffin, C. (2002). The voices of history: theorising the interpersonal semantics of historical discourses. *Text*, 22(4), 503-528.
- Cole, M. (1996). *Cultural Psychology*. Boston, MA: Harvard University Press.
- Colombi, M. (2002). Academic language development in Latino Students' writing in Spanish. En M. Schlepppegrell, & M. Colombi, *Developing Advanced Literacy in First and Second Languages*. (pp. 67-86). Mahwah, NJ: Erlbaum.
- Colomer, T., & Camps, A. (1996). *Enseñar a leer, enseñar a comprender*. Madrid: Celeste/MEC.
- Connor, U. (2002). New directions in contrastive rhetoric. *Tesol Quarterly*, 36(4), 493-510.
- Conrad, S. (1996). Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education*, 8, 299-326.
- Conrad, S. (2002). Corpus linguistic approaches for discourse analysis. *Annual Review of Applied Linguistics (Discourse and Dialog)*(22), 75-95.
- Conrad, S. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. En S. Conrad & D. Biber (Eds.) *Multi-dimensional studies of register variation in English* (p. 94-107). Harlow, England: Pearson Education
- Cope, B., & Kalantzis, M. (2000). *Multiliteracies: Literacy learning and the design of social futures*. London, UK: Routledge.
- Correig, M. (2008). ¿Qué es leer? ¿Qué es escribir? En M. Bigas, & M. Correig (Eds.), *Didáctica de la lengua en educación infantil* (pp. 125-156). Madrid: Síntesis Educación.
- Cortes, V. (2002). Lexical bundles in freshman composition. En R. Reppen, S. M. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 131-146). Philadelphia, PA: John Benjamins.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397-423.
- Coulmas, F. (1989). *The writing systems of the world*. Oxford, UK: Basil Blackwell.

- Coulmas, F. (2003). *Writing systems: An introduction to their linguistic analysis*. Cambridge, UK: Cambridge University Press.
- Crismore, A. (1989). *Talking with readers: Metadiscourse as rhetorical act*. New York, NY: Peter Lang.
- Crismore, A., Markkanen, R., & Steffensen, M. (1993). Metadiscourse in persuasive writing: A study of texts written by american and finnish university students. *Written Communication*, 10(1), 39-71.
- Croft, W., & Cruse, A. (2004). *Cognitive linguistics (Cambridge Textbooks in Linguistics)*. Cambridge, UK: Cambridge University Press.
- Cuenca, M. (2003). Two ways to reformulate: A contrastive analysis of reformulation markers. *Journal of Pragmatics*, 35(7), 1069-1093.
- Culicover, P., & Jackendoff, R. (2005). *Simpler Syntax*. Oxford, UK: Oxford University Press.
- Cutillas, L., Tolchinsky, L., Rosado, E., & Perera, J. (2014). Indicators of lexical growth throughout age, genre and modality for a Catalan L1 corpus. En A. Díaz-Negrillo, & F. Díaz-Pérez, *Specialisation and variation in language corpora* (pp. 161-188). Bern: Peter Lang.
- Donahue, C. (2011). Cross-cultural approaches to writing and disciplinary. En M. Deane, & P. O'Neill (Eds.), *Writing in the disciplines* (pp. 14-26). London, UK: Palgrave MCMillan.
- Dressler, W. (1978). *Textlinguistik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287-313.
- Dubois, B. (1982). The construction of noun phrases in biomedical journal articles. En J. Hoedt, *Pragmatics and LSP* (pp. 49-67). Copenhagen: Copenhagen School of Economics.
- Duschl, R. (2000). Making the nature of science explicit. En R. Millar, J. Leach, & J. Osborne (Eds.), *Improving science education: The contribution of research* (pp. 187-206). Philadelphia, PA: Open University Press.
- EAGLES Guidelines. Expert Advisory Group on Language Engineering Standards. (1996). <http://www.ilc.cnr.it/EAGLES96/browse.html>.

- Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journals of Second Language Writing*, 139-155.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki, FI: Orienta_Konsultit.
- Englert, C., Mariage, T., & Dunsmore, K. (2006). Tenets of socio-cultural theory in writing instruction research. En C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *The handbook of writing research* (pp. 235-247). New York, NY: Guilford Press.
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument for studying science discourse. *Science Education*, 88, 915-933.
- Estrada, F. (2013). Antropología social y cultural. En L. Tolchinsky (Ed.), *La escritura académica a través de las disciplinas* (pp. 47-74). Barcelona: Octaedro.
- Fang, Z., Schleppegrell, M., & Cox, B. (2006). Understanding the language demands of schooling: Nouns in academic registers. *Journal of Literacy Research*, 38(3), 247-273.
- Fairclough, N. (2004). *Analyzing discourse: Textual analysis for social research*. London, UK: Routledge.
- Ferreiro, E. (1999). *Cultura escrita y educación*. México: Fondo de Cultura Económica.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London, UK: SAGE.
- Fillmore, C. (1992). Corpus linguistics vs. computer-aided armchair linguistics. En J. Svartvik (Ed.), *Directions in corpus linguistics* (pp. 35-60). Berlin, D: DeGruyter.
- Flower, L., & Hayes, J. (1980). The cognition of discovery: defining a rhetorical problem. *College Composition and Communication*, 31, 21-32.
- Flowerdew, J. (2002). Genre in the classroom: A linguistic approach. En A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 91-102). Mahwah: Lawrence Erlbaum Associates.
- Flowerdew, J. (2014). Foreword. En L. Gil-Salom, & C. Soler- Montreal (Eds.), *Dialogicity in written specialised genres* (pp. IX-XVI). Amsterdam, The Netherlands: John Benjamins.
- Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for Specific Purposes*, 24, 321-332.

