

Universidad Autónoma de Querétaro
Facultad de Lenguas y Letras
Doctorado en Lingüística

El léxico en el dominio de la escritura académica

Tesis

Que como parte de los requisitos para obtener el grado de
Doctor en Lingüística

Presenta

Eva Margarita Godínez López

Dirigida por:

Dra. Luisa Josefina Alarcón Neve

Dra. Luisa Josefina Alarcón Neve
Presidente

Dra. Karina Hess Zimmermann
Secretario

Dra. Eva Velásquez Upegui
Vocal

Dra. Adriana Bolívar
Suplente

Dra. Rosa Margarita Galán Vélez
Suplente

Centro Universitario, Querétaro, Qro.
Septiembre de 2019
México

CONTENIDO

Agradecimientos	1
Resumen	3
Abstract	4
Índice de tablas	5
Índice de figuras	6
1. Introducción: El reto de la escritura académica.....	8
2. Antecedentes.....	14
2.1.Estudios sobre evaluación de la calidad de la escritura académica.....	15
2.2.Estudios enfocados al dominio léxico en la escritura.....	25
3. Marco teórico: Criterios para orientar la investigación	35
3.1.Primer criterio: Social	35
3.2.Segundo criterio: Epistemológico e interactivo	44
3.3.Tercer criterio: Lingüístico.....	48
4. Planteamiento del problema y objetivos de investigación.....	60
4.1.El problema.....	61
4.2.Delimitación y alcances prácticos de la tesis	63
4.3.Objetivos del estudio.....	64
4.4.Preguntas de investigación.....	65
4.5.Hipótesis	66
5. Metodología.....	68
5.1.Método y tipo de investigación.....	68
5.2.El corpus y la muestra.....	69
5.3.Variables del estudio	75
5.4.Procedimientos del análisis I: Indicadores léxicos	98
5.5.Procedimientos del análisis II: Sustantivos abstractos.....	99
5.6.Procedimientos del análisis III: Léxico especializado	101
5.7.Procedimientos del análisis IV: Léxico abstracto especializado.....	102

6. Resultados.....	105
6.1. Análisis I. Indicadores léxicos	105
6.2. Análisis II. Sustantivos abstractos	114
6.3. Análisis III. Léxico especializado.....	128
6.4. Análisis IV. Léxico abstracto especializado.....	137
7. Discusión de los resultados.....	150
8. Conclusiones.....	166
9. Bibliografía.....	171
10. Anexos	197
A. Resumen de los Antecedentes de investigación.....	197
B. Indicadores de literacidad según organismos internacionales y nacionales	201
C. Pauta de escritura de los textos	206
D. Rúbrica original y versiones adaptadas.....	207
E. Perfiles de los jueces.....	210
F. Postulados de Bioética de la Universidad Autónoma de Querétaro.....	211
G. Resumen de la propuesta de categorización de los sustantivos	213
H. Resumen de criterios de inclusión, exclusión y conteo de sustantivos.....	213

Agradecimientos

Esta investigación no habría sido posible sin el apoyo constante, la paciencia y la crítica de mi directora de tesis, la Dra. Luisa Josefina Alarcón Neve, así como las invaluable contribuciones de las profesoras miembros de mi Comité Doctoral, la Dra. Karina Hess Zimmermann, la Dra. Eva Velásquez Upegui, la Dra. Adriana Bolívar y la Dra. Rosa Margarita Galán Vélez, que con sus autorizados comentarios reorientaron esta investigación y su informe escrito más de una vez.

Asimismo, agradezco el apoyo económico e institucional de la beca para posgrado otorgada por el Consejo Nacional de Ciencia y Tecnología (CONACYT), fundamental para el logro de las metas de un posgrado de calidad con alto nivel de exigencia académica como este, así como la valiosa ayuda de las personas que estuvieron a cargo de la Coordinación del programa y de otras áreas de la Facultad, desde las cuales los estudiantes y becarios recibimos orientaciones cruciales para transitar exitosamente por el doctorado.

Ofrezco mi reconocimiento a aquellos profesores que generosamente contribuyeron a mi formación, especialmente al Dr. Ignacio Rodríguez Sánchez y la Dra. Haydée Carrasco; también a mi tutor, el Dr. Marco Ángel, y a los docentes de bachillerato que desinteresadamente donaron su tiempo y experiencia fungiendo como jueces de los textos de mi corpus y me ayudaron a afinar instrumentos importantes para este trabajo durante las fases de piloteo y aplicación de los mismos. También agradezco a mis compañeros de generación, que fueron un referente para sostener mis esfuerzos; a los compañeros que me precedieron, cuyo ejemplo me guió, y a los nuevos estudiantes con quienes tuve la fortuna de compartir las aulas y cuya juventud me sirvió de crítico espejo y de acicate.

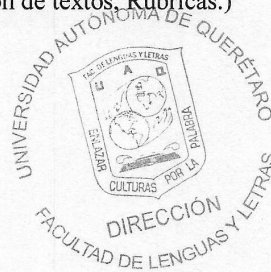
Gracias a mi familia; gracias eternamente a mi madre por haber confiado siempre en mí, hasta el último momento y con tanta fe. Gracias a todos mis hermanos, por sostener una red invisible que nunca me ha fallado y a cuyo esfuerzo y amor espero retribuir dignamente. Gracias a los que ya se fueron y me esperan. Gracias de manera especial a Marco Antonio Rivera Treviño, mi compañero en esta vida, por *tanto* y por *todo*.

Resumen

Resumen

Es común que en la escuela se solicite a los estudiantes elaborar textos con *un lenguaje académico variado*, pero sus escritos, a menudo, reciben una evaluación subjetiva. Por ello es necesario explorar qué características del lenguaje en los textos pueden construir objetivamente ese indicador de la escritura académica, para contribuir a la planificación de su enseñanza y a su evaluación. Con ese fin, se midió la densidad y variedad léxica y nominal de un corpus de 120 textos producidos por estudiantes de bachillerato; se analizó el léxico abstracto y especializado a través de la categorización de los sustantivos bajo criterios semánticos, y se modeló estadísticamente el efecto de estos indicadores sobre un promedio de calificaciones otorgadas por un grupo de jueces expertos con base en una rúbrica. El método fue cuantitativo y correlacional, y se emplearon herramientas de análisis léxico automatizado y un programa estadístico avanzado para realizar pruebas y gráficos. La hipótesis general fue que los textos mejor evaluados –y por tanto aquellos que los jueces consideraron evidencia de un mayor dominio de la escritura académica por parte de los estudiantes– serían aquellos que emplearan un léxico “más académico”, es decir, más denso, variado, abstracto y especializado, en concordancia con caracterizaciones previas del discurso académico. Los resultados muestran que los textos mejor evaluados presentan un balance entre su extensión y la diversidad léxica; aunque la densidad léxica y nominal parecen tener más peso que la variedad de vocabulario. Asimismo, se encontraron diferencias significativas en los textos mejor evaluados en el repertorio y uso de sustantivos abstractos y de léxico especializado, y del cruce de ambas categorías. De este modo se pudo verificar que estas características del léxico son, en efecto, predictores útiles de dominio de la escritura académica, que por tanto merecen ser objeto de enseñanza sistemática.

(Palabras clave: Discurso académico, Escritura académica, Léxico, Evaluación de textos, Rúbricas.)



Abstract

Abstract

In the school context, it is common for students to be asked to produce texts with *a varied academic language*, but their writing is often subjectively evaluated. For this reason, it is necessary to explore which language features in the texts can objectively construct that indicator of academic writing, in order to contribute to the planning of their teaching and their evaluation. To that end, lexical and nominal density and variety on a corpus of 120 texts produced by the same number of Spanish-speaking High School students were measured; also, the abstract and specialized lexicon were analyzed through the categorization of nouns under semantic criteria, and the effect of these indicators was statistically modeled over an average rate provided by a set of judges, based on a score rubric. Methods were quantitative and correlational, and automated lexical analysis tools and an advanced statistical program were used to perform tests and to do graphs. The general hypothesis was that the best scored texts –and therefore, those that the judges considered evidence of a greater mastery of academic writing– would be those that used lexicon in a “more academic” way, that is, dense, varied, abstract and specialized lexicon, in accordance with previous characterizations of academic discourse. The results show that the best scored texts present a balance between their extension and their lexical diversity; although lexical and nominal density seem to have more influence than the variety of vocabulary. Likewise, significant differences were found between low and high scored texts with regard to the repertoire and use of abstract nouns and specialized lexicon, and the crossing of both categories. In conclusion, it was possible to verify that these characteristics of the lexicon are, in fact, useful predictors of academic writing, which therefore deserve to be systematically taught.

(**Key words:** Academic Discourse, Academic Writing, Lexicon, Text Evaluation, Score Rubrics.)



Índice de tablas

Tabla 1. Principales dificultades lingüísticas de estudiantes de bachillerato	10
Tabla 2. Características léxicas del texto académico versus interacciones informales.	58
Tabla 3. Categorización semántica de los sustantivos empleada en esta investigación.	90
Tabla 4. Categorización de los sustantivos como léxico especializado.	93
Tabla 5. Categorización de los sustantivos como léxico abstracto especializado.	96
Tabla 6. Concentrado de variables.	97
Tabla 7. Léxico empleado en los textos del corpus.	106
Tabla 8. Mediciones de cada variable en el conjunto de los textos.	107
Tabla 9. Estimación de los efectos de las interacciones entre variables sobre PromEA.	108
Tabla 10. Medias de uso de las categorías y clases de sustantivos según el nivel de PromEA.	119
Tabla 11. Estimación de los efectos de las categorías y clases de sustantivos sobre el nivel de PromEA.	120
Tabla 12. Medias de uso de léxico no especializado y especializado por campo disciplinar, según el nivel de PromEA.	130
Tabla 13. Estimación de los efectos del léxico especializado sobre el nivel de PromEA.	133
Tabla 14. Frecuencia de uso de los cruces de categorías de sustantivos en la muestra.	138
Tabla 15. Frecuencia de uso de los cruces de categorías de sustantivos según el nivel de PromEA.	139
Tabla 16. Estimación de los efectos del léxico abstracto especializado sobre el nivel de PromEA.	144
Tabla 17. Posición, campo y procedencia de los sustantivos abstractos especializados más frecuentes en la muestra.	146
Tabla 18. Propuesta de jerarquía de los indicadores léxicos de dominio de la escritura académica.	163
Tabla 19. Aportes conceptuales y metodológicos de los estudios previos.	197
Tabla 20. Propuesta de la lexicología cognitivo-funcional de sustantivos de Schmid.	213

Índice de figuras

Figura 1. El continuo del discurso especializado y sus géneros.	41
Figura 2. Cruce de las categorías de sustantivos.	95
Figura 3. Efecto de la interacción entre la densidad nominal y la diversidad léxica sobre la calificación.	109
Figura 4. Efecto de la interacción entre la densidad léxica y la cantidad de palabras en el texto sobre la calificación.	110
Figura 5. Efecto de la interacción entre la densidad nominal en el léxico y en la cláusula sobre la calificación.	111
Figura 6. Efecto de la interacción entre la densidad y la diversidad nominal sobre la calificación.	112
Figura 9. Frecuencias y proporciones de las clases de sustantivos abstractos en el corpus.	116
Figura 10. Frecuencias y proporciones de las categorías de sustantivos en la muestra de textos.	117
Figura 11. Frecuencias y proporciones de las clases de sustantivos abstractos en la muestra de 60 textos.	118
Figura 13. Uso de sustantivos abstractos en la muestra, según el nivel de PromEA.	122
Figura 14. Uso de ‘Eventos del mundo físico y estados de cosas’ en la muestra, según el nivel de PromEA.	123
Figura 15. Uso de ‘Constructos teóricos y conceptos’ en la muestra, según el nivel de PromEA.	124
Figura 16. Uso de ‘Procesos del mundo mental y estados afectivos’ en la muestra, según el nivel de PromEA.	125
Figura 17. Uso de ‘Encapsuladores’ en la muestra, según el nivel de PromEA.	125
Figura 18. Proporciones del léxico especializado y no especializado en el corpus.	129
Figura 19. Proporciones del léxico especializado por campo disciplinar en el corpus.	130
Figura 20. Proporciones del léxico especializado en la muestra.	131
Figura 21. Proporciones del léxico especializado por campo disciplinar en la muestra.	131
Figura 22. Uso de términos especializados según el nivel de PromEA.	134

Figura 23. Uso de términos especializados de Lengua y literatura según el nivel de PromEA.....	135
Figura 24. Proporciones de los cruces de categorías de sustantivos en la muestra.....	138
Figura 25. Proporciones de los cruces de categorías de sustantivos en la muestra, según el nivel de PromEA.....	139
Figura 26. Uso de los sustantivos concretos comunes según el nivel de PromEA.....	141
Figura 27. Uso de los sustantivos concretos especializados según el nivel de PromEA.....	142
Figura 28. Uso de los sustantivos abstractos comunes según el nivel de PromEA.....	143
Figura 29. Uso de los sustantivos abstractos especializados según el nivel de PromEA.....	144
Figura 30. Pauta de escritura de los textos.....	206
Figura 31. Rúbrica original que se adaptó para esta investigación.....	207
Figura 32. Primera adaptación de la rúbrica.....	208
Figura 33. Versión final de la rúbrica.....	209

1. INTRODUCCIÓN: EL RETO DE LA ESCRITURA ACADÉMICA

Estar alfabetizado es un derecho humano fundamental porque esta condición dota a los individuos con un conjunto de habilidades que sientan las bases para aprender a lo largo de la vida, y, por ende, para su desarrollo en la sociedad del conocimiento (UNESCO, 2005). Unas de esas habilidades valiosas en el ámbito de la educación formal son saber estudiar y saber aprender, las cuales forman parte del capital cultural que todas las personas deberían poseer (Bourdieu, 1997). No obstante, en países con distintos grados de desarrollo, aun hoy existe una brecha entre el capital cultural deseado y el real, lo que convierte al saber en un privilegio (Ezcurra, 2011).

La extensión y las características de esta brecha en la literacidad han sido documentadas por organismos internacionales (OCDE, 2000; UNESCO, 2008, entre otros); los resultados de esos estudios colocan a muchos países latinoamericanos en una situación desfavorable en el logro de los objetivos de la educación básica y en la formación de competencias indispensables para la vida social y laboral (Bravo, 2007).

En México, los alcances del logro académico entre la población estudiantil son medidos y estudiados por instituciones nacionales de manera regular (ANUIES, 2014; CENEVAL, 2017; SEP, 2017; INEE, 2019); en particular el dominio de habilidades lingüísticas y el conocimiento del español, como aspectos integrantes de la literacidad académica, han merecido análisis acuciosos por sus efectos sobre los indicadores

educativos en los niveles básico, medio superior¹ y superior (Backhoff, Larrazolo & Tirado, 2011; Rivero, 2010; Zamudio, 2016, entre otros).

Para los jóvenes mexicanos, las posibilidades de ingresar y mantenerse en el nivel superior dependen, además de factores socioeconómicos, de sus habilidades para insertarse exitosamente en las prácticas de discurso que caracterizan el ámbito académico. Sin embargo, pruebas recientes demuestran que los conocimientos sobre el sistema de la lengua y las habilidades para la lectura y la escritura de textos de la mayoría de los egresados del nivel medio superior no son suficientes para asegurar su desarrollo académico y profesional a futuro (ANUIES, 2014; CENEVAL, 2017), lo cual los coloca en situación de desventaja frente a quienes poseen las destrezas lingüístico-discursivas que exige la universidad. En el año de cierre de esta tesis, la Misión Permanente de México ante la OCDE (2019) ha resaltado la importancia de la educación superior para el desarrollo cultural, social y económico del país. Dado que el 77% de los mexicanos entre 25 y 34 años no cuenta con estudios superiores, muy por encima de la media de los países miembros de ese organismo (44%), esta representación diplomática plantea que para promover la equidad en la educación superior es fundamental intensificar los esfuerzos para mejorar la educación en el bachillerato.

En razón de lo anterior, el bachillerato provee oportunidades inmejorables para la investigación aplicada en el campo del lenguaje, concretamente para la indagación acerca de los conocimientos y habilidades que garanticen a los estudiantes un nivel suficiente y adecuado de literacidad académica que facilite su ingreso y permanencia en la universidad. En este marco, el presente trabajo de tesis se propuso ofrecer elementos empíricos, basados

¹ Algunos de los exámenes de logros de aprendizaje y conocimiento del idioma con alcance nacional en México son, para el nivel medio superior: ENLACE-MS (Exámenes Nacionales del Logro Académico en Centros Escolares); EXHALING (Examen de Habilidades Lingüísticas); EXANI-II (Examen Nacional de Ingreso a la Educación Superior); EXCALE-12 (Exámenes de la Calidad y el Logro Educativo); EXHCOBA (Examen de Habilidades y Conocimientos Básicos), y PLANEA EMS (Plan Nacional para la Evaluación de los Aprendizajes).

en análisis de corpus en busca de indicadores de dominio de la escritura académica por parte de los estudiantes al final de la educación media superior.

Esta búsqueda se fundamenta en el hecho de que las pruebas estandarizadas reportan que más del 80% de los estudiantes de preparatoria en México se ubica en los niveles insuficiente y elemental de conocimiento de la lengua española (Rivero, 2010); alrededor de la mitad de ellos falla al identificar palabras incorrectas en un texto, en el empleo de sinónimos y antónimos, en tareas de extensión de conceptos y en la comprensión del significado textual de las palabras (Backhoff, Larrazolo & Tirado, 2011); muy pocos logran construir una argumentación, aunque cuenten con información que contribuya a tomar una postura, y la mayoría se concreta a expresar una opinión (Zamudio, 2016), y solo un 9% de ellos tiene las habilidades requeridas para realizar adecuadamente proyectos de investigación, planear escritos atendiendo a las propiedades textuales y a diferentes propósitos, e interpretar y valorar información de diferentes tipos de texto (INEE, 2019).

La Tabla 1 sintetiza algunos aspectos de la problemática planteada, sin un orden particular, con referencia a las pruebas estandarizadas de donde se obtuvo la información.