- Flowerdew, L. (2008). *Corpus-based analysis of the problem-solution pattern: A phraseological approach*. Amsterdam, Netherlands: John Benjamins.
- Flowerdew, L. (2012). *Corpora and language education*. (C. N. Candlin, & D. R. Hall, Eds.) New York, NY: Palgrave Mcmillan.
- Foucault, M. (1991). *The Foucault reader*. (P. Rabinow, Ed.) London, UK: Penguin.
- Francis, W. (1992). Language corpora BC. En J. Svartvik (Ed.), *Directions in corpus linguistics. Proceedings of the Nobel simposium* (pp. 17-32). Stockholm, SW.
- Freedman, A., & Medway, P. (1994). *Genre and the new rhetoric*. London, UK: Taylor & Francis.
- Freire, P. (1970). *Pedagogy of the oppressed*. (M. Bergman, Trad.) Harmondsworth, UK: Penguin.
- Galbraith, D. (1999). Writing as a knowledge-constituting process. En M. Torrance, & D. Galbraith (Eds.), *Knowing what to write: conceptual processes in text production* (pp. 139-159). Amsterdam, Netherlands: Amsterdam University Press.
- Galbraith, D. (2001). Writing Instruction. En N. Smelser, & P. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 16625-16629). Amsterdam, Netherlands: Elsevier.
- Gardner, S. (2012). Genres and registers of student report writing: An SFL perspective on texts and practices. *Journal of English for Academic Purposes*, 11(1), 52-63.
- Gardner, S., & Nesi, H. (2013). A classification of genre families in university student writing. *Applied Linguistics*, 34(1), 25-52.
- Gee, J. (1987). What is literacy? *Teaching and Learning*, 2(1), 1-11.
- Gee, J. (1990). *Social linguistics and literacies: Ideologies in discourses* (5th edition ed.). London, UK: Taylor & Francis.
- Gee, J. (1996). *Social linguistics and literacies: Ideology in discourses*. London, UK: Routledge Falmer.
- Gee, J. (2010). A situated sociocultural approach to literacy and technology. En E. Baker (Ed.), *The New Literacies: Multiple perspectives on research and practice* (pp. 165-193). New York, NY: The Guilford Press.
- Gee, J., & Handford, M. (Eds.). (2012). *The Routledge handbook of discourse analysis*. London, UK: Routledge.

- Giltrow, J., Gooding, R., Burgoyne, D., & Sawatsky, M. (2005). *Academic writing: An introduction* (Third ed.). Broadview Press.
- Gledhill, C. (1996). Science as collocation: Phraseology in cancer research articles. En S. Botley (Ed.), *Proceedings of teaching and language corpora conference* (Vol. 9, pp. 108-126). Lancaster, UK: University of Lancaster.
- Gonzálvez-García, F., & Butler, C. (2006). Mapping functional-cognitive space. *Annual Review of Cognitive Linguistics*(4), 39-96.
- Goody, J., & Watt, I. (1968). The consequences of literacy. En J. Goody (Ed.), *Literacy in traditional societies* (pp. 27-86). Cambridge, U.K.: Cambridge University Press.
- Grabe, W. (2002). Narrative and expository macro-genres. En A. Johns, *Genre in the Classroom: Multiple Perspectives*. (pp. 249-269). Mahwah, NJ: Lawrence Erlbaum Associates.
- Grabe, W., & Kaplan, R. (1996). *Theory and practice of writing*. Essex, UK: Longman.
- Gray, B., & Biber, D. (2012). Current conceptions of stance. En K. Hyland, & C. Sancho Guinda (Eds.), *Stance and voice in written academic genres* (pp. 15-33). London, UK: Palgrave McMillan.
- Gray, B. (2015). Linguistic variation in research articles: When discipline tells only part of the story. Amsterdam, Netherlands: John Benjamins.
- Gries, S. (2010). Corpus linguistics and theoretical linguistics. A love-hate relationship? Not necessarily.... *International Journal of Corpus Linguistics*, 15(3), 327-343.
- Gries, S. (2012). Corpus linguistics, theoretical linguistics, and cognitive/psycholinguistics: towards more and more fruitful exchanges. En J. Mukherjee, & M. Huber (Eds.), *Corpus linguistics and variation in English: Theory and description* (pp. 41-63). Amsterdam, Netherlands: Rodopi.
- Gross, A. G., & Chesley, P. (2012). Hedging, stance and voice in mediacy research articles. En K. Hyland, & C. Sancho Guinda (Eds.), *Stance and voice in written academic genres* (pp. 85-100). New York, NY: Palgrave Mcmillan.
- Gross, A. G., Harmon, J. E., & Reidy, M. (2002). *Communicating science: The scientific article from the 17th century to the present*. New York, NY: Oxford University Press.

- Guillén Galve, I. (1998). The textual interplay of grammatical metaphor on the nominalizations occurring in written medical English. *Journal of Pragmatics*, 30(3), 363-385.
- Habermas, J. (1984). *The theory of communicative action*. Boston, MA: Beacon Press.
- Halliday, M. (1973). *Explorations in the functions of language*. London, UK: Edward Arnold.
- Halliday, M. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Baltimore: University Park Press.
- Halliday, M. (1985). *An introduction of functional grammar* (3rd ed.). New York, NY: Routledge.
- Halliday, M. (1985). *Spoken and written language*. Victoria , Australia: Deakin University.
- Halliday, M. (1993). Some grammatical problems in scientific English. En M. Halliday, & J. R. Martin (Eds.), *Writing science:Literacy and discourse power* (pp. 69-85). Pittsburgh, PA: University of Pittsburgh Press.
- Halliday, M. (2002). *Linguistic studies of text and discourse*. (J. Webster, Ed.) New York, NY: Continuum.
- Halliday, M. (2004b). *Language of early childhood*. (J. Webster, Ed.) New York, NY: Continuum.
- Halliday, M. (2006). Afterwords. En G. Thompson, & S. Hunston (Eds.), *System and corpus: Exploring connections* (pp. 293-299). London, Uk: Equinox.
- Halliday, M. A. (2004a). *The language of science*. New York, NY: Continuum.
- Halliday, M. A., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective* (2nd ed.). Oxford: Oxford University Press.
- Halliday, M., & Martin, J. R. (1993). *Writing science:Literacy and discourse power*. Pittsburg, PA: University of Pittsburgh Press.
- Halliday, M., & Matthiessen, C. (1985). *Halliday's introduction to functional grammar* (4th ed.). New York, NY: Routledge.
- Hardy, J., & Römer, U. (2013). Revealing disciplinary variation in student writing: a multi-dimensional analysis of the Michigan Corpus of Upper-level Student Papers (MICUSP), *Corpora* 8(2), 183-107.
- Harris, Z. (1959). *Papers in structural and transformational linguistics*. Dordrecht, Netherlands: D. Reidel Publishing.