Tabla 1. *Principales dificultades lingüísticas de estudiantes de bachillerato*

Dificultad	Fuentes
Conocimiento del sistema de la lengua española	ENLACE-MS 2007, 2008, 2009
Dominio léxico	EXHCOBA 2006, 2007; PLANEA EMS 2017
Capacidad para argumentar	EXCALE-12 2016; PLANEA EMS 2017
Comprensión de textos de diversos tipos	EXCALE-12 2009, 2010; PLANEA EMS 2017
Planificación y redacción de textos escolares	EXCALE-12 2009, 2010; PLANEA EMS 2017

Fuente: Elaboración propia con base en Backhoff, Larrazolo y Tirado (2011), INEE (2019), Rivero (2010) y Zamudio (2016).

Las causas del problema –el bajo nivel de dominio de la lengua española y de las habilidades básicas de lectura y escritura– son de diversa índole: sociales, institucionales, contextuales y pedagógicas. Mucho se puede reflexionar al respecto, pero no es el interés principal del presente estudio. En la actualidad, además, están presentes en la discusión de políticas educativas las causas relacionadas con el acceso a los servicios educativos y a la cultura letrada, en las que entran en juego las condiciones del entorno social en que se desarrollan los jóvenes. Muchos de ellos han crecido en un mundo de imágenes, no de conceptos (Rivero, 2010), y se ha probado que distintos grados de acercamiento con la cultura escrita tienen un papel decisivo en la posibilidad de esos jóvenes para adquirir y emplear adecuadamente el léxico abstracto y especializado que caracteriza al lenguaje escolar (Hess, 2013). Para entender y producir textos académicos, el estudiante debe ser capaz de distanciarse de su registro lingüístico, familiarizarse con un nivel de discurso generado por expertos de carácter mucho más denso, abstracto y complejo que el empleado en las interacciones cotidianas informales (Schleppegrell, 2001, 2004), el cual muy raramente ha sido objeto de enseñanza.

Si de lo que se trata es de reducir la brecha entre los mejor provistos por el sistema educativo y aquellos que han sido menos favorecidos, una de las habilidades que se puede enfocar en la enseñanza es el dominio de la escritura académica, pues la competencia en escritura es fundamental para los estudiantes del nivel medio superior porque gran parte de la evaluación y acreditación se realiza a través de prácticas letradas escritas, y aun las prácticas orales tienen como base, en su mayoría, una variedad de textos escritos que los alumnos deben aprender a valorar y utilizar (Navarro & Revel, 2013; Navarro, Ávila & Gómez, 2019). Además, el dominio léxico y la competencia de escritura académica son requisitos que se infieren del perfil de egreso del bachillerato en México (SEP, 2016, 2017a), y que se desprenden de los estándares internacionales a los que se aspira en este país, de cara al mejoramiento de la calidad de la educación (Montoya, 2018; OCDE, 2019).

De esta circunstancia partió el interés por contribuir a la construcción de indicadores de dominio de la escritura académica a partir del léxico empleado por estudiantes del último semestre de un programa de bachillerato, porque este conocimiento puede ser útil para evaluar con mayor objetividad, a través de sus textos, su competencia de escritura académica, y para contribuir a la formación de la misma. La decisión de enfocar el estudio en el léxico se fundamenta en que su dominio es crucial tanto para la comprensión como para la producción de textos de alto valor conceptual y disciplinar que requieren el manejo de recursos lingüísticos para empaquetar información y para clasificar fenómenos y eventos en una palabra o una frase, además de emplear lenguaje disciplinar (Halliday & Martin, 1993; Schleppegrell, 2004; Snow & Uccelli, 2009; Zwiers, 2008). Se seleccionó un bachillerato general porque la mayoría de los egresados del nivel medio superior que presentan exámenes de ingreso a las universidades se forma en esta modalidad; y por otra parte se decidió que fuera uno de sostenimiento privado en razón de que, tradicionalmente, este medio escolar es el más favorecido en infraestructura educativa, en acceso a las tecnologías y, por lo general, también en el manejo de prácticas educativas innovadoras y de un profesorado capacitado para realizarlas. Todas estas condiciones, verificadas por la investigadora en el caso de los donantes de los textos, facilitan la aparición de los rasgos del lenguaje de interés para el presente estudio y extrapolables como rasgos deseables de cualquier estudiante en ese nivel (SEP, 2017a). Esta decisión metodológica también se apoya en la corriente que propone dirigir el análisis del discurso hacia corpus ecológicos y representativos (Parodi, 2007b).

Para el logro del objetivo central de identificar cuáles aspectos del uso del léxico pueden ser indicadores efectivos de dominio de la escritura académica, se extrajeron las listas de palabras y se clasificaron y cuantificaron los sustantivos de un corpus de textos producidos por estudiantes en un bachillerato del centro de México, agrupados en dos niveles de dominio según un promedio de calificaciones otorgadas por jueces expertos. Las

mediciones de extensión, densidad léxica, diversidad léxica y densidad nominal se correlacionaron con la calificación de los textos por medio de regresiones lineales múltiples, y en otra fase los sustantivos se categorizaron bajo criterios semánticos en concretos, abstractos y cuatro clases de abstractos (Lyons, 1977; Schmid, 1999b), luego en especializados y en abstractos especializados (Chung & Nation, 2004; Martín, 2004; Zwiers, 2008). Con estos datos se realizaron pruebas estadísticas para determinar si las variaciones en el uso predicen una variación en la calificación, con especial atención en los sustantivos abstractos y en el léxico especializado, por la complejidad cognitiva que la teoría le atribuye a este tipo de léxico (Bello, 2016; García, Hall & Marín, 2005; Halliday & Martin, 1993; Schleppegrell, 2004; Schmid, 1999a; Zwiers, 2008). Esta es una investigación de lingüística aplicada con el método cuantitativo correlacional, para el que se emplearon herramientas de análisis léxico automatizado y un programa estadístico avanzado para realizar pruebas y gráficos (AntConc, Freeling, R).

El presente documento se desarrolla en diez partes incluyendo esta introducción general al tema de investigación. En el capítulo 2 se ofrece una síntesis de los antecedentes en materia de calidad de la escritura académica y el dominio léxico en la escritura. En el capítulo 3 se presenta el marco teórico, a manera de criterios para orientar la investigación e interpretar sus hallazgos. En el capítulo 4 se plantea el problema y se establecen los objetivos de investigación, junto con las preguntas que guiaron las distintas etapas del análisis, así como la hipótesis de trabajo. El capítulo 5 se dedica a exponer la metodología, incluyendo la descripción del corpus, la definición de las variables del estudio y los procedimientos de obtención y análisis de los datos. En el capítulo 6 se presentan en cuatro apartados los resultados de cada etapa del análisis, lo cual se complementa con la discusión de los mismos en el capítulo 7. Finalmente, el capítulo 8 se dedica a aportar las conclusiones derivadas de los principales hallazgos de este trabajo de tesis.

2. ANTECEDENTES

En este capítulo se presenta una revisión de los principales artículos de investigación y estudios que sirvieron para precisar el problema y orientaron el enfoque teórico y las elecciones metodológicas de este trabajo de tesis. En una primera sección se presentan abreviadamente algunas investigaciones que han abordado la cuestión de la *calidad de la escritura académica* para identificar indicadores de evaluación del dominio en la escritura en español, en inglés como primera y segunda lengua, y en hebreo. Las preguntas centrales de este primer grupo de trabajos son: ¿qué hay que evaluar en la escritura académica? y ¿cómo evaluarla? Dichas investigaciones se organizan en tres incisos según la disciplina: los estudios lingüísticos, los pedagógicos y los multidisciplinarios; al interior de los incisos, las síntesis de los estudios se presentan en orden cronológico.

En la segunda sección de trabajos se resumen las aproximaciones al constructo *dominio léxico*; se trata de estudios de lingüística aplicada a la enseñanza de primera y segunda lengua, una propuesta pedagógica y algunos trabajos multidisciplinarios. Todos ellos indagan en las características del uso del léxico que pueden evidenciar el nivel de dominio de los aprendices en distintas etapas de su formación. Algunos comparan el desempeño léxico oral y escrito desde la perspectiva del desarrollo, mientras que otros realizan abordajes de enfoque sociocognitivo que atribuyen al dominio léxico un peso particular en el éxito académico, por su relación con la calidad de la escritura de textos en el ámbito escolar. Nuevamente, la organización de las síntesis se hace en tres incisos disciplinares: los estudios lingüísticos, los pedagógicos y los multidisciplinarios, y al interior de ellos el ordenamiento es cronológico.

2.1. Estudios sobre evaluación de la calidad de la escritura académica

2.1.1. Desde la lingüística

Torner y Battaner (2002) presentan los resultados de un estudio amplio de la competencia en lengua escrita² de los estudiantes que terminan el bachillerato y se postulan a las universidades españolas. Su objetivo fue describir, para aplicaciones pedagógicas, el nivel de competencia escrita real con fines académicos, al final de los estudios preuniversitarios, en la variedad estándar española. La recopilación del corpus PAAU 1992 cubrió todos los aspirantes universitarios a nivel nacional, pero el análisis se circunscribió al bachillerato privado con altos estándares educativos, de clase media alta, para que se pudiera evidenciar un mayor espectro de características. Se recolectaron las pruebas escritas de dos generaciones (2014 y 2015) y se utilizaron herramientas de análisis automatizado además de la descripción manual. Las categorías de análisis fueron la evidencia de procesos de escritura en cuanto a los movimientos retóricos específicos del texto expositivo (presentación del tópico, desarrollo y organización de la información, aclaración textual y ejemplificación y cierre textual) y los recursos léxico-gramaticales del producto.³ Entre los resultados de este estudio destacan, para los fines de esta tesis, los siguientes: los textos que fueron objeto de análisis están fuertemente condicionados por la situación comunicativa en que se produjeron (evaluación para ingreso a universidades, con tiempos fijos y consignas establecidas, sin planificación ni reelaboración); en ese sentido muestran un nivel de dominio que se puede caracterizar como una destreza acotada a la escritura de textos

² Torner y Battaner (2002) nombran de diversas maneras esta competencia: competencia de lengua escrita, competencia escrita real, competencia redactora, dominio de la competencia escrita en textos académicos, redacción del texto especializado en un determinado nivel, redacción de textos académicos escritos, redacción académica, redacción ofrecida al final del bachillerato, etc., que en general son operacionalizaciones semejantes a las que se manejan en pruebas estandarizadas mexicanas como el EXANI-II.

³ Para fines expositivos, el análisis específico del léxico hecho por Torner y Battaner (2002) se reseña en la sección 2.2, “Estudios enfocados al dominio léxico en la escritura”.

académicos y específicamente del subgénero examen. No obstante puede considerarse un índice de referencia que sobrepasa el conocimiento directo y particular que cada profesor puede llegar a tener de sus estudiantes, y puede contribuir al diseño de cursos de lengua española para mejorar la calidad de su enseñanza. Además, el trabajo de Torner y Battaner aporta escalas de valoración criterial útiles en la planificación didáctica y en la evaluación de textos académicos. En relación con el presente trabajo de tesis, interesa subrayar la importancia de analizar la competencia real por medio del estudio de corpus con criterios y categorías bien sustentados, que permitan hacer una descripción local, situada y restringida a los géneros relevantes para la formación pre-universitaria. Los autores resaltan el valor pedagógico de estos conocimientos para poder hacer explícito lo que hay que enseñar, además del valor didáctico de la construcción y socialización de pautas.

Por su parte, Berman y Nir (2009) abordan en su artículo el constructo *calidad del texto*, el cual, señalan, es accesible para expertos y no expertos de manera intuitiva aunque requiere precisiones basadas en evidencias. Las investigadoras israelíes toman como base los trabajos de Peterson y McCabe (1983) y Malvern, Richards, Chipere y Duran (2004), que confirman la complejidad del constructo desde dos puntos de vista: uno es que los jueces con diferentes bagajes y diferentes motivaciones atienden a diferentes aspectos de un texto y divergen en lo que consideran importante cuando lo leen, desde la originalidad, el estilo literario, hasta la ortografía y la puntuación; otro es que hasta los expertos parecen tener dificultades para definir criterios explícitos, no se diga precisos, para evaluar textos. Con la finalidad de despejar algunos aspectos de este complejo constructo, las autoras midieron la extensión de las palabras, la densidad léxica, el registro (definido como el nivel de uso formal/coloquial) y la abstracción nominal (en una escala adaptada de Ravid, 2006) en producciones escritas por individuos de distintas edades hablantes nativos de hebreo y por otra parte hablantes nativos de inglés norteamericano. Berman y Nir operacionalizaron la noción *calidad global del texto* con indicadores de competencia referentes a la estructura

del discurso y al desempeño en producción escrita y los correlacionaron con los datos de una evaluación que las propias investigadoras hicieron de los textos. Aunque solo encontraron correlaciones entre la calidad global del texto y su extensión medida en número de palabras y en número de cláusulas, el estudio es sumamente relevante para esta tesis por el diseño metodológico y por la consideración de que los jueces (en este caso las mismas investigadoras) se ven influenciados por la producción, es decir, por la extensión de los textos.

2.1.2.Desde la pedagogía

En el entorno escolar, Morales (2004) hace una propuesta de indicadores de evaluación de la escritura en el entorno escolar, especialmente dirigida a la enseñanza media del sistema venezolano. Discute, con base en la teoría sociocognitiva de la escritura, las prácticas arraigadas en los centros escolares que centran la atención en aspectos puramente gráficos de la escritura (ortografía, caligrafía, diseño de la página) y en la noción difusa de *redacción*. Arguye que todos los aspectos de la composición del escrito están indisolublemente relacionados, por lo que la evaluación con rúbricas analíticas no se corresponde con la realidad del proceso ni da cuenta equilibradamente de la calidad del texto producido. Hace hincapié en la necesidad de diseñar pautas holísticas que ayuden efectivamente a evaluar integralmente la escritura, incluyendo la secuencia de acciones que la generaron. En ese trabajo destacan los indicadores del léxico (variedad y amplitud, propiedad, entendido como significados correctos y adecuados al carácter formal del texto), y la adecuación a las características del género, en relación con el objeto de estudio de esta tesis.

También en el contexto de la educación, Atorresi (2005) aborda la cuestión de la redacción de consignas para evaluar la escritura como competencia para la vida, enfocándose en las composiciones escolares. Se basa en la teoría sociocognitiva de la

escritura para sostener que, en la evaluación de la capacidad de expresión escrita, se deben considerar tanto el producto escrito como el proceso de escritura, es decir, los discursos que un estudiante es capaz de redactar y los procedimientos y estrategias que el estudiante utiliza para producir esos discursos. Plantea que una evaluación realista ha de tomar en cuenta que escribir es una habilidad compuesta por una serie de microhabilidades que, aunque se manifiesten al mismo tiempo, no siempre pueden evaluarse en su totalidad, a detalle y de modo superpuesto. Por esa razón, en la construcción de la consigna el profesor debe orientarse por los aspectos que va a enfocar en cada prueba de escritura; proporcionar toda la información necesaria para facilitar los procesos de pensamiento que conducen a la textualización, y también la corrección de los borradores. Incluye en esta información los pasos que deben seguirse para lograr eficientemente la tarea, el propósito de la misma, el contexto en que circulará el escrito, el perfil del lector, el prototipo textual y el tema. En relación con la presente tesis, el trabajo de Atorresi brinda una visión amplia de lo que los profesores pueden necesitar saber acerca de qué evaluar en cada tipo de texto y cómo elaborar guías y rúbricas (“grillas”). Aunque la autora argentina no precisa el nivel educativo al que dirige su propuesta, resulta pertinente para el contexto de la educación media porque en ella no todos los docentes cuentan con la formación necesaria en lengua o en escritura para poder valorar la capacidad de los estudiantes para presentar por escrito la evidencia de sus aprendizajes.

En este tenor, Martínez (2010) aborda el problema de la falta de instrumentos con indicadores eficaces para evaluar la escritura en el sistema educativo chileno. El propósito de su estudio fue analizar los aportes y aplicaciones de los modelos de producción escrita implícitos o explícitos en el constructo con el que el sistema de medición de resultados de aprendizaje de ese país evalúa la escritura en el nivel medio (equivalente al bachillerato mexicano). Encontró que la evaluación que realizan los docentes no es muy profunda, que carece de la consolidación que se ha logrado en la evaluación de la comprensión lectora, así

como de criterios sólidos y que la prueba se orienta normativamente (con respuestas correctas/incorrectas). El investigador propone que la evaluación de la escritura incorpore los hallazgos de las ciencias cognitivas respecto de la producción escrita, las pruebas y materiales se articulen en todo el sistema, para que los tópicos de escritura se elaboren de manera suficiente y se incorpore el modelo “escribir para pensar”.

2.1.3.Desde la multidisciplina

McNamara, Crossley y McCarthy (2009) analizan un corpus de ensayos calificados por expertos con base en una rúbrica de seis niveles para encontrar las diferencias entre aquellos que fueron evaluados como de alta calidad y los que fueron evaluados bajos respecto de los siguientes indicadores: cohesión, complejidad sintáctica, diversidad de palabras (MTLD), frecuencia, concreción e imaginabilidad de las palabras. Su objetivo era establecer cuáles son los indicadores más eficientes para predecir la calidad de un ensayo. La medición de estos rasgos se hizo de manera automatizada (con el programa Coh-Metrix) y la evaluación fue manual, con piloteo y entrenamiento de los jueces. Sus resultados más importantes se encuentran en el hallazgo de que ninguno de los indicadores de cohesión mostró variación entre los textos de alto y bajo nivel de calidad, sino que las diferencias estaban en el uso de las palabras de baja frecuencia, con menor índice de concreción e imaginabilidad. Los investigadores concluyen que los índices textuales que caracterizan una buena escritura escolar no están alineados con los rasgos que facilitan la comprensión de la lectura; al contrario, los textos evaluados por su alta calidad contenían características lingüísticas asociadas con la dificultad del texto y un lenguaje más sofisticado. Para el estudio de la presente tesis, este artículo reviste la mayor importancia porque provee la rúbrica que se adaptó para esta investigación y porque ayudó a deslindar los indicadores léxicos que mostraron potencial predictivo respecto del constructo *calidad del texto*. Asimismo, fue la base para el diseño del método de evaluación por jueces que operó en esta

tesis y el procedimiento de validación estadística de las calificaciones otorgadas a los textos, que fue verificado con una guía para seleccionar y reportar coeficientes de confiabilidad entre jueces (Koo & Li, 2016).