- Hayes, J. (1996). A new framework for understanding cognition and affect in writing. En C. Levy, & S. Ransdell (Eds.), *The science of writing* (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hayes, J., & Nash, J. (1996). On the nature of planning in writing. En C. Levy, & S. Ransdell (Eds.), *The science of writing* (pp. 29-55). Mahwah, NJ: Lawrence Erlbaum Associates.
- Heinemann, M., & Heinemann, W. (2002). *Grundlagen der Textlinguistik: Interaktion-Text-Discurs*. Tübingen: Niemeyer.
- Heinemann, W. (s.f.). Textsorten. Zur Diskussion um Basisklassen des Kommunizierens. Rückschau und Ausblick. En K. Adamzik (Ed.), *Textsorten. Reflexionen und Analysen* (pp. 9-28). Tübingen: Stauffenburg.
- Henderson, A., & Barr, R. (2010). Comparing indicators of authorial stance in psychology students' writing and published research articles. *Journal of Writing Research*, 2, 245-264.
- Hinds, J. (1987). Reader versus writer responsibility: a new typology. En U. Connor, & R. Kaplan, *Writing across languages: Analyses of L2 Text*. (pp. 141-152). Reading, MA: Addison Wesley.
- Holtz, M. (2011). Lexico-grammatical properties of abstracts and research articles. A corpus-based study of scientific discourse from multiple disciplines. Tesis Doctoral. Technische Universität Darmstadt.
- Hood, S. (2004). Managing attitude in undergraduate academic writing: a focus on the introductions to research reports. En L. Ravelli, R. Ellis, L. Ravelli, & R. Ellis (Eds.), *Analysing Academic writing: contextualized frameworks* (pp. 24-44). London: Continuum.
- Hunston, S. (1994). Evaluation and organisation in a sample of written academic discourse. En M. Coulthard (Ed.), *Advances in written text analysis* (pp. 191-218). London, UK: Routledge.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, UK: Cambridge University Press.
- Hyland, K. (1996). Writing without conviction? Hedging in scientific research articles. *Applied Linguistics*, 17(4), 433-454.
- Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: John Benjamins.

- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London, UK: Longman.
- Hyland, K. (2003). Self-citation and self-reference: Credibility and promotion in academic publication. *Journal of the American Society for Information Science and Technology*, 54(3), 251-259.
- Hyland, K. (2004a). A convincing argument: Corpus analysis and academic persuasion. En U. Connor, & T. A. Upton (Eds.), *Discourse in the professions: Perspectives from corpus linguistics* (pp. 87-112). Amsterdam: John Benjamins.
- Hyland, K. (2004b). *Genre and second language writing*. Ann Arbor: University of Michigan Press.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192 doi:10.1177/1461445605050365.
- Hyland, K. (2008a). Persuasion, interaction and the construction of knowledge: Representing self and others in research writing. *International Journal of English Studies*, 8(2), 1-23.
- Hyland, K. (2008b). Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18(1), 41-62.
- Hyland, K., & Hamp-Lyons, L. (2002). EAP: issues and directions. *Journal of English for Academic Purposes*, 1(1), 1-12.
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6(2), 183-205.
- Hyland, K., & Milton, J. (1999). Assertions in students' academic essays: A comparison of English NS and NNS student writers. En R. Berry, B. Asker, K. Hyland, & M. Lam (Eds.), *Language analysis, description, and pedagogy* (pp. 147-161). Hong Kong: Hong Kong University of Science and Technology, Language Center.
- Hyland, K., & Sancho-Guinda, C. (2012). Introduction: a context-sensitive approach to stance and voice. En K. Hyland, & C. Sancho-Guinda (Eds.), *Stance and voice in academic writing genres* (pp. 1-14). New York, NY: Palgrave MacMillan.
- Hyland, K., & Tse, P. (2004 : 167). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25(2), 156-177.

- Hymes, D. (1994). Toward ethnographies of communication. En J. Maybin (Ed.), *Language and literacy in social practice: A reader* (pp. 11-22). Avon, UK: Multilingual Matters Ltd.
- Hymes, D. H. (1972). On communicative competence. En J. B. Pride, & J. Holmes (Eds.), *Sociolinguistics. Selected readings* (pp. 269-293). London, UK: Penguin.
- Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly*, 30(4), 693-722.
- Iñiesta, A. (2009). La regulació en l'escriptura d'articles de recerca. Estratègies d'escriptors experts en castellà com a primera llengua i en anglès com a llengua internacional. *Tesis Doctoral*. Universitat Ramon Llull.
- Ivanic, R. (1998). *Writing and identity. The discursual construction of identity in academic writing*. Amsterdam, Netherlands: Benjamins.
- Jiménez-Aleixandre, M., & Erduran, S. (2008). Argumentation in science education: An overview. En S. Erduran, & M. Jiménez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 3-29). Dordrecht, Netherlands: Springer.
- Jisa, H. (2004). Growing into academic french. En R. Berman (Ed.), *Language development across childhood and adolescence* (pp. 135-162). Amsterdam, Netherlands: John Benjamins.
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working Papers (Lund University)*, 53, 61-79.
- Johns, A. (1998). Discourse communities and communities of practice: Membership, conflict, and diversity. En A. Johns, *Text, Role, and context: Developing academic literacies* (pp. 51-70). Cambridge, NY: Cambridge University Press.
- Johns, A. (2001). *Genres in the classroom: relating theory and research to practice*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Jones, R. (2012). *Discourse analysis: A resource book for students*. London, UK: Routledge.
- Jørgensen, M., & Phillips, L. (2002). *Discursive analysis as theory and method*. London, UK: SAGE Publications.

- Kaplan, R. (1966). Cultural thought patterns in intercultural education. *Language Learning*, 16, 1-20.
- Karmiloff-Smith, A. (1992). *Beyond modularity: A development perspective on cognitive science*. Cambridge, MA: MIT Press/Bradford Books.
- Kellogg, R. (1994). *The psychology of writing*. New York, NY: Oxford University Press.
- Kellogg, R. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1(1), 1-26.
- Kelly, G. (2008). Inquiry, activity, and epistemic practice. En R. Duschl, & R. Grandy (Eds.), *Teaching scientific inquiry: Recommendations for research and implementation* (pp. 99-117). Rotterdam, Netherlands: Sense.
- Kelly, G., & Crawford, T. (1997). An ethnographic investigation of the discourse processes of school science. *Science Education*, 81(5), 533-559.
- Kintsch, W., & Van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2015). Students' learning activities while studying biological process diagrams. *International Journal of Science Education*, 37, 1915-1937.
- Kress, G. (2000). Multimodality. En B. Cope, & M. Kalantzis (Eds.), *Multiliteracies: Literacy learning and the design of social futures* (pp. 182-202). London, UK: Routledge.
- Kress, G., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.
- Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: Rhetorics of the science classroom*. London, UK: Continuum.
- Kristeva, J. (1980). *Desire in language: A semiotic approach to literature and art*. New York, NY: Columbia University Press.
- Kucer, S. (2001). *Dimensions of Literacy: A conceptual base for teaching reading and writing in school settings*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3), 319-337.
- Kuo, C. H. (1999). The Use of Personal Pronouns: Role Relationships in Scientific Journal Articles. *English for Specific Purposes*, 18(2), 121-138.

- Lafuente, E. (2008). Epistemic and approximative meaning revisited: the use of hedges, boosters and approximators when writing research in different disciplines. En S. Burgess, & P. Martín, *English as an Additional Language in Research Publication and Communication* (pp. 65-82). Bern: Peter Lang.
- Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, 2(4), 458-508.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Lee, D. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the bnc jungle. *Language Learning & Technology*, 5(3), 37-72.
- Leech, G. (1991). The state of the art in corpus linguistics. En K. Aijmer, & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 8-29). Harlow, UK: Longman.
- Leech, G. (1992). Corpora and theories of linguistic performance. En *Directions in corpus linguistics* (pp. 105-122). Berlin, D.: de Gruyter.
- Lemke, J. (1989). Social semiotics: A new model for literacy education. En D. Bloome, *Classrooms and literacy* (pp. 289-309). Norwood, NJ: Ablex.
- Lemke, J. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex.
- Lemke, J. (1998a). Multimedia demands of the scientific curriculum. *Linguistics and Education*, 10(3), 247-272.
- Lemke, J. (1998b). Multiplying meaning: Visual and verbal semiotics in scientific text. En J. Martin, & R. Veel (Eds.), *Reading science* (pp. 87-113). London, UK: Routledge.
- Lemke, J. (2002). Making meaning across textscales: A critical statement on reading comprehension. *RAND Reading Study Group*.
- León, I., & Divasson, L. (2006). Nominal domains in the biomedical research paper: A grammatico-rhetorical study of modification. En M. Gotti, & F. Salager-Meyer, *Advances in medical discourse analysis: Oral and written contexts* (pp. 289-309). Bern: Peter Lang.
- Lesko, N. (2001). *Act your age! A cultural construction of adolescence*. New York, NY: Routledge Falmer.
- Lévinas, E. (1982). *Ética e infinito*. Madrid: Visor.

- Lewin, B. (2005). Hedging: an exploratory study of authors' and readers' identification of 'toning down' in scientific texts. *Journal of English for Academic Purposes*, 4(2), 163-178.
- Lillis, T. (2001). *Student writing: Access, regulation, desire*. London, U.K.: Routledge.
- Llamas Saíz, C. (2014). Escribir en el contexto académico: conocimientos y estrategias. En I. Ballano, & I. Muñoz (Eds.), *Escribir en el contexto académico* (pp. 13-33). Publicaciones de la Universidad de Deusto, Serie Mayor.
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36-62.
- Lunsford, A. (2006). Writing, technologies, and the fifth canon. *Computers and Composition*, 23, 169-177.
- Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. New York, NY: Palgrave Macmillan.
- Markkanen, R., & Schröder, H. (Eds.). (1997). *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts*. Berlin: Walter de Gruyter.
- Márquez, C., Izquierdo, M., & Espinet, M. (2003). Comunicación multimodal en la clase de ciencias: el ciclo del agua. *Enseñanza de las Ciencias*, 21(3), 371-386.
- Martin, J. (2000). Beyond exchange: Appraisal systems in English. En *Evaluation in text* (pp. 142-175). Oxford: Oxford University Press.
- Martin, J. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*(20), 10-21.
- Martin, J. R. (1984). Language, register and genre. En F. Christie, *Children writing* (pp. 21-29). Geelong, Australia: Deakin University Press.
- Martin, J., & Rose, D. (2007). *Working with discourse. Meaning beyond the clause*. New York, NY: Continuum.
- Martín, P., Rey-Rocha, J., Burgess, S., & Moreno, A. I. (2014). Publishing research in English-language journals: Attitudes, strategies and difficulties of multilingual scholars of medicine. *Journal of English for Academic Purposes*, 16, 57-67.
- Martínez Lirola, M. (2007). *Aspectos esenciales de la Gramática Sistémica Funcional*. Alicante: Servicio de Publicaciones de la Universidad.

- Martínez, I. A. (2005). Native and non-native writers' use of first person pronouns in the different sections of biology research articles in English. *Journal of Second Language Writing*, 14(3), 174-190.
- Matthiessen, C. (2009). Léxico-gramática y colocación léxica: Un estudio sistémico-funcional. *Signos*, 42(71), 333-383.
- Mauranen, A. (1993). Cultural differences in academic discourse: Problems of a linguistic and cultural minority. En L. Löfman, L. Kurki-Suonio, S. Pellinen, & J. Lehtonen (Eds.), *The competent intercultural communicator* (Vol. 51, pp. 157-174). Tampere, Finland: AFinLA Series.
- McCarthy, P. M., & Jarvis, S. (2010). MTLT, VocD-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392 doi: 10.3758/BRM.42.2.381.
- McEney, T., & Hardie, A. (2011). *Corpus linguistics: method, theory and practice*. Cambridge, UK: Cambridge University Press.
- McEney, T., & Kifle, N. (2002). Epistemic modality in argumentative essays of second-language writers. En J. Flowerdew (Ed.), *Academic discourse* (pp. 182-195). London, UK: Longman, Pearson Education.
- McEney, T., & Wilson, A. (1996). *Corpus linguistics: An introduction* (2nd ed.). Edinburgh, UK: Edinburgh University Press.
- McEney, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. New York, NY: Routledge.
- McGrath, L., & Kuteeva, M. (2012). Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. *English for Specific Purposes*, 31(3), 161-173.
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, 27(1), 57-86.
- Mendiluce, M. (2005). La matización asertiva en el artículo biomédico: una propuesta de clasificación para los estudios contrastivos inglés-español. *Ibérica, Journal of the European Association of Languages for Specific Purposes*, 10, 63-90.
- Meyer, C. (2002). *English Corpus Linguistics: An introduction*. Cambridge, UK: Cambridge University Press.
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151-167.

- Milton, J. (1998). Exploiting L1 and interlanguage corpora in the design of an electronic language learning and production environment. En S. Granger (Ed.), *Learner English on computer* (pp. 186-198). London, UK: Longman.
- Mirzapour, F., & Ahmadi, M. (2011). Study on lexical cohesion in English and Persian research articles. *English Language teaching*, 4(4), 243-255.
- Moje, E. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of adolescent & adult literacy*, 96-107.
- Moje, E., Overby, M., Tysvaer, N., & Morris, K. (2008). The complex world of adolescent literacy: Myths, Motivations and Mysteries. *Harvard Educational Review*, 78(1), 107-155.
- Montague, N. (1995). The process oriented approach to teaching writing to second language learners. *New York State Association for Bilingual Education Journal*, 10, 13-24.
- Montolío, E. (2001). *Conectores de la lengua escrita*. Barcelona: Ariel.
- Muncie, I. (2006). Discourse analysis. En V. Jupp, *The SAGE dictionary of social research methods*. London, UK: SAGE Publications.
- Mur, P. (2007). 'I/we focus only': A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes*, 6(2), 143-162.
- Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, 10(1), 1-35.
- Neff, J., & Prieto, R. (1994). *First language influence on Spanish EFL university writing development*. Washington D.C.: Educational Resource Information Center.
- Nelson, N., & Castelló, M. (2012). Academic writing and authorial voice. En M. Castelló, & C. Donahue (Eds.), *University writing. Selves and texts in academic societies* (pp. 33-52). Bingley, U.K.: Emerald.
- Neumann, S. (2014). *Contrastive register variation: A quantitative approach to the comparison of English and German*. Berlin: De Gruyter.
- Newell, G. (2006). Writing to learn: How alternative theories of school writing account for student performance. En C. MacArthur, S. Graham, & J. Fitzgerrald (Eds.), *The handbook of writing research* (pp. 235-247). New York, NY: Guilford Press.