Crossley y McNamara (2012) se propusieron determinar qué indicadores de cohesión y de sofisticación lingüística pueden ser predictores de competencia de escritura en aprendices de inglés como segunda lengua. Partiendo de datos conocidos, como la diversidad léxica y la superposición (repetición) de palabras –a la que consideran un indicador de enlaces referenciales como mecanismos de cohesión–, se enfocaron en deslindar si una mayor diversidad léxica (que mina el solapamiento) es mejor predictor de competencia que la coherencia textual. Con ayuda del programa Coh-Matrix, hicieron mediciones léxicas en 1,200 ensayos de aprendices, los cuales fueron evaluados por universitarios de último año entrenados para ese propósito. Los hallazgos del estudio prueban la viabilidad de analizar diferencias en indicadores que atañen al uso del léxico y a los recursos cohesivos entre las producciones escritas de estudiantes de L2 evaluados con niveles alto y bajo de competencia. Afirmaron su hipótesis de que, como los escritores nativohablantes, los aprendices con alta calificación no parecen producir textos más cohesivos sino más sofisticados en el uso de la lengua, lo cual es visible a través del manejo de un vocabulario más amplio, compuesto por palabras de menor frecuencia en el uso cotidiano y menos familiares, que no suelen remitir a referentes concretos. Además, los textos mejor evaluados usan menos mecanismos de cohesión temporal y tienen niveles menores de repetición. Estos hallazgos fueron un antecedente importante para la presente investigación porque enfocan directamente al tipo de indicadores que interesan para el objetivo central de la tesis; además, se adaptó el diseño metodológico para analizar la escritura de hablantes del español.

Siguiendo esta línea, Uccelli, Dobbs y Scott (2012) analizaron el comportamiento de una selección de predictores de calidad de la escritura en un corpus de 51 ensayos de tipo

persuasivo que fueron evaluados por la calidad de su escritura por un grupo de profesores. El propósito de estas investigadoras fue identificar las características clave del lenguaje académico que serían relevantes para la enseñanza de la escritura; para ello, midieron con el software CLAN indicadores léxicos y gramaticales como la extensión en cláusulas, la diversidad léxica (VocD), la densidad léxica (la proporción de palabras de contenido por cláusula que refleja la densidad informativa compactada en el texto), la complejidad sintáctica y una serie de indicadores discursivos como los marcadores de posicionamiento y los ordenadores. Sus principales hallazgos arrojan luz sobre las expectativas implícitas de los profesores de nivel medio acerca del dominio del lenguaje académico por parte de los alumnos, en particular, en la escritura de ensayos persuasivos. El estudio confirma la relevancia de la diversidad léxica como indicador de dominio, junto con la complejidad sintáctica y el tejido discursivo; en este tipo de textos se espera una alta frecuencia de organizadores del discurso y marcadores epistémicos, específicamente de posicionamiento, los cuales mostraron ser predictores eficientes de la calidad del ensayo. Si bien el artículo no ofrece evidencia sobre la efectividad de determinadas prácticas de enseñanza, sí abona a la comprensión del constructo *calidad de la escritura* y a la identificación de aspectos léxico-gramaticales importantes que habría que tener en cuenta en la planificación de la enseñanza de la escritura. Estas conclusiones son los principales aportes de su trabajo a la presente tesis, que se enmarca en la misma línea de investigación.

Uccelli, Barr, Dobbs, Phillips Galloway, Meneses y Sánchez (2014) proponen una serie de indicadores del constructo *competencia en lenguaje académico*, el cual rebasa las habilidades léxicas, sintácticas y discursivas con las que ha sido caracterizado este tipo de lenguaje en estudios de desarrollo y lingüística aplicada. Parten de una perspectiva pragmática sociocultural del desarrollo del lenguaje para postular la ganancia tardía de una mayor “flexibilidad retórica” que permite a las personas adaptarse a diferentes contextos discursivos. Desde su posicionamiento, sostienen que la participación de los estudiantes en

comunidades de discurso académicas puede presentar dificultades para ellos a pesar de que sean hábiles y articulados para comunicarse en contextos coloquiales, pues el lenguaje académico tiene particularidades con las que no han tenido oportunidad de familiarizarse. Sobre la base de numerosos estudios que han dado evidencias de la importancia del dominio del vocabulario para comprender textos, estos investigadores asumen que el conocimiento del vocabulario puede servir como un indicador poderoso de un conjunto de habilidades lingüísticas que los individuos han desarrollado en sincronía. Por ejemplo, una persona letrada que ha desarrollado un extenso y rico repertorio de vocabulario académico también ha aprendido a usarlo de maneras apropiadas: a compactar información densa a través de la subordinación y la nominalización; a marcar relaciones conceptuales a través de los conectores apropiados; a usar diversas estrategias referenciales para relacionar los temas en un texto, etc. A partir de esta observación, los autores incluyen en su definición del constructo *competencia en el lenguaje académico* una serie de habilidades clave cuya aparición en tareas de aprendizaje orales escritas es fundamental, pues contribuyen positivamente a la comprensión y escritura de textos en la educación media y superior, en el entendido de que durante la adolescencia estas se siguen desarrollando. Estas habilidades forman el constructo Core Academic Language Skills (CALs), el cual amplían en Uccelli, Phillips Galloway, Barr, Meneses y Dobbs (2015), para incluir: el desempacar información condensada en palabras y oraciones complejas; conectar ideas lógicamente; rastrear participantes y temas en el discurso; organizar textos analíticos, especialmente los argumentativos, y reconocer el registro académico frente a alternativas más coloquiales. Aunque los trabajos de estos investigadores se adscriben a los estudios del desarrollo, sus conclusiones principales se alinean con la perspectiva teórica de este trabajo de tesis, que incluye entre sus criterios una base sociocognitiva.

Meneses, Uccelli, Santelices, Ruiz, Acevedo y Figueroa (2017), Figueroa (2018) y Figueroa, Meneses y Chandía (2018) adaptaron el modelo CALs de Uccelli et al. (2014,

2015), con el mismo conjunto de habilidades, para evaluar y comparar las habilidades en lenguaje académico de hablantes nativos de español; lo llamaron Habilidades Clave de Lenguaje Académico (HCLA) y de él desprendieron una serie de instrumentos para evaluar, en primer lugar, aptitudes para la lectura (S-CALS); luego, para identificar el uso de un repertorio de formas léxico-gramaticales y discursivas para organizar ideas y expresar una postura en una variedad de textos producidos en contextos escolares (S-AVoc). En este último caso, el objetivo fue probar si el nivel de manejo de vocabulario académico predice significativamente la calidad de las argumentaciones y explicaciones de estudiantes de educación básica en Chile. La *calidad de la escritura*, para Figueroa, Meneses y Chandía (2018), incluye el conocimiento del género, el desarrollo de la idea y la organización discursiva. Para el texto argumentativo se consideró el posicionamiento, argumentos y contraargumentos y la organización discursiva; para la explicación se añadió la exposición del fenómeno y el desarrollo de la idea; esas dimensiones se asentaron en una rúbrica para medir la calidad, que luego fue correlacionada con mediciones de un instrumento mixto que reúne las habilidades de lenguaje académico en español y el dominio de vocabulario académico (S-CALVS). Los resultados muestran una correlación positiva y significativa de los aspectos considerados en el modelo mixto con la calidad de ambos géneros discursivos. Aunque no existían precedentes que midieran las habilidades académicas en términos receptivos y luego las examinaran en relación con la producción escrita, las asociaciones encontradas entre los recursos léxico-gramaticales de los escritores en sus textos y la calidad de los mismos fue positivamente confirmada por esa investigación: tanto el dominio de habilidades clave como el manejo de vocabulario académico predicen la calidad de la escritura separadamente y de manera distinta según el género discursivo. El poder predictivo de estas habilidades fue sin embargo menor que cuando se les relaciona con la comprensión de lectura (Meneses et al., 2017), debido a que, según las autoras, en la producción de un texto influyen múltiples factores de índole extralingüística (motores,

cognitivos y metacognitivos). Estos hallazgos se extienden en Figueroa (2018), uno de cuyos resultados fue que la longitud de los textos –la productividad de los estudiantes medida en cláusulas– tuvo correlación positiva con las medidas de calidad y con la mayoría de los recursos del lenguaje académico en ambas tareas de escritura, por lo que la longitud se revela como una variable a controlar en modelos de regresión. De estas investigaciones se desprende la utilidad de separar conceptualmente las habilidades receptivas de lenguaje académico de las productivas, en el entendimiento de que las primeras se aprenden durante la escolaridad, se van especializando por medio de interacciones formales tanto orales como escritas y son la base de conocimiento necesaria para el despliegue de las habilidades productivas, vistas como recursos. La autora propone que la enseñanza escolar haga un andamiaje de los recursos de lenguaje académico específicos con tareas que demanden un fuerte repertorio lingüístico, como las explicaciones científicas, para dar oportunidad a los estudiantes de producir textos de calidad. Asimismo, sugiere atender la enseñanza de metalenguaje en el aula para que los alumnos puedan referirse a los fenómenos de la lengua por su nombre y así aumentar sus posibilidades de comprender las demandas de cada tarea, lo mismo que garantizar el acceso de los jóvenes al lenguaje del conocimiento apoyando la adquisición de vocabulario académico transdisciplinar y temático disciplinar para que puedan desarrollar las ideas de sus textos. En cuanto a las limitaciones, señala que hace falta analizar los recursos de lenguaje académico desplegados en la escritura de estudiantes que hayan tenido la oportunidad de revisar y corregir sus textos, para observar si hay un cambio en la calidad de la escritura.

Navarro, Ávila y Gómez (2019) se enfocan en las pruebas de escritura, que se suelen diseñar y aplicar con menor frecuencia que las de lectura, y se dirigen poco al nivel medio y superior. Destacan la importancia de estas pruebas estandarizadas porque pueden ser muy influyentes en la trayectoria de los estudiantes, en el diseño de carreras y en las prácticas pedagógicas de los profesores. Considerando que la escritura tiene una importante función

epistémica paralela a su conocida función semiótica, los autores sostienen que la competencia en escritura debe formar parte de los instrumentos de medición de competencias en los programas de estudio del nivel superior, y, por ello mismo, ser objeto de enseñanza explícito y planificado en el currículum, de una manera tal que tome en cuenta su complejidad, pues integra conocimientos, habilidades y estrategias distintas, vinculadas a factores culturales, personales, situacionales, lingüísticos y cognitivos. Tan solo el dominio lingüístico se compone del dominio léxico, gramatical, retórico, semántico y discursivo, ninguno de los cuales puede deslindarse del resto de los factores. Con base en ello, estos investigadores construyen y pilotean una propuesta para evaluar el desempeño escrito al inicio de la educación superior en Chile, basada en los siguientes principios: las pruebas deben evaluar de manera directa muestras reales de escritura de estudiantes; se deben solicitar tareas de escritura situadas y significativas, que coloquen al estudiante en una posición de autoridad para gestionar sus conocimientos, habilidades y experiencias previas; asimismo, las pruebas deben medir dimensiones escritas de alta complejidad, más allá de la ortografía o la puntuación; evaluar con rúbricas holísticas en lugar de analíticas, que pretenden evaluar dimensiones que no existen por separado. Por lo que toca a su relación con este trabajo de tesis, la definición del concepto de *competencia de escritura* resulta ser una orientación teórica útil: capacidad de construir y comunicar conocimiento por escrito para cumplir con propósitos específicos y adecuarse a una situación comunicativa particular a través de un texto coherente, cohesivo y ajustado a las convenciones lingüísticas.

2.2. Estudios enfocados al dominio léxico en la escritura

2.2.1. Desde la lingüística

Torner y Battaner (2002), cuyo trabajo fue presentado parcialmente en un apartado previo (2.1.1), analizan el conocimiento real del uso del vocabulario en estudiantes al final del

bachillerato. Su objetivo es ofrecer información que pueda servir como norma de referencia para la evaluación de esta competencia. Operan con algunas definiciones de las nociones estudiadas como *vocabulario conocido y usado en producciones de estudiantes*, que es el objeto a describir; *valor de término*, que es la función de la palabra cuando la primera acepción de un diccionario es terminológica, es decir, disciplinar; *sustantivo del léxico general*, aquel cuya primera acepción es no terminológica, o sea, no relacionada con un campo disciplinar; *nombres de alto valor conceptual*, sustantivos que construyen conceptos operativos manejados en el discurso de las humanidades o de las ciencias, etc. Entre los resultados de la descripción del corpus, estos investigadores encontraron que el léxico que usan los estudiantes en exámenes escritos está muy ligado al tema (materia examinada), se expresa por medio de sustantivos abstractos que designan conceptos, por medio de nombres concretos empleados para la descripción de objetos, y por medio de vocabulario disciplinar y términos científicos. Este vocabulario conocido y empleado pone en evidencia el dominio de un tipo de texto específico escrito con fines académicos (más concretamente con fines de evaluación), pero no se puede saber si los estudiantes saben utilizar ese vocabulario con otros fines. No obstante, concluyen los autores, siendo el conocimiento del léxico uno de los componentes implicados en la calidad textual, es muy importante conocer la competencia real de los alumnos antes de incorporar acciones didácticas en el aula. Una conclusión de la investigación que resulta de la mayor relevancia es que la calidad del texto implica el dominio de varios recursos, que incluyen el buen uso de los mecanismos léxicos, retóricos y sintácticos. El presente trabajo de tesis, aunque no se aboca a todos los niveles de análisis lingüístico que comprende el estudio de Torner y Battaner (2002), converge con el objetivo de describir el uso lingüístico real de los estudiantes, relacionarlo con indicadores de evaluación y derivar propuestas pedagógicas.

Strömqvist, Johansson, Kriz, Ragnarsdóttir, Aisenman y Ravid (2002), desde la psicolingüística, hacen un análisis contrastivo de las mediciones léxicas de diversidad,

densidad léxica y extensión de la palabra, sobre tareas de producción oral y escrita realizadas por sujetos escolarizados de distintos niveles, con el programa CLAN. Parten del supuesto de que las diferencias en las condiciones comunicativas y las restricciones de procesamiento entre habla y escritura conducen a una diferenciación en términos de organización de la información mediante recursos lingüísticos. En su estudio encontraron diferencias en el léxico y en el contenido proposicional entre las producciones orales y escritas, específicamente puntajes más altos en los tres indicadores en la escritura: diversidad léxica, densidad léxica y tamaño de la palabra. Sus hallazgos concuerdan con los de Ure (1971), que estableció una mayor densidad léxica en la escritura y una mayor proporción de palabras diferentes con respecto de corpus orales; y con el trabajo de Chafe (1985) que propone como rasgo distintivo de la oralidad una mayor restricción de las opciones léxicas a causa de la velocidad de la interacción cara a cara, y, por otro parte, atribuye a la escritura un mayor empaquetamiento de la información, que conduce al uso de recursos léxicos diferentes. Este trabajo, junto con sus antecedentes, aunque toma la perspectiva del desarrollo contribuye a la presente investigación porque sienta precedentes de algunas unidades de análisis relevantes.

Chung y Nation (2004) problematizan las dificultades para identificar de manera precisa las palabras que funcionan como términos técnicos en dos corpus, uno de anatomía y uno de lingüística aplicada. Los autores neozelandeses compararon la precisión de las escalas de puntuación del nivel de especialización de una palabra, los diccionarios técnicos y las pistas presentes en el texto como ayudas para identificar los términos técnicos en ambos corpus. La investigación tuvo como objetivo determinar cuál método ofrece un mayor grado de precisión y confiabilidad en dicha identificación. Partiendo del conocimiento de que el vocabulario técnico está estrechamente relacionado con la asignatura o área disciplinar, proponen cuatro métodos para identificar las palabras en función de término: las escalas de valoración, que clasifican las palabras de acuerdo a cuán

cerca están de algún área en particular; el uso de diccionarios técnicos, equivalente a la consulta a un especialista en el campo; la búsqueda de pistas del escritor, quien, como especialista también y como miembro de la comunidad académica, sigue convenciones disciplinares para marcar, resaltar e incluso definir los términos que emplea en su discurso; y, dado que el vocabulario técnico ocurre en dominios especializados, el cuarto método es comparar la frecuencia de ocurrencia de las palabras en un corpus de un área en particular respecto de las ocurrencias o no ocurrencias en otros corpus. Los investigadores concluyen que el uso de escalas de puntuación es el método que arroja resultados más precisos, seguido por las pistas del autor, aunque conceden un alto valor práctico al uso de analizadores léxicos automatizados, que permiten combinar el conteo de frecuencias con el análisis de las colocaciones en el discurso como pistas del académico. Un resultado paralelo a este trabajo comparativo tiene particular interés para la presente tesis, pues al poner a prueba los cuatro métodos con dos corpus, uno de anatomía y uno de lingüística aplicada, encontraron que el vocabulario técnico puede llegar a ocupar una significativa proporción del total de palabras de un texto: del 20% (en lingüística) al 30% (en anatomía). Además, según la disciplina pueden albergar palabras de clases y registros diferentes: mientras que en anatomía los términos son estrictamente empleados en esa disciplina, en lingüística se encontraron palabras que pueden ocurrir en otros usos del lenguaje (*negociación, interacción, incidental*) y que en ese campo cobran un valor de término. Estas conclusiones fueron de interés como antecedente teórico y metodológico para dos de los análisis conducidos en esta investigación, los cuales se dedicaron al estudio de los sustantivos con referente especializado.

Ravid (2006) examina desde una perspectiva psicolingüística el desarrollo en la infancia y la adolescencia del léxico nominal en producciones orales y escritas, narrativas y expositivas, en hebreo. Para ello diseñó una escala nominal de diez niveles con base en la semántica de los sustantivos de Lyons (1977); con ella, puntuó todos los ítems nominales

de 320 textos y los comparó entre edades, niveles de escolaridad, modalidades y género, sobre la base teórica de que el léxico juega un papel central en la adquisición del lenguaje y es esencial para el desarrollo de la sintaxis, dado que una mayor proporción de palabras de contenido, especialmente sustantivos y adjetivos, conduce a una mayor complejidad en la arquitectura sintáctica y, por otra parte, la densidad nominal refleja la adquisición de algunas habilidades cruciales en el desarrollo lingüístico como la de empaquetar información densa y cohesionar el texto. Su hipótesis fue que los textos escritos por los niños mayores y adolescentes contendrían más sustantivos (y nominalizaciones) y que estos serían cada vez más categóricos y más abstractos; por lo que el aumento de estos recursos se observaría en función de la edad y la escolaridad. Tal como Ravid había hipotetizado, los textos escritos resultaron ser más densos nominalmente que los orales, y también desplegaron un vocabulario nominal más abstracto y categórico que la producción oral, de tal modo que la densidad nominal resulta ser un buen indicador de desarrollo léxico y sintáctico porque subyace a la arquitectura de los textos. Las aportaciones de este trabajo a la tesis que aquí se expone fueron, concretamente, la orientación teórica hacia los fundamentos de su escala nominal y el enfoque a los sustantivos abstractos.