- Nippold, M. A. (2004). Research on later language development: International perspectives. En R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 1-8). Amsterdam: Benjamins.
- Norris, S., & Phillips, L. (2009). Scientific literacy. En D. Olson, & N. Torrance, *The Cambridge handbook of literacy* (pp. 271-285). Cambridge, UK: Cambridge University Press.
- Nuyts, J. (2005). Brothers in arms? On the relation between cognitive and functional linguistics. En F. Ruiz de Mendoza Ibáñez, & S. Peña Cervel (Eds.), *Cognitive linguistics: Internal dynamics and interdisciplinary interaction (Cognitive Linguistics Research)* (pp. 69-100). New York, NY: Mouton de Gruyter.
- Ochs, E., & Schieffelin, B. (1989). Language as a heart. *Text*, 9, 7-25.
- Oliver del Olmo, S. (2004). Análisis contrastivo Español/ Inglés de la atenuación retórica en el discurso médico. EL artículo de investigación y el caso clínico (Doctoral dissertation). Retrieved from <http://dialnet.unirioja.es/servlet/tesis?codigo=6271>.
- Olson, D. (2009a). Language, literacy and mind: The literacy hypothesis. *Psyke*, 18(1), 3-9.
- Olson, D. R. (2009b). Literacy, literacy policy, and the school. En D. R. Olson, & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 566-576). Cambridge: Cambridge University Press.
- Olson, D. R., & Torrance, N. (Eds.). (2009). *The Cambridge Handbook of Literacy*. Cambridge University Press.
- Ong, W. (1982). *Orality and literacy*. London, U.K.: Methuen & Co. Ltd.
- Orlikowski, W., & Yates, J. (1994). Genre repertoire: The structuring of communicative practices in organizations. *Administrative Science Quarterly*, 39, 541-574.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492-518.
- Ouaknin, M. (1994). *Bibliothérapie. Lire, c'est guérir*. Editions du Seuil.
- Padró, L., & Stanilovsky, E. (2012). FreeLing 3.0.: Towards wider Multilinguality. *Proceedings of the Language Resources and Evaluation Conference (LREC)*. Istanbul, Turkey.
- Paltridge, B. (2012). *Discourse analysis: An introduction*. (B. Paltridge, Ed.) London, UK: Bloomsbury.

- Payrató, L. (1997). *De professió, lingüista: Panorama de la lingüística aplicada*. Barcelona: Empúries.
- Pérez-Llantada, C. (2012). *Scientific discourse and the rhetoric of globalization: The impact of culture and language*. London, UK: Continuum.
- Perry, K. (2012). What is Literacy? – A critical overview of sociocultural perspectives. *Journal of Language and Literacy Education*, 8(1), 51-71.
- Phillips, N., & Hardy, C. (Eds.). (2002). *Discourse analysis: Investigating processes of social construction*. London, UK: Sage Publications.
- Pinker, S. (2014). *The Sense of style: The thinking person's guide to writing in the 21st century*. New York, NY: Penguin.
- Porter, J. (1986). Intertextuality and the discourse community. *Rhetoric Review*, 5(1), 34-47.
- Prior, P. (1998). *Writing /disciplinarity: A sociohistoric account of literate activity in the academy*. Mahwah, NJ: Erlbaum.
- Prior, P. (2003). Are communities of practice really an alternative to discourse communities? *Paper presented at the American Association of Applied Linguistics*. Washington, 22-25 March. Extraído de <http://www.researchgate.net/publication/228784036>.
- Prior, P. (2006). A sociocultural theory of writing. En C. McArthur, S. Graham, & J. Fitzgerald, *Handbook of writing research* (pp. 54-66). New York, NY: Guilford Press.
- Prior, P. (2009). From speech genres to mediated multimodal genre system: Bakhtin, Voloshinov, and the question of writing. En C. Bazerman, A. Bonini, & D. Figueiredo (Eds.), *Genres in a changing world* (pp. 17-34). Fort Collins, CO: WAC Clearinghouse.
- Proverbio, A., & Zani, A. (2005). Development changes in the linguistic brain after puberty. *Trends in Cognitive Science*, 9, 164-167.
- Pujol Dahme, A. (2014). Llegim imatges: La gramàtica visual als treballs de recerca de batxillerat. (L. Tolchinsky, & M. Fons, Eds.) *Articles de Didàctica de la Llengua i la Literatura*.(62), 45-53.

- Pujol Dahme, A. (2015). La alfabetización académica desde una perspectiva multidimensional. *Fuentes. Revista de la Facultad de Ciencias de la Educación de la Universidad de Sevilla*, 17, 67-87.
- Ravid, D., & Berman, R. A. (2009). Developing linguistic register across text types. The case of modern Hebrew. *Pragmatics and Cognition*, 17(1), 108-145.
- Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language*, 29(2), 417-447.
- Recchio, T. (1991). A Bakhtinian reading of student writing. *College Composition and Communication*, 42, 446-454.
- Rienecker, L., & Stray Jörgensen, P. (2003). The (Im)possibilities in teaching university writing in the Anglo-American tradition when dealing with Continental student writers. En L. Björk, G. Bräuer, L. Rienecker, & P. Stray Jörgensen (Eds.), *Teaching academic writing in European higher education* (Vol. 12, pp. 101-112). Dordrecht, Netherlands: Kluwer Academic.
- Rijlaarsdam, G., & van den Bergh, H. (2006). Writing process theory: A functional dynamic approach. En C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 41-53). New York, NY: Guilford Press.
- Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssen, T., Kieft, M., Broekkamp, H., y otros. (2005). Psychology and the teaching of writing in 8000 and some words. *BJEP Monographs Series II*, 3, 127-153.
- Rodríguez, J. (2012). Marc-Alain Ouaknin: Un rabino entre la filosofía francesa contemporánea. En L. Llevadot, & J. Riba (Eds.), *Filosofías postmetafísicas. 20 años de filosofía francesa contemporánea* (pp. 179-197). Barcelona: UOC.
- Russell, D. R. (1995). Activity theory and its implications for writing instruction. En J. Petraglia (Ed.), *Reconceiving writing, rethinking writing instruction* (pp. 51-78). Hillsdale, NJ: Lawrence Erlbaum.
- Russell, D. R., & Cortes, V. (2012). Academic and scientific texts: the same or different communities? En M. Castelló, & C. Donahue (Eds.), *University writing. Selves and texts in academic societies* (pp. 3-18). Bingley, UK: Emerald.
- Russell, D., Lea, M., Parker, J., Street, B., & Donahue, C. (2009). Exploring notions of genre in 'academic literacies' and 'writing across the curriculum': approaches across countries and contexts. En C. Bazerman, A. Bonini, & D. Figueiredo (Eds.), *Genre*