Dentro de la tradición que compara habla y escritura en los estudios del desarrollo, Berman y Nir (2010) realizaron mediciones del tamaño de la palabra, el registro (nivel de vocabulario), densidad, diversidad léxica y abstracción (con base en Ravid, 2006), en textos orales y escritos, narrativos y expositivos, producidos por personas de distintos grupos de edad y cuatro niveles de escolaridad; indicadores que según las investigadoras reflejan un estilo de expresión más elevado y cuidadosamente monitoreado. En este examen del léxico las autoras asumen que el dominio de la lengua escrita, como distintivo de la alfabetización, es un proceso que se extiende mucho más allá de la escuela primaria, y que aprender a escribir es más que solo expresar las ideas en un medio alterno al habla, además de que las modalidades oral versus escrito, las circunstancias comunicativas y sus diferentes

restricciones de procesamiento tienen impacto tanto en la forma como en el contenido de los textos que se producen. En busca de correlaciones entre los niveles de desarrollo, las modalidades y el género, y la influencia de la escolaridad con respecto a los usos del léxico, encontraron que en la escritura adulta priva un estilo más elevado, más abstracto y más complejo morfológicamente, hallazgos que constituyen la principal contribución a este trabajo, porque compilan algunas de las variables de interés para la tesis.

Berman, Nayditz y Ravid (2011) inscriben su trabajo en la línea del desarrollo lingüístico tardío; su objetivo es comparar producciones escritas por niños y adolescentes en hebreo, de dos niveles educativos y con diferentes historiales académicos. Se enfocan en el léxico y la sintaxis como indicadores de madurez (extensión del texto en cláusulas y en cantidad de palabras, densidad sintáctica, diversidad y densidad léxica), para probar esos rasgos como aspectos de un *diagnóstico del texto*, equiparado a la *calidad léxica en la escritura*. Se basan en estudios previos que destacan el papel del conocimiento léxico en el desarrollo tardío del lenguaje y en la alfabetización en general, como un componente crucial en actividades de alto nivel cognitivo, crítico para el logro académico durante los años escolares. Uno de sus indicadores es la presencia de sustantivos entre las palabras de contenido, que son cruciales para hacer referencia a las entidades en el discurso, siendo indicativa de un mayor desarrollo la aparición en el texto de sustantivos más específicos y más sustantivos abstractos, lo mismo que sustantivos derivados (nominalizaciones), los cuales son considerados muy importantes para la construcción de los textos expositivos propios de la escolarización. Este trabajo fue relevante para la investigación del presente estudio porque se acerca, desde su perspectiva, al objetivo de esta tesis de conocer el peso de ciertas características del léxico, y particularmente de los sustantivos, en la valoración de los textos; asimismo, las conclusiones acerca de la importancia de ciertos procesos de empaquetamiento de información en los sustantivos para la construcción de textos en el ámbito escolar ayudaron a interpretar algunos de los resultados del presente trabajo.

Hess (2013) estudió narraciones orales y escritas de adolescentes escolarizados, agrupados según su acceso previo a la cultura escrita, con la finalidad de observar la forma en que ellos realizan elecciones léxicas en cada contexto discursivo. Para ello recurrió al análisis de contenido semántico-pragmático de los sustantivos con la escala nominal de Ravid (2006) y a la identificación de elementos del lenguaje académico literario, partiendo del supuesto de que el entorno escolar provee a los niños y adolescentes de un nuevo tipo de vocabulario compuesto por palabras más abstractas y menos transparentes semánticamente que aquellas a las que están acostumbrados en las interacciones cotidianas. Esta característica del léxico se combina con estructuras sintácticas más complejas para una estructuración más densa de la información en el texto, el cual ofrece una mayor carga conceptual en menos palabras. La autora enfoca el análisis a los sustantivos en razón de que estos desempeñan un papel crucial en la organización discursiva, dado que portan en gran medida el flujo de la información textual, y porque permiten observar la incorporación de términos técnicos, vocabulario especializado (en este caso literario), nominalizaciones y uso del sentido figurado. En los dos grupos estudiados no se encontraron diferencias importantes debidas a la edad, pero sí las hubo en cuando a la modalidad y al contacto previo con la escritura. Los textos escritos presentaron una mayor cantidad de información en menos cláusulas respecto de los orales, lo cual refleja un mayor empaquetamiento de información; ambos grupos mostraron predominancia de sustantivos de primer orden (en la escala de Lyons, 1977, adaptada por Ravid, 2006) en la narrativa oral, mientras que la modalidad escrita favorece la aparición de sustantivos de segundo y tercer orden, aunque con diferencias relacionadas con la escolaridad previa. Otros resultados refieren el uso más difundido de expresiones no literales en la modalidad escrita que en la oral. Hess (2013) pudo verificar también un uso más frecuente de sustantivos propios del lenguaje académico literario y un mayor nivel de abstracción de los sustantivos en las narraciones escritas, particularmente en las producidas por los adolescentes más inmersos en la cultura escrita.

El enfoque al análisis de los sustantivos en busca de los rasgos de lenguaje más académico, más especializado y abstracto es el aporte principal de este artículo a la presente investigación.

2.2.2.Desde la pedagogía

En cuanto a estudios de corte pedagógico que revisan el dominio léxico, se encuentra el trabajo de Gómez, Sotomayor, Bedwel, Domínguez y Jéldrez (2016), quienes estudiaron la relación entre la calidad del vocabulario y el desarrollo de la habilidad de escritura en los niños chilenos. Dispusieron de una muestra de más de dos mil textos narrativos, persuasivos e informativos escritos en español por estudiantes de cuarto grado de educación primaria. Operacionalizaron la noción de *calidad léxica* como el conjunto de indicadores compuesto por la diversidad léxica (el número de palabras diferentes usadas en el texto), la sofisticación léxica (el empleo de palabras de baja frecuencia y vocabulario complejo) y la densidad léxica (la cantidad de vocabulario de contenido presente en el texto), y modelaron el efecto de las mediciones sobre una calificación global obtenida con base en una rúbrica. Los resultados de los modelos predictivos les permitieron establecer relaciones entre géneros de la siguiente manera: la diversidad léxica fue factor de diferenciación en los textos narrativos y persuasivos; la densidad léxica fue determinante en los tres géneros, mientras que la sofisticación léxica lo fue únicamente para los textos narrativos y expositivos, no para los persuasivos. Aunque hubo más hallazgos relacionados con el sexo y la condición socioeconómica de los participantes, para el presente trabajo de tesis resulta relevante la capacidad de los índices léxicos como determinantes en la *calidad de la escritura* de acuerdo con los géneros.

2.2.3.Desde la multidisciplina

En el mismo tenor de encontrar indicadores para evaluar la calidad académica de la escritura, Crossley, Salsbury, McNamara y Jarvis (2010) pusieron a prueba un modelo de competencia léxica basado en índices léxicos relacionados con la amplitud del vocabulario, la profundidad del conocimiento léxico y la accesibilidad de ítems léxicos fundamentales para la producción de escritura académica. Usaron Coh-Metrix para medir el tamaño de las palabras, la diversidad léxica, la frecuencia de las palabras, hiperonimia, polisemia, concreción de las palabras, entre otras unidades, y un equipo de evaluadores fue entrenado para calificar 240 textos usando una rúbrica específica para valorar el *dominio léxico* mediante cinco niveles. Los investigadores asumen que la competencia léxica es crucial (especialmente en el aprendizaje de una segunda lengua) para el logro académico, por lo que buscaron correlacionar los indicadores léxicos con la calificación para determinar su potencial predictivo. Encontraron que la diversidad léxica, la hiperonimia y la frecuencia de las palabras de contenido tienen efecto sobre la variación de las evaluaciones hechas por humanos. Estos indicadores se relacionan con el tamaño del vocabulario y la profundidad del conocimiento de las características del léxico. No encontraron, sin embargo, correlaciones con el acceso a los ítems léxicos fundamentales. Fueron aspectos orientadores para esta investigación el método de jueces, la rúbrica especializada en el léxico, el método de validación de los datos y algunas de las mediciones de las variables.

Crossley, Salsbury y McNamara (2011) emplearon el análisis discriminante para predecir el nivel de categorización de textos producidos por estudiantes de inglés como L2 según el puntaje de TOEFL de su autor, con base en indicadores léxicos como la extensión del conocimiento léxico, su profundidad y el acceso a ítems léxicos fundamentales. Los predictores más fuertes de la categorización resultaron ser la imaginabilidad de las palabras, la frecuencia, la diversidad léxica y la familiaridad de las palabras. Los investigadores concluyen que los índices automatizados pueden ser predictores de la *competencia léxica* de aprendices de segunda lengua basándose en muestras de escritura. Sostienen que sus

hallazgos apoyan la relación entre la competencia léxica y la competencia lingüística general. Las conclusiones de este estudio, además de su diseño metodológico, fueron orientadores para el presente trabajo de tesis.

En resumen, la presente investigación se ubica dentro de una tradición existente dedicada a estudiar las relaciones entre los usos del lenguaje y la calidad de la escritura en el ámbito académico. Por medio de la revisión de literatura previa se ha podido enfocar de manera precisa el léxico como un aspecto lingüístico que requiere mayor profundización para conocer su incidencia en la construcción y evaluación de textos académicos.⁴

⁴ En el Anexo A se presenta la Tabla 19, que sintetiza la información de esta sección de la tesis.

3. MARCO TEÓRICO: CRITERIOS PARA ORIENTAR LA INVESTIGACIÓN

El propósito de este capítulo es presentar de manera puntual una serie de conceptos que ayudaron a delimitar y fundamentar el objeto de estudio: el léxico en el dominio de la escritura académica, así como a enfocar el tipo de texto cuyas características lingüísticas se decidió enfocar problemáticamente para lograr el objetivo principal del presente estudio, que fue determinar qué características del léxico pueden ser indicadores efectivos del nivel de dominio de la escritura académica, y jerarquizar su importancia para futuros tratamientos didácticos y de evaluación.

Para cumplir el propósito indicado, la exposición de conceptos en este marco teórico se guió por tres criterios: uno *social*, que da cuenta de los fines de los actos discursivos que se realizan entre sujetos en un contexto dado; uno *epistemológico e interactivo*, que considera los tipos de conocimiento que son producidos y validados por medio del discurso en esos contextos y por esos sujetos interactuantes; y uno *lingüístico*, que ayuda a caracterizar los recursos del sistema de la lengua que particularizan los textos que funcionan como artefactos en esas interacciones discursivas.

Parodi (2017a, 2018) y Navarro y Brown (2014) orientan dichos criterios con algunas preguntas: ¿dónde se produce y dónde circula el discurso?, ¿quién lo produce, para quién, para conseguir qué? y ¿qué opciones del sistema de la lengua realiza ese discurso? En las secciones que siguen se enmarca el objeto de estudio con apoyo en estas orientaciones.

3.1. Primer criterio: Social

3.1.1. El discurso en el contexto social

Desde una perspectiva funcionalista, el lenguaje es entendido como una herramienta semiótica que se usa para conseguir determinados fines en contextos sociales (Halliday, 2013). Las teorías funcionalistas “sostienen que la naturaleza del lenguaje se explica a partir de los objetivos sociales que guían su uso [...] son teorías interesadas por entender las relaciones entre los contextos socioculturales y situacionales del uso lingüístico y las características propias del lenguaje” (Navarro & Brown, 2014, p. 55).

El funcionalismo postula que la interacción entre lenguaje y contexto construye significados dentro del sistema de una lengua; por ello, se enfoca en las opciones lingüísticas sistemáticamente disponibles para los hablantes en diferentes contextos de uso. De este enfoque teórico se desprende que los enunciados orales y escritos reflejan las esferas comunicativas donde estos se producen, y que, a la vez, las opciones lingüísticas disponibles en el sistema ayudan a configurar esos contextos.

Considerada en este marco, la escuela es un contexto institucional en que los estudiantes se familiarizan con formas de usar el lenguaje diferentes a las que conocen, y, concomitantemente a esta inculturación, aprenden nuevas formas de pensar. En ese contexto formal, los alumnos estudian los contenidos escolares de manera imbricada en el discurso, por medio de los géneros que contribuyen a construirlos; los miembros de la institución esperan de ellos que lleguen a dominar a su vez algunos de esos géneros, y les evalúan con base en la forma en que expresan sus aprendizajes a través del lenguaje. Por todo ello, se puede decir que la escolarización es un proceso primordialmente sociolingüístico (Schleppegrell, 2004).

A partir de este posicionamiento sociocognitivo, interaccionista y funcional ante el discurso, se considera que entre todas las prácticas sociales que se realizan en las comunidades humanas existe un subtipo que se caracteriza por el uso situado y particular de ciertos “artefactos comunicativos letrados” (discursos escritos) que las personas utilizan

para conseguir objetivos relevantes en sus vidas. Estos artefactos son el resultado de una práctica histórica y cultural situada en cada comunidad. El manejo de esos artefactos letrados y de las prácticas letradas en que estos se usan se relaciona con la construcción de identidades sociales y con el ejercicio del poder al interior de esas comunidades (Cassany & López-Ferrero, 2010).

Las prácticas de escritura son múltiples y tienen lugar en cualquier grupo social que comparta propósitos comunicativos específicos. Una comunidad específica en la que se realizan prácticas letradas es la academia, donde se configuran diversas culturas disciplinares (Navarro & Revel, 2013). Aunque el ámbito académico ha sido concebido tradicionalmente como el contexto característico de aprendizaje, en el que se desarrollan prácticas letradas que usualmente van sofisticándose y especializándose conforme se incrementa la complejidad del conocimiento, no es el único espacio donde se realizan; de hecho existen prácticas de lectura y escritura en casi todos los ámbitos de la vida.

Existen manifestaciones vernáculas de prácticas letradas, más asociadas a los espacios personal, familiar y cultural donde se desenvuelven las vidas de las personas; estas son privadas y autogestionadas, a diferencia de las prácticas académicas, que adquieren un carácter dominante y son estandarizadas y reguladas por las instituciones educativas (Cassany & López-Ferrero, 2010). A pesar de esta diversidad, las concepciones dominantes asociadas a la academia, acerca de lo que significa leer y escribir, están tan interiorizadas que no es fácil “despedagogizarlas” (Street, 1995).

Atendiendo el carácter situado de los usos del lenguaje y de la escritura, diversos autores han identificado distintos tipos de discurso: académico, profesional, especializado o científico, entre otros atributos. El *discurso académico* se puede identificar como el lenguaje y las prácticas de discurso que se generan en los centros de investigación, las universidades, las escuelas y en cualquier contexto de educación formal en diferentes

niveles y disciplinas, y cuyos géneros son, correspondientemente, llamados *géneros académicos*. Este tipo de discurso tiene sobre todo propósitos persuasivos, divulgativos y didácticos; expresa credibilidad y prestigio como discurso autorizado; se usa para difundir conocimiento disciplinar, para apoyar la formación de los miembros de una comunidad de especialistas y para mediar las interacciones de esa comunidad (Parodi, 2007a).

El *discurso profesional*, por su parte, sería aquel que media en las interacciones del campo laboral, cuyas demarcaciones pueden ser trazadas por las diferentes prácticas de la vida productiva y económica, lo mismo que por las disciplinas que subdividen a estas prácticas, las cuales le imprimen una impronta característica, formando tradiciones y hábitos. Este discurso se realiza a través de *géneros profesionales* tan diversos como el memorando, el proyecto, el oficio, el plan de negocios o el informe técnico, que surgen en respuesta a las necesidades y fines de esas prácticas (Cassany & López-Ferrero, 2010; Navarro, 2012; Navarro & Brown, 2014; Parodi, 2008), y circula entre los integrantes (practicantes) de una misma profesión, pero también entre miembros de profesiones distintas y entre la comunidad profesional y un público lego (Parodi, 2007a).⁵

De acuerdo con Bolívar y Parodi (2015) y Parodi (2007a), el discurso académico y el discurso profesional forman parte de un constructo más amplio, una abstracción mayor que los contiene, llamado *discurso especializado*, que en términos generales se caracteriza por su alto grado de formalidad y su disciplinariedad, en oposición a las interacciones informales, y también en oposición a otras prácticas institucionales pero que no están relacionadas con el quehacer científico, académico o profesional.

⁵ Estos dos tipos de discurso se diferencian por sus contextos de uso y circulación, aunque existen puntos de contacto entre ellos: los textos profesionales plasman un conocimiento específico del mundo que fundamenta los propósitos y los hábitos de una comunidad profesional, por lo cual es viable entenderlos también como textos académicos; y, por otra parte, desde un punto de vista interactivo, el discurso de los académicos es discurso profesional porque estos docentes e investigadores son, a su vez, personas que actúan y cumplen fines en un medio laboral.

Algunos autores usan los constructos *discurso científico* (Parodi, 2008) o *discurso científico-académico* (Navarro & Revel, 2013) para designar al lenguaje y las prácticas de discurso realizadas por los investigadores –los *scholars* de la tradición anglosajona–, deslindándolo de cualquiera otra forma de discurso que se produzca o circule en el ámbito escolar.⁶ Su fin es generar y validar conocimientos a través de *géneros de investigación* (Swales, 2013), entre los que descuellan el artículo de investigación, el capítulo de libro, la conferencia y la ponencia (Castelló, 2007). La diferencia principal entre el discurso académico y el discurso científico radica en la prevalencia del andamiaje didáctico con fines educativos que caracteriza al primero, por el cual algunos autores le llaman también discurso pedagógico (Jarpa, 2015).

El criterio de este deslinde en este caso es interactivo y epistémico: existe otro tipo de discurso al interior de la academia que se produce con fines de enseñanza, aprendizaje y evaluación, y que los estudiantes de diferentes niveles tienen que aprender y dominar tan solo para transitar con éxito por el sistema educativo, pero no necesariamente para producir conocimiento nuevo y original (Navarro & Revel, 2013), como se verá más adelante.

Concediendo un valor práctico a estas consideraciones sobre distintos tipos de discurso, el presente estudio se adhiere a una definición amplia de discurso académico que abrigue tanto “el discurso de los académicos” hecho por investigadores (Bolívar, 2004), como la “lectura y escritura en la universidad”, hecha por estudiantes y profesores de licenciatura y posgrado (Bolívar & Beke, 2014), lo mismo que el “lenguaje de la escolaridad” (Schleppegrell, 2004), asimilado al “discurso escolar general” (Parodi, 2007a), que se configura por las prácticas de oralidad, lectura y escritura que se realizan desde la educación básica hasta la educación media superior.

⁶ “El discurso científico-académico se considera una variedad funcional del lenguaje con rasgos lingüísticos co-ocurrentes propios que circula en textos orales y escritos ligados a organizaciones específicas de las esferas científicas y académicas, y que presenta como objetivo general la producción de conocimiento, especificado a su vez en los objetivos de los géneros discursivos propios” (Navarro & Revel, 2013, p. 49).

Para precisar el objeto de estudio del presente trabajo de tesis, se propone usar el constructo *discurso académico de formación* para referirse a las prácticas de lectura y escritura con propósitos de aprendizaje que implican “interpretar la información de manera crítica, integrar información obtenida en fuentes diversas y crear información” (Cubo, 2014, p. 17), pero no necesariamente generar nuevos conocimientos. Estas prácticas se realizan a través de “discurso pedagógico” y se evalúan con “géneros evaluativos” (Jarpa, 2015).

Estos géneros comprenden textos que en todos los casos presentan una relación de asimetría porque en ellos los estudiantes tienen el rol de evaluado (Cubo, 2014). Incluso si son considerados como *géneros académicos de formación* (según la denominación de Navarro & Revel, 2013), su propósito intrínseco es evaluar las competencias, habilidades y dominio disciplinar adquirido por los estudiantes. En ellos interactúan un profesor (experto) y unos estudiantes (semilegos), en el marco de una situación comunicativa de carácter didáctico-pedagógico, al interior de un sistema de actividad académico, donde se emplean para realizar fines educativos (Jarpa, 2015).

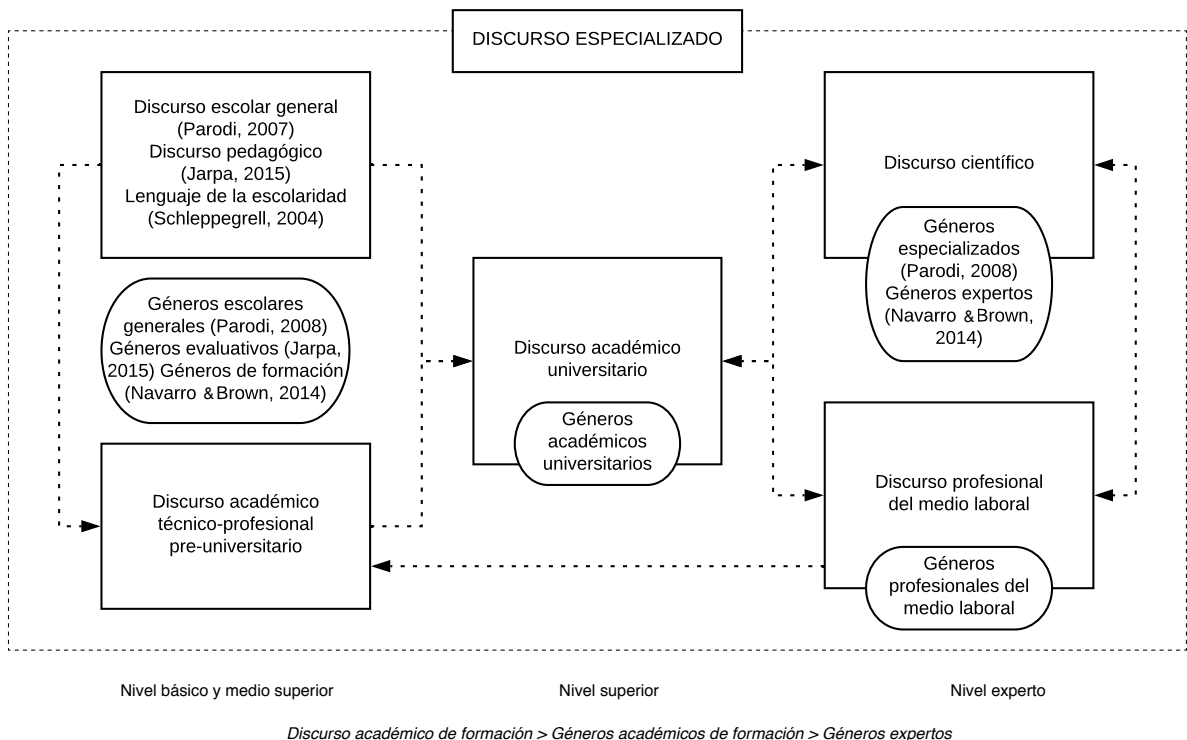


Figura 1. El continuo del discurso especializado y sus géneros.

Fuente: Elaboración propia con base en Jarpa (2015), Navarro y Brown (2014), Parodi (2007a, 2008) y Schleppegrell (2004).

La Figura 1 sintetiza esquemáticamente las comunicaciones y trayectos entre los contextos que hacen uso del discurso especializado y sus géneros. Tales contextos educativos constituyen sistemas de actividad que desarrollan prácticas letradas usando discursos provenientes del ámbito científico y que habilitan a sus estudiantes en la producción de los textos considerados relevantes por la comunidad por ser el puente hacia el ejercicio profesional y sus géneros específicos (Navarro, 2012, 2014).

En el diagrama de la Figura 1 se propone que el discurso especializado es un continuo que va aumentando el nivel de especialización, desde uno menos alto a uno más alto. La columna de la izquierda comprende las prácticas discursivas de la escolaridad básica y media, que junto con la columna central de la formación universitaria representan el campo en que se emplean el *discurso académico de formación* y los *géneros académicos de formación*, cuyas características ya se han explicado. A partir de la diversificación disciplinar que opera en la universidad, y a causa de la profundización en los conocimientos que implica la obtención un grado profesional, se puede distinguir, en la columna de la derecha, un nivel más complejo del discurso especializado: el nivel experto, con sus géneros.

3.1.2. Literacidad, eventos letrados y prácticas letradas

El concepto de *literacidad* ha sido definido de diversas maneras.⁷ Street (1984, 2005) explica que ha habido cuatro enfoques sobre el concepto: a) la *literacidad como aprendizaje*, que la asimila a la alfabetización básica llevada a cabo en la infancia; b) los

⁷ En el Anexo B se presentan las principales definiciones de literacidad según organismos internacionales, además de los indicadores considerados en el sistema educativo mexicano para el campo del lenguaje.

enfoques cognitivos, a los que subyace la idea de “la gran división” entre (pueblos e individuos) letrados y no letrados, asociada a los procesos de inculturación escolar basados en la escritura; c) el enfoque de los Nuevos Estudios de la Literacidad (NEL, en lo subsecuente), fundado en el concepto de *práctica social*, que considera las prácticas letradas como parte de contextos sociales y culturales, situadas históricamente e imbuidas por relaciones de poder (Barton & Hamilton, 2010), d) y, finalmente, la perspectiva de la *multimodalidad*, que postula la existencia de literacidades múltiples, un repertorio de prácticas dentro de las cuales quedarían incluidas las prácticas escritas convencionales (Anstey & Bull, 2018; Cassany, 2012; UNESCO, 2014).

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) define literacidad como: “the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated what varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve their goals, to develop their knowledge and potential, and to participate fully in their community and wider society” (Montoya, 2018). A este enfoque se adscribe la definición de literacidad como el conjunto de prácticas letradas articuladas entre sí que pueden asociarse a un contexto institucional específico, como la familia, la escuela, el trabajo, la iglesia, la comunidad, etc. (Zavala, 2009).

La Organización para la Cooperación y el Desarrollo Económico (OCDE), en la operación de un programa de seguimiento a las habilidades de adultos, define este mismo concepto como entender, evaluar, usar e involucrarse con textos para participar en la sociedad, para lograr objetivos propios y desarrollar el propio conocimiento y potencial (Montoya, 2018).

Dada esta variedad de perspectivas, es necesario especificar que en esta investigación se adopta la posición de los NEL y de las multiliteracidades (considerada en la definición

de la UNESCO), que establece que las prácticas letradas escolares se inscriben al interior de un gran espectro de prácticas en que se ven involucrados la escritura y los textos, en diferentes contextos sociales, realizadas por comunidades vinculadas por objetivos comunes.

Los conceptos centrales del enfoque de los NEL son *eventos letrados y prácticas letradas* (Barton & Hamilton, 2010; Street, 2005). Los primeros son cualquier ocasión en que las interacciones de los participantes involucran materiales escritos y por lo tanto implican procesos de interpretación de los mismos; estos eventos son repetitivos e incluso ritualizados, como sucede en cualquier práctica institucional. Las segundas hacen énfasis en el entendimiento de estas interacciones como actos sociales situados y cargados de ideologías, es decir, de concepciones sobre el papel de la escritura, sus efectos, sus metas, su valor o prestigio, etc., por lo cual algunas prácticas resultan más visibles e influyentes que otras (Barton & Hamilton, 2010). En función de estas nociones y de lo que implican socialmente, algunas prácticas letradas han sido dominantes y otras no dominantes, unas son percibidas como formales y otras como informales o vernáculas (Cassany & López-Ferrero, 2010).

De lo anterior se desprende que, en el contexto escolar, se llame *literacidad escolar o académica* a la capacidad de los estudiantes para adaptarse a las diversas situaciones en que se usa el lenguaje especializado que caracteriza los procesos de formación, y a las prácticas mediadas por ese lenguaje.

Según Bazerman (2005), *literacidad académica* es un término que combina la lectura y la escritura, pues estas nunca ocurren por separado; siempre son parte de un mismo campo de actividad. Además, se tiene que entender como una competencia transversal al currículo, ya que cada materia curricular tiene un lenguaje propio y en cada una se habla, se lee y se escribe para aprender (Álvarez, 2010).

Coincidentemente con lo señalado por Russell (1991), en el sentido de que cada nivel y organismo educativo desarrolla su propia literacidad, Bazerman (2005) distingue dos procesos de inculturación de las prácticas letradas dependiendo de los propósitos de cada etapa de la formación académica. Por un lado, una suerte de “iniciación a la escolaridad”:

The term *academic literacy* is most widely used in reference to the lower and middle grades of schooling, to distinguish the kinds of reading and writing students are expected to do in school from the kinds of reading and writing children might do in their daily life outside of school (Bazerman, 2005, p. 7, cursivas añadidas).

Por otro lado, se refiere a un complejo proceso de inserción a la vida académica como puerta de acceso al campo profesional:

Academic language socialization is the process by which individuals learn to enter into the discussions and gain access to the resources of academic disciplines through learning specialized language use and participating in academic activity settings. Learning to read and write in academic settings occurs through extended experiences in those settings, by meeting the expectations of those situations, and gaining from the opportunities for participation they offer (Bazerman, 2005, p. 8, cursivas añadidas).

Este segundo proceso es lo que Parodi (2008) llama *alfabetización disciplinar o especializada*, en un medio en que los géneros escritos constituyen “accesos al conocimiento y a las prácticas especializadas escritas” (p. 31). En el marco de la presente investigación, el carácter de “puente” de la escolaridad media superior hacia las “prácticas especializadas escritas” de la universidad sitúa el objeto de estudio (la escritura de estudiantes de bachillerato) como producto del primer proceso, por lo que, en adelante, el concepto de literacidad académica se restringirá al sentido expresado en la primera cita de Bazerman (2005, p. 7) respecto de la alfabetización escolar o académica, en cuyo interior se encuentra la competencia en escritura académica (Álvarez, 2010).

3.2. Segundo criterio: Epistemológico e interactivo

3.2.1.Literacidad académica, escritura y conocimiento

Como se ha explicado, pertenecer a una cultura disciplinar implica “participar de un conjunto de prácticas de lectura y escritura consensuadas” (Navarro & Brown, 2014, p. 67), y una participación exitosa en dicha comunidad, asimismo, exige el manejo competente de dichas prácticas. La cuestión del “manejo competente de esas prácticas” (Navarro & Brown, 2014, p. 67) –entre ellas la escritura académica– como constitutivas de la literacidad escolar, y el desempeño previsto en tales prácticas como indicador de logro, es el asunto que se aborda en este apartado.

Para comenzar, se retoma el continuo del discurso especializado empleado en contextos académicos esquematizado en la Figura 1, el cual dibuja un gradiente de menor a mayor nivel de especialización, para abordar las interacciones implícitas entre los sujetos que forman parte de esos contextos institucionales. Se trata de una sucesión de eventos letrados por los cuales un sujeto en formación debe transitar para llegar a construir un dominio discursivo en los ámbitos académico y profesional. Dicho recorrido representa un proceso de alfabetización permanente: desde la perspectiva de esa persona en formación, esta progresión es experimentada como un ambiente de inculcación disciplinar, a través del cual va enfrentando escenarios y géneros diversos que retan sus habilidades y conocimientos previos (Parodi, 2007a, 2008).

Al interior de dicho proceso, algunos de estos géneros solo serán leídos como medios de acceso a conocimientos específicos (Parodi, 2008); otros constituirán tareas de escritura que persigan el propósito de comunicar informaciones también específicas, pero algunos de ellos serán seleccionados por la comunidad como aquellos que el aprendiz debe dominar. Se trata de géneros “puente” a los géneros expertos (Navarro, 2012) que cumplen funciones relevantes en las prácticas sociales de esa comunidad.

En este mismo marco de ideas, y dado que el tema de interés de este trabajo es el dominio de la escritura académica por parte de estudiantes del nivel medio superior, es necesario tomar en consideración las exigencias el discurso académico en este nivel para comprender los desafíos que presentan a los estudiantes de bachillerato al estar sujetos al mencionado proceso de alfabetización académica disciplinar, especialmente en lo que se refiere a las prácticas de escritura.

3.2.2. Los géneros como accesos al conocimiento y como evidencia formativa

En primer lugar, hay que notar que en todos los niveles educativos circulan discursos con algún grado de especialización, incluso en los libros de texto, manuales y exámenes de la educación básica y media. Es un hecho que ese tipo de artefactos didácticos fueron escritos por expertos de diversas disciplinas para conseguir propósitos pedagógicos en sus usuarios, que son los aprendices. Esos textos son, en este sentido y contexto, textos “iniciáticos” que apoyan los fines de la escolarización (Cubo, 2014).

En segundo lugar, es preciso tener presente que entre los fines de las instituciones educativas está, justamente, ofrecer a sus miembros oportunidades de aprendizaje (y por tanto experiencia) en esta sucesión progresiva de eventos de inculcación disciplinar y profesional (con matices en los distintos niveles), por lo que en estos contextos cobran particular importancia las nociones de *construcción de conocimientos* y de *evaluación*. Tomando esto en cuenta, resulta extremadamente útil pensar en la existencia de unos *géneros académicos de formación* que conducen y sostienen el trayecto de los estudiantes hacia el dominio de una selección de géneros expertos, o, por lo menos, de unos *géneros evaluativos* que acreditarán su competencia a criterio de sus profesores.

Al estar inmersos en un ambiente institucional que valora la adquisición de prácticas epistémicas (como regular, dirigir, explicar, analizar y valorar el pensamiento) dentro de un determinado marco disciplinar, los estudiantes asimilan una cultura académica, es decir,

unos modos de socializar el conocimiento con sus pares y con sus profesores que van formando en ellos modelos de pensamiento y, concomitantemente, esa inculturación también modela el uso de un lenguaje adecuado a esos modos de razonamiento (Navarro & Revel, 2013).

Por ello es necesario enfatizar la importancia de la formación preuniversitaria en la adquisición de habilidades cruciales para la vida de las personas. Este nivel se caracteriza por ser una etapa preparatoria que, potencialmente, facilita el acceso de los estudiantes a la comunidad de los expertos; de hecho, adopta las divisiones disciplinares del ámbito universitario (Bazerman, 2005). Innegablemente, el dominio de las prácticas letradas más exigentes y especializadas de la educación superior tiene que entenderse como un trayecto largo cuyo arranque no puede situarse a las puertas de la universidad.

Al respecto, Russell (1991) puntualiza que en un sistema de educación masiva ni la academia es una comunidad unificada (sino más bien una colección de comunidades, de disciplinas profesionales que compiten, cada una con su propio discurso especializado), ni la educación secundaria y media son un bloque homogéneo. Hoy en día la formación preparatoria incluye un amplio rango de instituciones que forman y acreditan a los estudiantes para cumplir una variedad de roles sociales. De este modo, cada tipo de institución desarrolla, en palabras de este autor, “su propia literacidad”, sus propias expectativas, generalmente tácitas, de cómo deberían escribir los miembros de esa comunidad y sus estudiantes.

Así pues, se considera que en el contexto de la escolaridad se da una literacidad particular, pero que también al interior de ella hay múltiples literacidades, no solo en función de la diversidad de soportes y planos de lo escrito, que por sí mismos movilizan habilidades diversas para su manejo eficiente, sino también porque cada nivel educativo, cada modalidad escolar y cada disciplina alientan la multiplicidad de recursos cognitivos,

lingüísticos y retóricos de los que deben disponer los miembros de una comunidad académica para comunicarse y conseguir sus metas.

Desde el punto de vista de los NEL, se entiende que involucrarse en procesos de literacidad es siempre un acto social. En el caso de la literacidad escolar o académica, estos actos sociales consisten en una variedad de interacciones entre estudiantes y profesores, las cuales están revestidas por relaciones asimétricas de poder y se conducen por intereses de las partes. Por ejemplo, los intereses de parte de los profesores pueden estar orientados a procurar que los alumnos accedan a determinados conocimientos y demuestren sus aprendizajes, y por el lado de los estudiantes, la motivación puede ser aprobar materias, acreditar niveles académicos y alcanzar metas personales o profesionales.