- in a Changing World. Perspectives on Writing* (pp. 459-491). Colorado: WAC Clearinghouse/Parlor Press.
- Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes*, 13(2), 149-170.
- Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education*, 92(3), 447-472.
- Sandoval, W. A., & Reiser, B. J. (2004). Explanation driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, 88(3), 345-372.
- Sandoval, W., & Millwood, K. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction*, 23(1), 23-55.
- Scarpa, F. (2007). Using an Italian diachronic corpus to investigate the "core" patterns of the language of science. En K. Ahmad, & M. Rogers (Eds.), *Evidence-based LSP. Translation, text and terminology* (Vol. 47, pp. 75-95). Bern, Switzerland: Peter Lang.
- Schiffrin, D., Tannen, D., & Hamilton, H. (Eds.). (2001). *The handbook of discourse analysis*. Malden, MA: Blackwell Publishers.
- Schlaggar, B., Brown, T., Lugar, H., Visscher, K., Miezin, F., & Petersen, S. (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science*, 296, 13476-1479.
- Schleppegrell, M. (2000). How SFL can inform writing instruction: The grammar of expository essays. *Revista Canaria de Estudios Ingleses*(40), 171-188.
- Schleppegrell, M. (2001). The linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431-459.
- Schleppegrell, M., & Colombi, M. (2002). Theory and practice in the development of advanced literacy. En M. Schleppegrell, & M. Colombi (Eds.), *Developing advanced literacy in first and second languages: meaning with power* (pp. 1-19). Mahwah, NJ: Lawrence Erlbaum.
- Schmandt-Besserat, D., & Erard, M. (2008). Origins and forms of writing. En C. Bazerman (Ed.), *Handbook of research on writing: History, society, school, individual, text* (pp. 7-22). New York, NY: Roudledge.

- Siegel, H. (1992). On defining "critical thinker" and justifying critical thinking. En H. Alexander (Ed.), *Philosophy of education* (pp. 72-75). Urbana, IL: Philosophy of Education Society.
- Simpson, J. (2007). *Burning to read: English fundamentalism*. Cambridge, M.A.: Harvard University Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.
- Sinclair, J. (2000). Lexical grammar. *Naujoji Metodologija*, 24, 191–203.
- Sinclair, J. (2004). Design and management of corpora. Workshop. University of Birmingham.
- Slobin, D. (1996). Two ways to travel, verbs of motion in English and Spanish. En S. Masayoshi, & S. Thompson, *Grammatical constructions. Their form and meaning* (pp. 195-219). Oxford: Oxford University Press.
- Smith, F. (1990). *Para darle sentido a la lectura*. Madrid: Aprendizaje Visor.
- Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. En D. R. Olson, & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge: Cambridge University Press.
- Snow, C., & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* New York, NY: Carnegie Corporation.
- Solé, I. (1992). *Estrategias de lectura*. Barcelona: Graó/ICE.
- Solé, I. (2014). Alfabetización académica y lectura. En I. Ballano, & I. Muñoz (Eds.), *Escribir en el contexto académico* (pp. 35-51). Publicaciones de la Universidad de Deusto. Serie Mayor.
- Stella, E. (2004). *Translated texts: Properties, variants, evaluations*. Frankfurt am Main: Peter Lang.
- Street, B. (1984). *Literacy in theory and practice*. New York, NY: Cambridge University Press.
- Street, B. (2012). *Literacy and Multimodality*. STIS Lecture: Inter-Disciplinary Seminars o Laboratório SEMIOTEC da FALE/UFM. Belo Horizonte: Faculdade de Letras.
- Strömquist, S., Johansson, V., Kriz, S., Ragnarsdóttir, H., Aisenman, R., & Ravid, D. (2002). Toward a cross-linguistic comparison of lexical quantity in speech and writing. *Written Language and Literacy*, 5(1), 45-67.

- Stubbs, M. (1993). British traditions in text analysis: From Firth to Sinclair. En M. Baker, F. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 1-46). Amsterdam, Netherlands: John Benjamins.
- Stubbs, M. (1996). *Text and corpus analysis: computer-assisted studies of language and culture*. Oxford, UK: Blackwell.
- Swales, J. M. (1990). *Genre analysis: English in academic and research and research settings*. New York, NY: Cambridge University Press.
- Swales, J. M. (2004). *Research genres: Explorations and applications*. New York, NY: Cambridge University Press.
- Szmrecsanyi, B. (2004). On operationalizing syntactic complexity. En G. Purnelle, C. Fairon, & A. Dister (Ed.), *Le poids des mots. Proceeding of the 7th International Conference on Textual Data Statistical Analysis*.2, pp. 1032-1039. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Talmy, L. (1991). Path to realization, a typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Pasasession on The Grammar of Event Structure*, (pp. 480-519).
- Tardy, C. (2012). Current conceptions of voice. En K. Hyland, & C. Sancho-Guinda (Eds.), *Stance and voice in written academic genres* (pp. 34-50). New York, NY: Palgrave MacMillan.
- Teberosky, A. (2007). El texto académico. En M. Castelló (Ed.), *Escribir y comunicarse en contextos científicos y académicos. Conocimientos y estrategias* (pp. 17-45). Barcelona: Graó.
- Teberosky, A., & Tolchinsky, L. (1994). Introducción: Acerca de la interpretación. (A. Teberosky, & L. Tolchinsky, Eds.) *Substratum. Temas fundamentales en psicología y educación*, 2(4), 9-22.
- Teubal, E. D. (2007). *Notational knowledge developmental and historical perspectives*. Netherlands, Rotterdam: Sense.
- Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10(1), 1-13.
- Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics*, 22(1), 58-78.

- Thompson, G., & Hunston, S. (2000). Evaluation: an introduction. En G. Thompson, & S. Hunston (Eds.), *Evaluation in text*. Oxford: Oxford University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam, Netherlands: John Benjamins.
- Tolchinsky, L. (2003). *The cradle of culture. What children know about writing and numbers before being taught*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tolchinsky, L. (2014). L'ús de corpus lingüístics com a eina pedagògica. (L. Tolchinsky, & M. Fons, Eds.) *Articles de Didàctica de la Llengua i de la Literatura*.(62), 9-17.
- Tolchinsky, L., & Martí, M. (2013). La lingüística de corpus, què és i quina utilitat educativa té. L'explotació didàctica de corpus lingüístics. (L. Tolchinsky, Ed.) *Quaderns per a l'anàlisi*, 7-22. Horsori.
- Tolchinsky, L., & Simó, R. (2001). Escribir y leer a través del currículum. En I.C.E, *Cuadernos de educación* (pp. 159-165). Barcelona: Horsori.
- Tolchinsky, L., Castrechini, A., Cañedo, M., Estrada, F., Gracia, M. C., Navarro, M., et al. (2013). *La escritura académica a través de las disciplinas*. Barcelona: Octaedro.
- Torrado M., & Berlanga, V. (2013). Análisis Discriminante mediante SPSS. *REIRE Revista d' Innovació i Recerca en Educació*, 6(2),150-166.
- Toulmin, S. (1958). *The uses of argument*. New York, NY: Cambridge University Press.
- Tribble, C. (2002). Corpora and corpus analysis: New windows on academic writing. En J. Flowerdew (Ed.), *Academic discourse* (pp. 131-149). London: Longman, Pearson Education.
- Tusón, J. (2006). *Lletres sobre Lletres*. Barcelona: Grup Editorial 62.
- Uccelli, P., Dobbs, C., & Scott, J. (2013). Mastering academic language: Organization and stance in the persuasive writing of high school students. *Written Communication*, 30(1), 36-62.
- Upton, T., & Cohen, M. (2009). An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies*, 11(5), 585-605.
- Van Dijk, T. (Ed.). (1997). *Discourse as structure and process. Volume 1*. London, UK: SAGE Publications.
- Vande Kopple, W. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication*, 26(1), 82-93.

- Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes*, 20, 83-102.
- Vázquez, G., Fernández, A., & Martí, M. A. (2000). Dealing with lexical semantic mismatches between English and Spanish. *Proceedings of the International Conference of Knowledge Based Computer Systems*, (pp. 308-319). Mumbai, India.
- Vázquez, I. (2010). A contrastive analysis of the use of modal verbs in the expression of epistemic stance in Business Management research articles in English and Spanish. *Ibérica*, 19, 77-96.
- Vázquez, I., & Giner, D. (2008). Beyond mood and modality: Epistemic modality markers as hedges in research articles. A cross-disciplinary study. *Revista Alicantina de Estudios Ingleses*, 21, 171-190.
- Veel, R. (1997). Learning how to mean – scientifically speaking: apprenticeship into scientific discourse in the secondary school. En F. Christie, & J. Martin (Eds.), *Genre and institutions. Social processes in the workplace and school* (pp. 161-195). London: Cassell.
- Veel, R. (1997). Learning how to mean_scientifically speaking: apprenticeship into scientific discourse in the secondary school. En F. Christie, & J. R. Martin, *Genre and Institutions: Social Processes in the Workplace and School* (pp. 161-195). London, U.K.: Cassell.
- Ventola, E. (1996). Packing and unpacking of information in academic. En E. Ventola, & A. Mauranen, *Academic Writing. Intercultural and textual issues* (pp. 153-194). Philadelphia, PA: Benjamins.
- Vigotsky, L. (1978). *Pensamiento y lenguaje*. Buenos Aires, Argentina: La Pleyade.
- Villayandre, M. (2008). Lingüística con corpus. *E.H. Filología*, 329-349 Recuperado de <http://mobile.buleria.unileon.es/bitstream/handle/10612/2921/Ling%C3%BC%C3%ADstica%20p.%20239-349.pdf?sequence=1>.
- Volosinov, V. (1986). *Marxism and the philosophy of language*. Cambridge, MA: Harvard University Press.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge, UK: Cambridge University Press.
- Wertsch, J. (1993). *Voces de la mente. Un enfoque sociocultural para el estudio de la acción mediada*. Madrid: Visor.

- White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16, 3-118.
- Whitehead, T. (2005). Basic classical ethnographic research methods. *Ethnographically Informed Community and cultural Assessment Research Systems (EICCARS) working paper series*. Recuperado de <http://www.cusag.umd.edu/documents/workingpapers/classicalethnomethods.pdf>.
- Widdowson, H. (1998). Context, community and authentic language. *TESOL Quarterly*, 32(4), 705-716.
- Williams, J. (1981). *Style: ten lessons in clarity and grace*. Chicago, IL: University of Chicago Press.
- Wineburg, S., & Martin, D. (2004). Reading and rewriting history. *Educational Readership*, 62(1), 42-45.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.
- Woerfel, T., & Yilmaz, S. (2011). Lexical Development of German-Turkish bilinguals: of German-Turkish bilinguals: A comparative study in written discourse. En C. Cummins (Ed.), *Proceedings of the 6th Cambridge Postgraduate Conference in Language Research* (pp. 240-251). Cambridge: Cambridge Institute of Language Research.
- Woodward-Kron, R. (2005). The role of genre and embedded genres in tertiary students' writing. *Prospect*, 20(3), 24-41.
- Xanthos, A. (2014). Textable: Programmation visuelle pour l'analyse de données textuelles. *Actes des 12èmes Journées Internationales d'Analyse Statistique de données textuelles (JADT)*.
- Zorraquino, M., & Portolés, J. (1999). Los marcadores del discurso. En I. Bosque, & V. Demonte (Eds.), *Gramática descriptiva de la lengua española (3)* (pp. 4051-4213). Madrid: Espasa-Calpe.

Anexo 1

Autorizaciones de los coordinadores para hacer uso de los trabajos de máster universitarios.



Dos campus d'excel·lència internacional:



Màster Oficial Interuniversitari en Immunologia
Facultat de Biologia
Universitat de Barcelona
Av. Diagonal 643, 3ª planta
Barcelona 08028
Espanya

Barcelona, 27 de abril de 2012

A QUIEN CORRESPONDA:

Como coordinador del Máster Interuniversitario en Inmunología de la Universitat de Barcelona y de la Universitat Autònoma de Barcelona, me dirijo a ti como director de algún/os de los Trabajos Final de Máster que se han llevado en tu grupo desde el curso 2006-2007 hasta el curso 2010-2011.