Esta alfabetización, no obstante, se operacionaliza de diferente manera según el sistema y el nivel educativo, la modalidad, el plan de gobierno de un país, etcétera, por lo que resulta pertinente matizar el alcance de la literacidad en tanto alfabetización académica disciplinar o especializada como el “manejo experto” que se espera al final del nivel medio superior; asimismo, conviene recordar sus propósitos sociales y delimitar las nociones de “competencia” y de “experiencia” a lo que efectivamente se le puede exigir a un joven al término del bachillerato, para lo cual en el Anexo B se ofrecen las definiciones y los estándares de organizaciones internacionales y del sistema educativo mexicano respecto del concepto de literacidad que opera en programas regionales o locales, además de los atributos deseables en materia de lenguaje y comunicación que forman el perfil de egreso de la educación media superior en México.

3.3. Tercer criterio: Lingüístico

3.3.1. El constructo *texto académico*

Dado que esta tesis se enfocó al análisis de las producciones escritas de estudiantes de bachillerato para encontrar indicadores certeros del dominio de la escritura académica, interesa saber cuáles son las características lingüísticas esperables del texto académico que pudieran representar retos de consideración para los bachilleres al enfrentar tareas de escritura. Para ello, es necesario recuperar algunos aspectos de la relación discurso-género-texto que se han expuesto en los apartados previos y puntualizar el nivel de análisis lingüístico que operó en esta investigación.

Como se ha explicado, los textos son la realización concreta de unos géneros que a su vez configuran una entidad abstracta que es un determinado tipo de discurso. También es sabido que los géneros son constructos en los que interactúan una dimensión social, una dimensión cognitiva (epistémica) y una dimensión lingüística, que las enlaza y las vehicula (Parodi, 2008); por lo tanto, los textos que dan realidad a esos géneros no deben verse como meras “constantes estructurales” o “patrones textuales” que un estudiante ha de aprender a llenar.

En su manifestación concreta, los géneros son variedades de una lengua que operan a través de las tramas de un texto, y que se circunscriben lingüísticamente en virtud de propósitos comunicativos, participantes implicados (escritores y comprendedores), contextos de producción, ámbitos de uso, modos de organización discursiva, soportes y medios, etc. (Parodi, 2008, p. 26).

Estos criterios sociocognitivos y lingüísticos sitúan el discurso académico como una variedad funcional específica dentro de una lengua natural (Navarro & Brown, 2014), que se actualiza en un gradiente de textos de menor a mayor especialización, en función de la mayor complejidad de los significados que construyen para cumplir fines específicos (Parodi, 2008). En el entendido de que el texto académico cristaliza ciertas organizaciones retóricas por medio de recursos lingüísticos específicos, aprender a escribir este tipo de textos implica tener conciencia de estos recursos y aprender a controlarlos, para, de ese

modo, “poder formar parte de la disciplina con aportes propios” (Navarro & Brown, 2014, p. 77).

Hecha esta contextualización, para continuar es conveniente asentar algunos supuestos de partida, mencionados en los apartados anteriores: primero, que la escritura es el medio preferente mediante el cual se crea, fija y transmite el conocimiento disciplinar que configura un discurso especializado, cristalizado en géneros prototípicos que se realizan, a su vez, en textos concretos (Parodi, 2008); segundo, que no obstante la preeminencia de la escritura, todo texto es un artefacto multisemiótico que permite la construcción y el uso de esos conocimientos avanzados por medios diversos (Cassany, 2012; Lemke, 1998; Parodi, 2010); tercero, que como tal, el texto puede ser enfocado meramente como objeto lingüístico o bien como un artefacto semiótico implicado en eventos y prácticas sociales (Bolívar, 2007), y cuarto, que los textos funcionan como accesos al conocimiento pero también a las prácticas discursivas de una comunidad, y por ello, aunque la comprensión de los textos es fundamental, solo la producción escrita efectiva de los géneros requeridos por una comunidad particular puede revelar el máximo nivel de competencia discursiva de un sujeto experto en la disciplina (Parodi, 2008).

Dicho lo anterior, en la presente investigación se enfoca el discurso académico escrito y se atiende solo uno de los sistemas de significación que construyen el texto, esto es, el lenguaje verbal, cuyo análisis se dirige al nivel léxico-gramatical, específicamente al empleo del léxico en relación con la percepción de dominio de la escritura académica de un sujeto en formación por parte de expertos.

Así, en este trabajo el *texto académico* se conceptualiza como un artefacto lingüístico que media en eventos y prácticas letradas en el ámbito académico, aunque se acepta la co-participación de la imagen, la tipografía, la nomenclatura científica, la notación matemática, los esquemas, los enlaces activos a las redes y cualquier otro sistema simbólico

en la construcción del discurso escrito, de acuerdo con las perspectivas de la multimodalidad y las multiliteracidades (Anstey & Bull, 2018; Cassany, 2012; Lemke, 1998); e igualmente se acepta que el contexto escolar no es el único en que suceden prácticas letradas y por tanto el texto cumple múltiples funciones sociales fuera de la academia.

3.3.2. Características lingüísticas del texto académico

Una vez situado el carácter del texto que interesa en esta investigación, enseguida se abordan sus características discursivas y los recursos lingüísticos que se emplean para realizarlas.

Un principio de estilo de la escritura académica es que todo lo que el lector necesita saber ha de estar explícito en el texto; por ello las ideas toman la posición central y la personalidad del escritor se desplaza al fondo, de modo que las afirmaciones y argumentos sean aceptados por el lector a causa de la lógica o la evidencia, y no por quién escribe. El esfuerzo por mantener a la persona del escritor en el fondo es observado a través de los medios por los que el autor se refiere a sí mismo, generalmente usando la tercera persona, el pronombre *se* impersonal o el plural, pues este distanciamiento aporta credibilidad al discurso (Monippally & Shankar, 2010).

La formalidad requerida para la escritura académica implica cambiar de registro y evitar las expresiones coloquiales, así como seguir de cerca las convenciones de los géneros. Una de las convenciones más destacadas es la forma de citar las fuentes o los puntos de vista de otros; el grado de rigor con que se hace referencia a las fuentes refleja la intención del escritor del texto académico de lograr la objetividad, no apoyarse en la opinión o la intuición (Beke, 2011).

Otros rasgos de la escritura académica son el tono tentativo, más matizado que categórico, que se observa a través de recursos atenuadores; también la claridad y la precisión, que se materializan por medio de las elecciones léxicas y sintácticas; por ello el lenguaje tiende a ser especializado y las oraciones y párrafos tienden a ser más largos y complejos que en otros tipos de escritura (Monippally & Shankar, 2010).

Para evitar ambigüedades e interpretaciones erróneas, y lograr la claridad en el mensaje, en los textos académicos se emplean organizadores textuales que guían al lector. Por un principio de economía e informatividad, se emplean sustantivos y frases nominales que concentren el máximo de información, por lo que hay un elevado número de nominalizaciones, más que en otros discursos no académicos (Cademártori, Parodi & Venegas, 2006; Parodi, 2007a), y una mayor densidad léxica. Para evitar la redundancia, en los textos académicos se cuida que no haya repeticiones, dando paso a un vocabulario diverso; por precisión, se eliminan los adjetivos vacíos; la sintaxis se controla para mantenerla en su orden habitual o canónico, y se emplean abundantes expresiones formulaicas y metalenguaje (Hyland & Tse, 2004).

En los textos académicos co-ocurren de manera sistemática rasgos lingüísticos particulares en torno a temáticas específicas no cotidianas que exigen experiencia previa. La función del lenguaje que predomina en ellos es la comunicativa referencial, y uno de sus rasgos prototípicos es el empleo de un léxico especializado. Se caracterizan por su alto grado de generalización y abstracción semántica; son textos altamente elaborados, utilizan un registro formal de la lengua y procuran ser objetivos (Parodi, 2007a).

Los lectores de textos académicos esperan que el escritor manifieste procesos ordenados en su escritura, que presente una tesis, una idea o una afirmación y la sustente con evidencia, en función de ciertos hábitos de construcción del conocimiento que dependen de la disciplina.

Different disciplines have different ways of viewing the world, gathering information, interpreting data, and organizing knowledge. Each discipline uses variations (subregisters) of academic language that branch off to accomplish these purposes. This branching starts during the school years, as students start to learn, for example, that thinking and talking in science differ from thinking and talking in history. These overlapping sub registers that branch off further and become more specialized as they form the highly technical languages of higher education and the professional world (Zwiers, 2008, p. 69).

Algunos de los recursos lingüísticos que se emplean para organizar el conocimiento disciplinar en los textos académicos son el encapsulamiento de hipótesis o afirmaciones en etiquetas nominales, y el empleo, como se ha dicho, de un léxico especializado (Halliday & Martin, 1993; Navarro & Revel, 2013). Además, ya se ha advertido que en algunas disciplinas los textos presentan información no verbal por medio de lenguaje matemático y nomenclaturas de alto grado de especialización, así como tablas, figuras, diagramas, etc. (Parodi, 2007a).

Dado que esta investigación se enfocó el uso del léxico como posible indicador de dominio de la escritura académica, a continuación se especifican algunas de las características léxicas del texto académico.

En este punto es pertinente señalar que el léxico es, junto con la morfosintaxis, uno de los principales niveles o componentes del análisis lingüístico (Alcaraz & Martínez, 1997). “El léxico en una lengua es el repertorio de palabras establecidas que los hablantes pueden usar cuando hablan y a las que pueden recurrir para entender lo que oyen” (Clark, 1995b, p. 2); aunque en principio los términos *léxico* y *vocabulario* pueden ser intercambiables, se suele usar *léxico* para nombrar la clase abierta de todas las palabras portadoras de significado disponibles en una lengua, mientras que *vocabulario* alude a una clase cerrada, una reducción del léxico para fines de análisis. Desde un enfoque paradigmático, en este trabajo se alude al léxico empleado por los estudiantes en sus textos como la serie de

elecciones realizadas por ellos en la construcción del discurso a partir del sistema, y no a la clase cerrada del vocabulario específico de cada texto.

En particular, el tipo de léxico que se espera en la escritura académica es especializado, técnico, denso; se espera que las entidades o sujetos sean nombrados explícitamente (es decir, que los participantes del discurso sean lexicalizados, no asumidos), y que haya abundantes nominalizaciones (Schleppegrell, 2001). Asimismo, en este tipo de textos se espera una mayor diversidad y densidad léxica con respecto a interacciones informales (Chafe & Danielewicz, 1987; Snow & Uccelli, 2009). También es esperable el uso de expresiones prestigiosas, propias de la comunicación formal, alejada de coloquialismos. Del escritor académico se espera que sus elecciones léxicas sean precisas, a diferencia de la ambigüedad que caracteriza al habla y la escritura coloquial; que elija emplear tanto términos técnicos como conceptos abstractos, en lugar de los conceptos del sentido común y del discurso no escolar (Snow & Uccelli, 2009).

Respecto del tipo de elecciones léxicas, Halliday y Martin (1993) destacan el recurso de la nominalización para condensar información en el discurso académico; por ello, la vasta mayoría de los términos técnicos son sustantivos, muchos de los cuales resultan de la abstracción de procesos verbales (p. ej., *lixiviación*, *precipitación*, *calcificación*, *inflación*, *descripción*, *narración*). Estos procesos se abstraen para usarlos como principios explicativos y clasificatorios en las diferentes disciplinas (p. ej., el término *erosión* condensa y clasifica una serie de procesos que tendrían que explicarse con una serie de pasos, los cuales, a su vez, condensan procesos previos y efectos posteriores: *deforestación*, *abrasión*, *sedimentación*, *desertificación*).

Ahora bien, es necesario destacar la complejidad cognitiva del recurso de la nominalización; al respecto, Halliday y Martin (1993) señalan que para muchos estudiantes la abstracción es un mayor problema que el empleo de tecnicismos, lo cual, dicen, está

directamente relacionado con las distintas tradiciones discursivas disciplinares. Diversos autores han evidenciado la complejidad de los procesos de abstracción que subyacen a la interpretación y producción de textos académicos (Bello, 2016; Fang, Schleppegrell & Cox, 2006; García, Hall & Marín, 2005). Esta complejidad tiene su origen en los distintos valores y funciones que cobran principal, pero no únicamente, los sustantivos, las frases nominales y las nominalizaciones en el tejido del discurso: no solo nombran sus entidades y participantes, sino que construyen significados de alto grado de especialización disciplinar, condensan información y cohesionan anafórica y catafóricamente segmentos textuales, además de aportar una carga modal, atenuante o evaluativa, a ciertas palabras y frases, formando un metalenguaje característico de este discurso (Jiang & Hyland, 2015, 2016; Hyland & Tse, 2004).

A la vez, la dificultad de reconocer, emplear y monitorear estos recursos con funcionalidades múltiples se relaciona con los diferentes tratamientos didácticos que se le dan al lenguaje en las materias escolares en los distintos niveles educativos. Halliday y Martin (1993) sostienen que mientras en unos casos la tradición de enseñanza de alguna disciplina facilita la adopción de términos como si fueran requisitos, en otros se pierde de vista la utilidad del término para construir todo un concepto complejo. Para ejemplificar esta idea, se puede suponer el caso de que los profesores de ciencias naturales no pueden evitar enseñar directamente los conceptos y términos que conforman el discurso científico; por ejemplo, al explicar la *precipitación*, aunque es requisito que se desarrollen explicaciones con lenguaje común, estas necesariamente desglosarán los procesos verbales concomitantes que harán aparente la relación proceso-nombre; o la *experimentación* en el método científico, que se introduce en un sistema ordenado de pasos y resultados, lo cual probablemente facilita a los estudiantes entender la relación entre acciones y resultados.

Halliday y Martin (1993) contrastan con ello los hábitos de enseñanza de los profesores de ciencias sociales o de humanidades, que generalmente no enseñan

explícitamente el papel de la nominalización en la abstracción de conceptos como su principal herramienta interpretativa.⁸ Un caso que podría ejemplificar lo que señalan estos autores es que ciertos procesos como la *migración*, o ciertas características de la población como la *movilidad*, que los expertos usan para clasificar la realidad social, quedaran insertos en el discurso escolar de las ciencias sociales sin un tratamiento didáctico específico en la clase que permitiera a los estudiantes inferir el recurso gramatical y su capacidad para etiquetar fenómenos complejos.

La importancia de destacar el empleo de términos técnicos como una parte esencial de toda escritura técnica y científica radica en el hecho de que cada campo del conocimiento y cada especialidad emplean típicamente un determinado léxico que transmite una variedad de conceptos especializados usando lenguaje técnico. Estos términos especiales comunican significados concentrados que han sido construidos a lo largo de periodos significativos de estudio en un campo, y su valor yace en la forma en que cada término condensa una masa de información en una sola palabra (Perelman, Paradis & Barret, 1998).

El carácter técnico o especializado de una palabra es un aspecto funcional de la misma, lo cual supone que en un determinado contexto una palabra realiza la función de condensar significados de un dominio especializado –y por lo tanto entra a formar parte del sistema de ese campo de conocimiento–, mientras que en otros contextos puede cumplir funciones diferentes. La estrecha relación del léxico técnico y especializado con las disciplinas y sus materias se refleja en las distintas proporciones de este tipo de léxico que

⁸ Es necesario matizar esta aseveración de Haliday y Martin (1993), porque precisamente en el problema central de esta tesis se asume que la falta de tratamiento sistemático y planificado del lenguaje académico disciplinar al interior de las materias escolares ha dificultado la adquisición de habilidades discursivas básicas y conocimientos sobre las opciones del sistema de la lengua a disposición de los estudiantes para adecuarse a las exigencias del contexto académico.

se han encontrado en corpus de distintas especialidades (Chung & Nation, 2004; Coxhead, 2017).⁹

Para Halliday y Martin (1993) la importancia de la terminología técnica en la escritura académica y científica radica en que gracias al uso de léxico especializado se pueden realizar taxonomías del mundo, de sus fenómenos y entes, mucho más precisas y comunicables que las clasificaciones del sentido común, que se basan en lo que puede ser observado directamente con los sentidos. La escritura técnica está fuertemente asociada a las disciplinas o campos de la ciencia; y la ciencia se aboca a entender cómo está organizado el mundo y cómo ha llegado a ser así, lo cual se comunica a través de explicaciones y desarrollo de procesos. Ambos modos del discurso requieren terminología especializada del campo para interpretar y comunicar eficientemente dichas taxonomías y procesos, lo cual implica dominar los crecientes grados de abstracción que particularizan el discurso académico.

Coxhead (2017), finalmente, hace énfasis en el reto que constituye el aprendizaje del léxico técnico especializado. Señala que ser capaz de expresar por escrito conceptos académicos de manera precisa, empleando un léxico apropiado y de manera fluida es una prueba de la pertenencia a una comunidad académica. Sin embargo, usualmente los estudiantes no adquieren este tipo de léxico fuera del ámbito académico y es posible que no lo aprendan solo a través de la lectura.

En resumen, el léxico que se emplea en la escritura de textos académicos se caracteriza por ser más variado, más informativamente denso y preciso en relación con otros usos del lenguaje coloquiales e informales, y por constituir un vocabulario especializado, de contenido disciplinar y más abstracto que el que los estudiantes usarían en

⁹ Chung y Nation (2004) encontraron que la variación va de un 20% a un 30% de términos técnicos o especializados en relación con el total de palabras en sendos corpus de lingüística aplicada y de anatomía respectivamente, que también difieren en el tipo de palabras que contienen.

sus interacciones cotidianas. Por ello, el empleo de un léxico con estas características puede ser considerado un indicador de dominio de la escritura académica.

En la Tabla 2 se comparan las características léxicas del lenguaje académico esperables en la escritura de textos, en contraste con el lenguaje coloquial informal.

Tabla 2. *Características léxicas del texto académico versus interacciones informales.*

Interacciones informales, coloquiales	Textos académicos
Baja diversidad léxica	Alta diversidad léxica
Baja densidad léxica	Alta densidad léxica
Expresiones coloquiales	Expresiones formales, prestigiosas
Ambigüedad	Precisión
Conceptos concretos y del sentido común	Conceptos abstractos y técnicos, especializados
Sujetos pronominalizados, asumidos como presentes, y participantes conocidos	Sujetos y participantes léxicos con abundantes nominalizaciones y frases nominales expandidas.

Fuente: Elaboración propia con base en Schleppegrell (2001) y Snow y Uccelli (2009).