Hace unos meses una profesora de la Facultad de Filología, la Dra. Liliana Tolchinsky del Departamento de Lingüística General se dirigió a mí para el desarrollo de una de una tesis de una alumna suya, Ana Pujol, poder acceder a los trabajos de final de máster (TFM) de inmunología desde el curso 2006-7 hasta el 2010-11 y realizar un análisis lingüístico. Como coordinador he recabado la autorización de la Comisión del Máster que ha sido positiva y por lo tanto se ha autorizado a iniciar a la Sra. Ana Pujol el contacto con los directores de cada uno de los 128 trabajos que tenemos en el archivo para que pida el permiso específico que se adjunta a esta carta que te envío.

Por lo tanto, te rogamos atiendas la petición que se te hace pero siempre teniendo en cuenta que es una decisión que no te obliga a nada y es absolutamente voluntaria. Que conste que la carta adjunta ha sido aprobada por la Comisión del Máster, intentando preservar la confidencialidad y el buen uso de los documentos de los TFM.

Atentamente.

Dr. Jorge Lloberas

Coordinador del Máster

Em dirigeixo a vostè per a sol·licitar la seva autorització per a consultar els treballs de fi del màster d'estudis històrics (UB).

Aquesta consulta ve motivada pel meu treball de doctorat en Lingüística Teòrica, Computacional i Aplicada (UB) sota la direcció de la Dra. Liliana Tolchinsky, coordinadora del programa de doctorat.

L'objectiu general d'aquest treball és descriure com es desenvolupa la competència en la redacció científica, al llarg de l'educació postobligatòria i fins a la publicació científica, en diferents àmbits de coneixement i en 3 nivells acadèmics: treball de recerca de batxillerat, treball fi de màster i articles publicats en revistes indexades.

Es contrastaran els estils retòrics d'especialitat, dins de l'àmbit de ciències i d'humanitats, analitzant aspectes lingüístics, com ara, l'ús de marcadors discursius, així com diferents apartats del text: marc teòric, formulació d'hipòtesi, anàlisi i discussió de resultats.

Em comprometo a:

Utilitzar aquests treballs amb una finalitat estrictament científica, citar les fonts corresponents i respectar la propietat intel·lectual dels treballs.

És per això que,

SOL·LICITO: autorització per a consultar els treballs de recerca de fi del màster d'estudis històrics.

Nom del coordinador: Dr. Llorenç Ferrer

Autoritzo



Signatura:

Màster d'Estudis Històrics

Cordialment,

Ana Pujol Dahme - Estudiant de Doctorat en Lingüística

Liliana Tolchinsky - Directora de la tesi

Barcelona, 28 de febrer de 2012

Ejemplo de las autorizaciones que se les envió a cada uno de los tutores de máster de inmunología



Em dirigeixo a vostè per a sol·licitar la seva autorització per a consultar els treballs de fi de màster d'immunologia de la Facultat de Biologia (UB).

Aquesta consulta ve motivada pel meu treball de doctorat en Lingüística Teòrica, Computacional i Aplicada (UB) sota la direcció de la Dra. Liliana Tolchinsky, coordinadora del programa de doctorat.

L'objectiu general d'aquest treball és descriure com es desenvolupa la competència en la redacció científica, al llarg de l'educació postobligatòria i fins a la publicació científica, en diferents àmbits de coneixement i en 3 nivells acadèmics: treball de recerca de batxillerat, treball fi de màster i articles publicats en revistes indexades. Es contrastaran els estils retòrics d'especialitat, dins de l'àmbit de ciències i d'humanitats, analitzant aspectes lingüístics, com ara, l'ús de marcadors discursius, així com diferents apartats del text: marc teòric, formulació d'hipòtesi, anàlisi i discussió de resultats.

* **Condicion de la consulta:** la informació científica recollida en el Treballs de fi de Màster d'Immunologia és confidencial, no es pot copiar, ni utilitzar en qualsevol medi electrònic en cap sentit que l' estrictament dirigit a la seva anàlisi lingüística.

És per això que,

SOL·LICITO: autorització per a consultar els treballs de recerca de fi de Màster d'Immunologia, el títol dels quals figura a continuació (*VEURE LES CONDICIONS EN QUE ES FARÀ LA CONSULTA DE LES MEMÒRIES DELS TREBALLS DE MÀSTER):

Títol del treball:

Genotipatge del gen Trex2 a tumors de cap i coll (2009-2010)

Nom del professor/tutor: Dra. Soler

Autoritzo



Signatura: _____

Cordialment,

Ana Pujol Dahme - Estudiant de Doctorat en Lingüística

Liliana Tolchinsky - Directora de la tesi

Barcelona, 01 d'octubre de 2012

Anexo 2

Títulos de los trabajos de máster en inmunología y de las publicaciones en revistas de impacto, sub-corpus TARBUC para el estudio 2.

List of master's theses in the discipline of immunology

1. Efecte d'un nou probiòtic en un model d'asma al·lèrgica en el ratolí
2. Estudi de la resposta al superantigen SEB a nivell de citocines i fosforilació d'STATs en un model de shock sèptic murí.
3. Estudi dels efectes coestimuladors de la interacció CD26-ADA en l'expressió de marcadors d'activació de la cèl·lula T
4. Estudi d'una doble estratègia per a induir i expandir el repertori de cèl·lules T basada en l'administració d'hormona de creixement i vacuna en pacients amb infecció pel VIH
5. Disminució dels nivells de TCR en la membrana dels limfòcits T CD4+ humans per efecte de sindecàn-2
6. Generació de cèl·lules dendrítiques tolerogèniques per terapia cel·lular en esclerosi múltiple
7. Modificació de l'expressió de CD36 i CD206 per lligands de receptors toll-like (TLRs) i citocines
8. Estudi de la influència dels pèptids naturals presentats per MHC de classe II en autoimmunitat tiroidea

List of published articles

1. Activity of the cyclooxygenase 2-prostaglandin-E prostanoid receptor pathway in mice exposed to house dust mite aeroallergens, and impact of exogenous prostaglandin E2
2. Increased Susceptibility to Skin Carcinogenesis in TREX2 Knockout Mice
3. Increased α -Defensins 1-3 Production by Dendritic Cells in HIV-Infected Individuals Is Associated with Slower Disease Progression
4. The reconstitution of the thymus in immunosuppressed individuals restores CD4-specific cellular and humoral immune responses
5. Syndecan-2 can promote clearance of T-cell receptor/CD3 from the cell surface
6. Specific T-cell proliferation to myelin peptides in relapsing-remitting multiple sclerosis
7. Functional consequences of CD36 downregulation by TLR signals
8. The peptide-binding motif of HLA-DR8 shares important structural features with other type 1 diabetes-associated alleles