Con base en las consideraciones vertidas en este marco teórico, se puede afirmar que el léxico es uno de los aspectos clave en la construcción y valoración de textos académicos durante la formación preuniversitaria porque revela la forma en que los estudiantes se van adecuando a las prácticas académicas valoradas por la institución educativa, por medio de un proceso de inculcación de carácter social, cognitivo y lingüístico. Una de las prácticas a las que mayor valor epistémico se le confiere en ese contexto es la escritura de textos que respondan a las expectativas de la comunidad académica, y entre esas expectativas se encuentra el manejo de un discurso especializado, diferente del que los estudiantes emplean en su vida cotidiana fuera del contexto escolar.

En la presente investigación se asume que el dominio de la escritura académica está vinculado al conocimiento y al uso en la escritura propia de unos rasgos lingüísticos que se aceptan como característicos de los géneros académicos, o al menos de aquellos que gozan de mayor prestigio como vehículos de comunicación de significados disciplinares. El manejo eficiente de esos recursos lingüísticos relevantes en el texto académico demuestra a los miembros expertos de una comunidad académica que el estudiante ha adquirido unas habilidades y unos conocimientos que le permiten participar de manera efectiva en las prácticas discursivas que sostienen y dan sentido a esa comunidad.

De este modo, los textos producidos en el ámbito escolar se presentan al investigador como artefactos lingüísticos con los que los estudiantes participan en determinadas prácticas letradas escolares, y a la vez son un reflejo de la literacidad alcanzada por ellos, pues, como se dijo antes, aunque la comprensión de textos académicos es importante para los procesos de escolarización e inculcación disciplinar, es la capacidad de escribirlos adecuadamente la que se considera el rito de paso a la academia, a la disciplina y, más adelante, a la universidad y al mundo laboral.

En síntesis, la escritura es una de las competencias que se desea formar en el ámbito académico. Su dominio está relacionado con la competencia en comunicación lingüística de los estudiantes, pues estos requieren utilizar el lenguaje no solo para comunicarse oralmente y por escrito, sino para representar, interpretar y comprender la realidad, para construir conocimientos, organizar y regular el pensamiento, las emociones y la conducta (Álvarez, 2010). Por todo ello, el dominio de la escritura académica, como parte de la literacidad escolar –esta competencia tan amplia y decisiva para los estudiantes–, requiere apoyo pedagógico planificado y sistemático.

4. PLANTEAMIENTO DEL PROBLEMA Y OBJETIVOS DE INVESTIGACIÓN

Esta tesis se enmarca en la concepción del discurso académico como un uso particular del lenguaje que responde a necesidades y objetivos específicos al interior de un tipo particular de prácticas sociales. A su vez, las diversas prácticas sociales configuran comunidades a las que las personas pertenecen, por las que transitan o en las que son admitidas o excluidas con base en sus conocimientos o habilidades. Y, en el ámbito educativo, tales conocimientos y habilidades están sujetos a evaluación constante, principalmente a través del lenguaje (Beke, 2011; Schleppegrell, 2004).

En este sentido, las relaciones entre la construcción de saberes, la escritura y la evaluación de aprendizajes en el ámbito escolar son, desde los planteamientos de esta tesis, una cuestión de mayor relevancia para la pedagogía y para la lingüística aplicada.

En esta investigación se propone que el léxico es uno de los aspectos clave para la construcción y valoración de textos en el ámbito escolar porque, como señala Schleppegrell (2001), las opciones léxicas de los estudiantes en los textos dejan ver la manera en que se sitúan a sí mismos como miembros de una comunidad de discurso particular. Esto significa que el léxico empleado por los estudiantes puede revelar el alcance de su capacidad para distinguir entre el uso académico del lenguaje y los usos cotidianos y coloquiales, y poner en evidencia, concomitantemente, su habilidad para adecuarse a las exigencias de un tipo de discurso cada vez más especializado. Estas capacidades y habilidades son consideradas como integrantes de la literacidad escolar o académica.

Por todo lo anterior, el presente estudio se propone examinar qué aspectos del uso del léxico pueden servir como indicadores de dominio de la escritura académica, dominio que

puede llegar a ser determinante en la inclusión o exclusión de los jóvenes en el sistema educativo (Zavala, 2009). Se espera que los resultados puedan contribuir, primeramente, a los diversos procesos de evaluación de la escritura académica, de manera más general; y de manera más específica, a la enseñanza del léxico académico. La motivación central del presente estudio es contribuir, desde la lingüística aplicada, al desarrollo de acciones didácticas que promuevan la competencia en escritura académica en los estudiantes de bachillerato, quienes se preparan para estudios universitarios en donde el dominio de esta práctica de lenguaje es fundamental.

4.1. El problema

Como se explica en el marco teórico del presente estudio, el discurso especializado comprende los modos en que se usa el lenguaje en el ámbito escolar para la formación de personas, y también el empleado en el ámbito científico-académico para la comunicación y validación de conocimientos generados por profesionales de alguna disciplina, además del empleado en algunas áreas del campo laboral. Estas esferas de la actividad humana se comunican entre sí y se retroalimentan (Bolívar y Parodi, 2015; Parodi, 2007a, 2008).

Algunos aspectos del uso del lenguaje que particularizan al discurso especializado se relacionan con el uso del léxico: este es variado, emplea expresiones formales y socialmente prestigiosas, busca la precisión y es informativamente denso; además, emplea terminología especializada, así como conceptos abstractos (Halliday & Martin, 1993; Schleppegrell, 2004; Snow & Uccelli, 2009; Zwiers, 2008). Este tipo de discurso difiere de la forma en que se emplea la lengua con otros fines, por ejemplo, para las interacciones cotidianas familiares, privadas o públicas; para la convivencia social y la acción civil, para el comercio, el culto, la comunicación de masas, el arte, etcétera, con las que el individuo se familiariza en contextos no escolares.

Por las razones antes expuestas, algunos autores llaman lenguaje científico, lenguaje académico o lenguaje de la escolaridad a los usos de la lengua con propósitos específicos en el ámbito académico (Navarro & Revel, 2013; Schleppegrell, 2004); y proponen la existencia de sub-registros configurados por las distintas disciplinas, en razón de que los contenidos, las actividades y los objetos mismos de esas prácticas añaden especificidad y complejidad a estos lenguajes (Martín, 2004; Zwiers, 2008).

A esta complejidad y especificidad se agrega, además, el uso predominante del medio escrito para las interacciones entre los miembros de comunidades de discurso altamente especializadas, cuyas producciones se configuran en géneros propios, algunos de los cuales circulan transversalmente: de unas ramas de la ciencia a otras; de las ciencias a la vida académica, a la enseñanza y a la preparación para la vida laboral (Russell, 1991; Swales, 1990, 2013).

En el entendido de que “aprender lengua es aprender a usarla de forma adecuada en cada uno de los ámbitos de la actividad humana” (Camps & Castelló, 2013, p. 19), la literacidad escolar o académica implica aprender a hablar, a leer y a escribir, en cada etapa de la escolarización, de maneras que desafían los conocimientos y habilidades previamente adquiridos por los estudiantes.

En el caso de la escolaridad media superior, a medida que los estudiantes transitan por la currícula, entran en conocimiento de una gama de vocabulario disciplinar, consistente en un léxico influenciado o incluso determinado por las materias, cuya fuente es la escuela pero que proviene de las distintas ramas de la ciencia y la tecnología, incluyendo en ciertos casos el discurso profesional. Las prácticas letradas de la enseñanza media conducen a los estudiantes, a su vez, a emplear en diversas medidas ese léxico nuevo para la composición de textos escritos que permiten a los profesores y a otros agentes educativos verificar sus aprendizajes.

Esta ineludible relación de la escritura con la evaluación en el ámbito educativo es el centro del problema que en esta tesis se ha querido afrontar, pues, aunque existe una base firme de conocimiento lingüístico en el campo del discurso especializado y de la escritura académica, la valoración de textos escolares –específicamente en el bachillerato, aunque no únicamente en ese nivel educativo– sigue siendo, la mayoría de las veces, una acción subjetiva, por más que se haga uso de instrumentos y pautas (Navarro, Ávila & Gómez, 2019).

Por otra parte, está documentado que la literacidad escolar en general se halla en un estado de estancamiento o incluso de retroceso a nivel mundial, a pesar de que ha sido el centro de atención de las instituciones públicas internacionales dedicadas a la educación por al menos cincuenta años, a causa de su importancia para el desarrollo social y económico de los pueblos (OCDE, 2000; UNESCO, 2005, 2008), por lo que este problema tiene relevancia no solo para el caso mexicano, aunque en este trabajo se enfoca este contexto.

4.2. Delimitación y alcances prácticos de la tesis

Con base en lo anterior, para este trabajo fue importante observar lo que un conjunto de jueces expertos identificó como un buen desempeño en los textos académicos producidos por estudiantes de bachillerato, y relacionar esta percepción de dominio con el léxico que despliegan dichos estudiantes en la construcción de sus textos. Para el análisis se eligió un tipo de texto que es considerado como género de formación de carácter evaluativo, en razón de que fue producido como evidencia de aprendizaje en una materia escolar y porque fue escrito empleando una pauta que expresaba los estándares esperados para esa producción textual específica, entre los cuales se hallaba un uso particular del lenguaje.

Se ha tomado la decisión de enfocarse a la escritura de alumnos del último semestre del bachillerato porque, aunque la escritura académica no es un reto exclusivo de los

jóvenes, esta etapa de la escolaridad es considerada crucial por su carácter preparatorio, tanto para la vida de los estudiantes como para su ingreso y permanencia en la universidad.

Se ha tomado el léxico en la construcción y evaluación de textos académicos porque, como se planteó en la Introducción, uno de los aspectos que han sido reportados como problemáticos para los estudiantes de bachillerato es el manejo de un vocabulario suficiente para comprender lo que leen en las materias escolares y para escribir los textos que sus profesores les solicitan o que les son exigidos en pruebas estandarizadas de acceso a la universidad (ANUIES, 2014; Backhoff, Larrazolo & Tirado, 2011; CENEVAL, 2017; Zamudio, 2016).

Además, el dominio léxico y la competencia de escritura académica son requisitos que se infieren del perfil de egreso del bachillerato en México (SEP, 2016, 2017a), y que se desprenden de los estándares internacionales a los que se aspira en este país, de cara al mejoramiento de la calidad de la educación (Montoya, 2018; OCDE, 2019).

En síntesis, esta investigación se propuso contribuir al conocimiento general del discurso académico, de sus características lingüísticas, de los retos que estas les plantean a los estudiantes de bachillerato y en particular acerca del papel del léxico en el dominio de la escritura académica. Su intención fue obtener evidencias de la viabilidad de algunos aspectos del léxico como indicadores del nivel de dominio de la escritura académica, que puedan orientar la evaluación de la escritura académica, así como futuras aplicaciones pedagógicas.

4.3. Objetivos del estudio

El objetivo general de este trabajo fue determinar qué características del léxico pueden ser indicadores efectivos del nivel de dominio de la escritura académica, para poder proponer una jerarquía de su importancia en la valoración del texto académico, por medio del análisis

de las relaciones entre el léxico desplegado en un corpus de textos producidos por estudiantes de bachillerato y la variación en las calificaciones que les fueron otorgadas por un conjunto de jueces expertos.

Para lograr este objetivo general se plantean los siguientes objetivos particulares:

I) Identificar en un corpus de textos producidos por estudiantes de bachillerato cuáles características en el uso del léxico (extensión, diversidad léxica, densidad léxica, diversidad nominal y densidad nominal) tienen potencial para predecir una variación en la calificación otorgada por jueces expertos, y en qué condiciones se observa esa relación.

II) Determinar si las diferencias en el uso de sustantivos abstractos, de términos especializados y de sustantivos abstractos especializados en los textos tienen potencial para discriminar el nivel de calificación otorgada a los textos por los jueces.

III) Establecer, con base en la evidencia, una jerarquía de importancia de los indicadores del léxico en la valoración de los textos académicos para hacer recomendaciones acerca de su tratamiento didáctico como componentes de la escritura académica, acotadas al tipo de textos analizados y al nivel educativo en que estos se produjeron.

4.4. Preguntas de investigación

Se partió de una pregunta central: ¿Qué características del léxico en los textos escritos por estudiantes de bachillerato son indicadores efectivos del nivel de dominio de la escritura académica y en qué orden de importancia?

Para poder responderla, fue preciso contestar a las siguientes preguntas específicas que condujeron las distintas etapas del trabajo de tesis:

1) ¿Cómo interactúan las mediciones de extensión de los textos, de diversidad y densidad léxica, de diversidad y densidad nominal, y cuáles son sus efectos sobre el promedio de calificación otorgada a los textos por jueces expertos?

2) ¿En qué proporciones se encuentran los sustantivos abstractos y sus clases en el corpus y qué efecto tiene el uso de este léxico abstracto en los textos sobre la variación en la calificación de una muestra agrupada por su nivel de dominio de la escritura académica?

3) ¿Qué proporción de los sustantivos en el corpus son términos especializados de algún campo disciplinar y qué efecto tiene el uso de este léxico especializado en los textos sobre la variación en la calificación de una muestra agrupada por su nivel de dominio de la escritura académica?

4) ¿Qué proporción de los sustantivos abstractos en la muestra son a la vez términos especializados, y qué efecto tiene el uso de este léxico abstracto especializado en los textos sobre la variación en la calificación otorgada por jueces expertos?

5) ¿En qué orden se deberían jerarquizar las características del léxico para su tratamiento didáctico como componentes de la escritura académica en el bachillerato?

4.5. Hipótesis

La hipótesis general de esta investigación fue que los textos mejor evaluados –y por tanto aquellos que los jueces expertos consideraron evidencia de un mayor dominio de la escritura académica por parte de los estudiantes– serían aquellos que emplearan el léxico de una forma más cercana a las convenciones académicas, es decir, con una mayor densidad, variedad, abstracción y especialización, en concordancia con las caracterizaciones de este tipo de discurso revisadas en el capítulo dedicado al marco teórico.

De verificarse la hipótesis, las características del léxico analizadas en los textos demostrarían ser buenos indicadores de dominio de la escritura académica y un análisis

detallado de la forma en que se correlacionan con una buena evaluación proveería datos suficientes para establecer una jerarquía entre ellos.

Esta jerarquía expresaría el diferente peso o influencia que tiene cada uno de estos aspectos –densidad léxica y nominal, variedad léxica y nominal, abstracción de los sustantivos y especialización del léxico empleado– en la evaluación de los textos. En la medida en que estos datos se obtuvieron de un corpus de textos escritos por estudiantes del último semestre de un bachillerato de alta exigencia académica y de un entorno favorecido en términos socioeconómicos, los niveles de calidad alcanzados por estos estudiantes se podrían asumir como rasgos deseables para ese momento del nivel escolar en que los jóvenes se encuentran preparándose para transitar a la universidad.

Con las reservas del caso, la jerarquía de indicadores léxicos podría servir para hacer recomendaciones en cuanto al espacio y al énfasis que debería tener su enseñanza en el nivel medio superior.

El cumplimiento del objetivo general de esta investigación puede abonar a un mejor conocimiento del papel del léxico en la construcción y evaluación de textos académicos, lo que podría contribuir a enfocar este importante componente de la escritura académica en futuros tratamientos didácticos.

5. METODOLOGÍA

En el presente capítulo se especifica el método, el tipo de investigación y se delimitan sus alcances; se describen la composición del corpus que fue objeto de análisis y el criterio de obtención de una muestra del mismo para realizar calas sucesivas con las distintas variables, las cuales también son definidas y operacionalizadas en el presente capítulo. Asimismo, se explican los procedimientos de obtención de los datos y se especifica el tratamiento estadístico dado a cada conjunto de ellos.

5.1. Método y tipo de investigación

Esta es una investigación aplicada, empírica, basada en corpus (*corpus-driven*); cuantitativa por la naturaleza de sus datos, no experimental (*ex post facto*)¹⁰ y de alcance correlacional (Kerlinger & Lee, 2002).

En su desarrollo se emplearon herramientas de análisis léxico automatizado que se irán definiendo a medida que se explican los procedimientos; también se usó un programa estadístico avanzado para realizar pruebas y gráficos, y se atendieron principios básicos de análisis del discurso, como lo es el observar cualquier recurso lingüístico siempre en su contexto de uso. La perspectiva desde la cual se codificaron e interpretaron los datos fue funcional.

El estudio se realizó en cuatro etapas, en las que se pusieron a prueba diferentes grupos de variables para atender el objetivo general del estudio, que como se expuso en el capítulo anterior, fue determinar qué características del léxico pueden ser indicadores de

¹⁰ Kerlinger y Lee (2002) definen este tipo de investigación de la siguiente manera: “La investigación no experimental es la búsqueda empírica y sistemática en la que el científico no posee control directo de las variables independientes, debido a que sus manifestaciones ya han ocurrido o que son inherentemente no manipulables. Se hacen inferencias sobre las relaciones entre las variables, sin intervención directa, de la variación concomitante de las variables independiente y dependiente” (p. 504).

dominio de la escritura académica. Las preguntas de investigación, que también fueron expuestas en dicho capítulo dedicado al planteamiento del problema, son retomadas como guía para la presentación de los resultados en el capítulo correspondiente.

5.2. El corpus y la muestra

5.2.1. El corpus

El corpus está formado por 120 textos académicos recolectados en una escuela de nivel medio superior, no elicitados para los propósitos específicos de esta investigación. Según lo definido en el marco teórico, los textos que aquí se abordan se consideran textos de formación (Navarro & Brown, 2014), más específicamente textos evaluativos (Jarpa, 2015), ya que fueron producto de una tarea de escritura realizada en situación escolar cuyo fin era evidenciar determinados aprendizajes, por lo que se sitúan en el marco del discurso escolar general (Parodi, 2007a) y de la escritura académica en el nivel medio superior. Como manifestación del lenguaje de la escolaridad (Schleppegrell, 2004), este tipo de textos incorpora –a la manera de rasgos deseables, de adecuación o de pertenencia– ciertas convenciones en la organización del discurso y ciertos usos del lenguaje que particularizan el discurso disciplinar o especializado (Parodi, 2007a). En tanto género, las producciones de estos estudiantes se pueden tipificar como comentario de texto, pues tuvieron como propósito explícito comparar dos obras literarias, analizarlas y dar una conclusión general.

Sus autores son estudiantes que en el momento de la recolección de los textos cursaban el último semestre del bachillerato en una escuela preparatoria privada del centro de México, dirigida a un estrato socioeconómico medio a medio-alto y de alta exigencia académica. Cada uno de los estudiantes (60 hombres y 60 mujeres, entre 17 y 18 años de edad) produjo un texto en condiciones de clase, disponiendo de una pauta de escritura¹¹ con

¹¹ La pauta de escritura original puede consultarse en el Anexo C.

los siguientes requisitos: mantener una estructura introducción-desarrollo-conclusión; comparar dos obras literarias leídas en la clase de español y literatura durante el semestre, con base en algunos términos de comparación tomados del libro de texto;¹² no enfocarse en la historia que transcurre en la obra literaria sino en los aspectos comparados; respaldar el análisis con un número mínimo de citas textuales referidas con el estilo APA; usar un lenguaje académico variado evitando repeticiones y ambigüedades, cuidar la ortografía y evitar la personalización. No había una extensión mínima requerida, por lo que las producciones difieren en cantidad de palabras.¹³

Al final, los productores donaron su texto de manera voluntaria e informada para este proyecto, el cual fue referido como un análisis de la escritura escolar.

Los textos fueron anonimizados, digitalizados y normalizados. Se les corrigió la ortografía,¹⁴ por no ser parte del objeto de estudio y con el fin de hacer eficiente el uso de analizadores léxicos automatizados. Se conservaron los títulos pero se eliminaron citas textuales¹⁵ y epígrafes, por ser escritura no propia de los estudiantes; además de referencias y bibliografía, codificadas en los textos mediante una simbología.¹⁶ No se intervino sobre la puntuación original de los textos.

¹² Estos términos de comparación eran *género, contexto, estilo, intertextualidad, escuela literaria, efecto, interpretación, influencia, motivo*, entre otros (Betancourt, 2009).

¹³ Los textos tienen una media de 820 palabras ortográficas después de la normalización (ver Tabla 8, Apartado 6.1.1).

¹⁴ Referente: RAE 21a. edición, que fue la norma común en la clase en que se produjeron los textos.

¹⁵ Se consideró cita textual aquella colocada entre comillas o separada del párrafo en un bloque, aun sin comillas. Las paráfrasis se conservaron por ser una reelaboración propia de una cita.

¹⁶ Simbología paratextual/intertextual: Cambio de párrafo ¶, Sin título (- -), Cita (&), Nota (*), Notas (**), Epígrafe (\$), Referencia (+), Lista de referencias (++), Metadata (=).

5.2.2. Calificación de los textos por cinco jueces expertos

Los fundamentos del método de evaluación de textos por jueces han sido documentados de manera extensa en la vertiente de los estudios del desarrollo y la psicolingüística –por ejemplo, Crossley y McNamara (2011), Berman y Nir (2009), Nir y Berman (2010), Berman, Nayditz y Ravid (2011), McNamara, Crossley y McCarthy (2009), Crossley, Salisbury, McNamara y Jarvis (2010)– y ha sido empleado también por pedagogos como Meneses, Uccelli, Santelices, Ruiz, Acevedo y Figueroa (2017), Figueroa, Meneses y Chandía (2018) y Figueroa (2018). Consiste en la valoración de textos –orales o escritos– por parte de expertos, para asignar una calificación bajo criterios definidos previamente con algún instrumento. Dependiendo de los intereses de cada investigación, los criterios se amplían o se restringen para enfocar la evaluación en determinados aspectos, como en la calidad global del texto, el estilo retórico, el vocabulario, la arquitectura sintáctica, etcétera, o bien, como en este caso, en la percepción de un determinado nivel de dominio de la escritura académica.

El procedimiento para la calificación de los textos fue el siguiente: Primero, se adaptó al español una rúbrica de los investigadores McNamara, Crossley y McCarthy (2009), que evalúa el dominio de la escritura académica. La primera versión de la adaptación fue revisada por dos lingüistas; las modificaciones de fondo y de forma hechas a raíz de sus observaciones constituyeron la segunda versión. Luego, esta fue sometida a una prueba piloto con tres profesores de educación media con experiencia en el uso de rúbricas y otros instrumentos de evaluación, durante un periodo de una semana, sobre 12 de los textos del corpus. Estos jueces piloto no participaron en la aplicación subsecuente. Con base en sus observaciones se revisó la redacción de cada nivel de la rúbrica para lograr la versión final,¹⁷ que consiste en una escala con seis niveles de dominio de la escritura académica,

¹⁷ En el Anexo D se adjuntan tanto la rúbrica original de McNamara et al. (2009) como las adaptaciones subsecuentes.

donde 1 es el nivel más bajo y 6 el más alto. Este último considera como dominio “notablemente consistente” el manejo de los aspectos que se citan a continuación. Los subrayados y negritas son de la versión utilizada con los jueces.

Puntuación de 6. Demuestra un **dominio notablemente consistente de la escritura académica**, aunque puede tener algunos errores menores. El texto se **estructura** a partir de una introducción en que presenta el propósito del escrito y plantea de manera explícita un punto de vista, que es desarrollado y sustentado utilizando ejemplos apropiados, razones y otras evidencias para apoyar su posición, misma que recupera en la conclusión. Muestra una **organización claramente enfocada al desarrollo del tema**, demostrando coherencia y una eficiente progresión de las ideas. Exhibe un **uso hábil del lenguaje**, utilizando un vocabulario variado, preciso y adecuado al propósito del texto. Demuestra una habilidad gramatical notable que se muestra en la estructuración y conexión de las oraciones, misma que favorece el significado.

Los niveles de 5 a 2 describen actuaciones menos eficientes de los mismos rasgos. El dominio más bajo se expone enseguida en el nivel 1:

Puntuación de 1. Demuestra **poco o ningún dominio en la escritura académica**, cuya calidad se ve severamente afectada por varias de las siguientes debilidades: aunque el texto muestre cierta **estructura** no desarrolla ningún punto de vista viable sobre el tema; proporciona poca o ninguna evidencia para apoyar una posición, que puede no ser explícita, o no hay una conclusión claramente definida. El texto está **desorganizado** o desenfocado, resultando en un escrito inconexo o incoherente. Presenta una visible **dificultad en el uso del lenguaje**, con errores importantes en la variedad, la elección y el uso del vocabulario. Demuestra severas fallas en la estructura de las oraciones y contiene errores gramaticales que persistentemente interfieren con el significado.

La estrategia de aplicación del instrumento se estableció luego de verificar el nivel de confiabilidad (acuerdo) entre los jueces piloto, para lo cual se corrió la prueba estadística

Intraclass Correlations Coeficient (ICC)¹⁸ con el paquete Psych en R.¹⁹ Según Koo y Li (2016), el parámetro de confiabilidad de esa prueba es el siguiente: con un ICC menor a 50% la confiabilidad es pobre; entre 50% y 75% es moderada; entre 75% y 90% es buena, y más de 90% es excelente. Estudios previos de corte experimental que emplean el método de evaluación por jueces proceden con un 80% de acuerdo entre ellos, incluyendo dos entrenamientos en el uso de la rúbrica de evaluación. Para alcanzar un nivel de aceptable a bueno en la confiabilidad, es decir, un ICC_{2k} de 0.70 a 0.80 con alfa de 0.05, se necesitan al menos 4 jueces que evalúen más de 20 textos a la vez.

Esto condujo a las siguientes decisiones metodológicas: 1) los mismos cinco jueces debían valorar todos los textos; 2) debían ser profesionales en el área educativa y su perfil debía pertenecer al campo de las humanidades;²⁰ 3) debían estar en servicio activo, con al menos un año de experiencia frente a grupo; 4) habrían de tener experiencia en el manejo de instrumentos de evaluación, para que aun sin entrenamiento fueran capaces de emitir un juicio objetivo; 5) los textos se les habrían de entregar en archivo electrónico con formato de documento portátil (PDF), completos, con la ortografía normalizada y toda la configuración homogeneizada en cuanto a tipografía, cuerpo y márgenes, para evitar que características diferentes de los textos interfirieran en la valoración; y, finalmente, 6) los

¹⁸ Este coeficiente calcula la confiabilidad entre una cantidad k de jueces; la prueba para obtenerlo fue diseñada por Shrout y Fleiss (1979) sobre la base de un análisis de varianza de una vía. Sus resultados se interpretan así: un número k de jueces, evaluando los mismos n textos, acordarían en x porcentaje en evaluar del mismo modo un determinado texto. Con ello no se buscó medir la viabilidad del instrumento empleado, sino calcular el grado de acuerdo entre los jueces, que deben poder ser remplazados por otros en tanto son una muestra representativa de una población con perfil semejante.

¹⁹ R es un entorno de software libre para computación y gráficos estadísticos. Se compila y se ejecuta en una amplia variedad de plataformas: UNIX, Windows y MacOS. Para esta investigación se usó el sistema iOS. Sostenido por ©The R Foundation, se puede descargar de <https://www.r-project.org>

²⁰ En el Sistema Educativo Mexicano no existe una formación única estandarizada como requisito para ejercer la docencia en el Nivel Medio Superior. Del espectro profesional de un profesor de bachillerato, se eligió el campo de las Humanidades porque a los profesores de esta área se les encomienda tradicionalmente la enseñanza de las materias relacionadas con el lenguaje, la comunicación y la literatura. El perfil completo de los jueces puede ser consultado en el Anexo E.

textos tendrían que estar anonimizados con una clave secundaria (diferente a la codificación inicial que conservaba metadatos relevantes para la investigadora), sin pistas sobre la identidad del autor, su sexo, edad o cualquier otro dato personal o escolar, por lo que se decidió usar un número consecutivo simple. Los cinco jueces recibieron una gratificación económica por su colaboración.²¹

Después de la calificación por el grupo de jueces expertos, las puntuaciones obtenidas fueron validadas con la misma prueba antes mencionada.

Vale aclarar que en el estudio de McNamara, Crossley y McCarthy (2009) los jueces recibieron entrenamiento en el uso de la rúbrica. Para el presente trabajo no hubo entrenamiento, sino que se buscaron profesores con un perfil de experiencia en el uso de rúbricas, para que la calificación reflejara la percepción y el criterio propios del docente sin mediación de la investigadora.

5.2.3. La muestra

La muestra se compone de 60 textos del corpus: los 30 de nivel más alto y los 30 de nivel más bajo en el promedio de las calificaciones otorgadas por jueces expertos, a partir de su valoración del dominio de la escritura académica manifiesto en el texto. Los niveles alto y bajo se determinaron a partir de la mediana del promedio general (3.8), según la escala de 6 puntos que se explicó en el apartado previo.

²¹ El Código de Bioética de la Universidad Autónoma de Querétaro considera entre los requisitos de los proyectos de investigación que se informe sobre los “incentivos económicos” dados a los sujetos o participantes en la investigación. En el Anexo F se citan los postulados atinentes a las remuneraciones no coercitivas, en el apartado “Retribución aceptable”, en que entra la gratificación recibida por los jueces, dado que fue una cantidad de dinero relativamente pequeña (menor a un 10% del monto de la beca de un estudiante de doctorado a finales de 2017); además, ninguno de ellos era estudiante o becario; todos tenían empleos remunerados en ese momento, por lo que no entran en el rubro de población vulnerable.

5.3. Variables del estudio

Partiendo del objetivo central de esta investigación, que fue determinar qué características del léxico en los textos producidos por estudiantes en situación escolar pueden ser indicadores eficientes para determinar el nivel de dominio de la escritura académica, hemos tomado como base otros trabajos que han perfilado bien los rasgos del texto académico que resultan relevantes para observar su especificidad, y que, asimismo, son útiles para relacionarlos con la percepción de su calidad –como expresión de un nivel de dominio– por parte de un conjunto de jueces expertos.

Como se expuso en los capítulos dedicados a los antecedentes y al marco teórico, existen estudios que relacionan la calidad del texto académico con indicadores sintácticos, pragmáticos, discursivos y léxicos en conjunto; para los fines de esta investigación se tomaron primeramente las descripciones generales del lenguaje académico de Schleppegrell (2001), Snow y Uccelli (2009) y Uccelli, Dobbs y Scott (2012), que cubren todo ese espectro; y luego, para poder enfocar el análisis de manera particular en el léxico, se acudió a descripciones más puntuales como apoyo para la operativización de las unidades de análisis.

Enseguida se explica el fundamento para la construcción de un indicador de dominio de la escritura académica y para la delimitación de las variables predictoras: diversidad y densidad léxica (Ávila, 1991; Halliday, 1990; López, H., 2011; O’Loughlin, 1995; Ure, 1971), diversidad y densidad nominal (Johansson, 2008; Ravid, 2006), empleo de sustantivos abstractos (Bosque, 2011; Lyons, 1977; Schmid, 1999b), empleo de léxico especializado y de léxico abstracto especializado (Halliday & Martin, 1993; Martín, 2004; Zwiers, 2008).

5.3.1. Variable dependiente: el constructo *dominio de la escritura académica* (PromEA)

Nociones como *calidad del texto* o *dominio* de los componentes del lenguaje académico o bien, específicamente, de la escritura académica, son objeto de discusión en varios artículos, como ya se refirió (ver Apartado 2.1). Entre otros, Berman y Nir (2009) operacionalizan la calidad del texto como un conjunto de rasgos retóricos, gramaticales y léxicos que aportan estructura y eficiencia al texto de acuerdo con el género discursivo.

Por su parte, López, H. (2011) incluye en el constructo *calidad del texto* una serie de factores como la riqueza léxica, la madurez sintáctica, los esquemas de cohesión y la coherencia discursiva. Otros autores optan por remitir a juicio de expertos una selección de estas unidades de observación según su interés investigativo, de modo que al constructo *calidad* o *dominio* se le pueda adjudicar un nivel dentro de una escala estimativa respecto de un logro dado (Crossley & McNamara, 2011).

Fue con base en esta última posición que para el presente trabajo se estableció como indicador del nivel de dominio de la escritura académica la valoración objetiva del texto hecha por personas distintas al investigador.

Como se expuso en el marco teórico (ver Capítulo 3), en tanto escritura escolar el texto académico de formación, y específicamente el texto evaluativo, es considerado y tratado como evidencia de aprendizaje, por lo que, en un sentido pedagógico, su valoración incluye aspectos que, aunque estén mediados por el discurso, van más allá de él (hacia la comprensión o apropiación de los contenidos de la materia). En este tenor, y atendiendo al hecho de que los textos del corpus no fueron elicitados para la investigación sino que fueron producidos en respuesta a una tarea de escritura en situación de clase y con fines de acreditación curricular, la evaluación de los mismos tuvo que ser deslindada de su contexto

original de producción –el aula–, para dar confiabilidad al indicador de dominio como parámetro de investigación. Como ya se explicó en el apartado 5.2.2, esto se logró a través del distanciamiento por parte de la investigadora, precisamente para que la acción de los jueces expertos representara, por efecto de ese distanciamiento y algunos controles, una situación ideal de evaluación que permitiera un cierto grado de generalización de los resultados.

De este modo, la variable dependiente *dominio de la escritura académica* se definió en este trabajo como el puntaje (promedio o nivel) obtenido en una valoración criterial realizada por profesores experimentados, con base en una rúbrica, acerca del dominio de la escritura académica manifiesta en un texto producido en el ámbito escolar, y se le asignó el acrónimo PromEA (Promedio de Escritura Académica), por ser el promedio de las cinco calificaciones obtenidas por cada texto, una por juez, con base en una rúbrica (ver Apartado 5.2.2 y Anexo D). En la primera etapa del trabajo, PromEA se operativizó como variable numérica continua: un número racional con un decimal, y a partir de la segunda etapa como variable categórica: una posición en el nivel alto (por encima de la mediana de 3.8) o en el nivel bajo (por debajo de la mediana) del indicador.²²

5.3.2. Variables independientes (predictores)

5.3.2.1. Extensión del texto

Esta variable refiere al tamaño del texto y se operativizó de dos maneras: a) como *la cantidad total de palabras en el texto*, en donde *palabras* significa palabras ortográficas (cada token contable en el texto), según la definición de Johansson (2008), para salvar las diferencias establecidas por Halliday (1990) acerca de los ítems léxicos y los ítems gramaticales, útil en las mediciones de densidad; y b) como *la cantidad de cláusulas en el*

²² Los estadísticos completos de las mediciones hechas sobre el corpus pueden consultarse en el Apartado 7.1.1 de este trabajo.

texto, con base en la definición de cláusula de Berman y Slobin (1994): “We define a clause as any unit that contains a unified predicate. By unified, we mean a predicate that expresses a single situation (activity, event, state). Predicates include finite and nonfinite verbs, as well as predicate adjectives” (p. 660).

De este modo, la cláusula, como unidad de medida de la extensión de un texto, es una unidad de discurso que expresa una situación simple, con un solo sujeto, un solo predicado principal, una sola marca de tiempo/aspecto/modo (TAM) por predicado, una sola marca de concordancia de sujeto por predicado (en los verbos finitos) y una sola marca de polaridad (Palancar & Alarcón, 2007). Los criterios específicos de delimitación de la unidad *cláusula* fueron los siguientes:

a) Forman una sola cláusula los enunciados que cumplen los criterios expuestos por Palancar y Alarcón (2007), incluyendo los complementos y adjuntos intraclausales (Aissen, 2006) con la condición de referir a un solo sujeto, dentro de la definición de Berman y Slobin (1994) ya citada.

b) Forman una cláusula distinta los complementos clausales y los adjuntos de cláusula (Aissen, 2006), en tanto cumplen funciones oracionales formalizadas como oraciones ellas mismas (Bogard, 2004), en relación de subordinación a una predicación principal.

c) Los casos de predicación secundaria (Palancar & Alarcón, 2007), específicamente frases adjetivales y de gerundio, se consideran modificadores intraclausales y por tanto no se separan de su controlador en la predicación principal.

Aunque la extensión en palabras o en cláusulas no es un indicador léxico, la medida del texto se incluyó en razón de que para fines prácticos no es viable descartar la interacción de esa magnitud con el resto de las variables explicativas, según lo mostró un estudio de Berman y Nir (2009).

5.3.2.2. Diversidad léxica

Como se expuso en el capítulo dedicado al marco teórico (ver Capítulo 3), el texto académico se caracteriza por una alta diversidad léxica; esta se define como la proporción de palabras diferentes (tipos) respecto del total de palabras ortográficas (tokens) que constituyen un texto (López, H., 2011). Para la medición de esta característica del léxico en los textos existen varias magnitudes, las principales –TTR, VocD y MTL– se describen a continuación.²³

La *Type-token ratio* (TTR), que significa “proporción tipos/tokens” (Ávila, 1991; López, H., 2011), se puede expresar como proporción o como porcentaje indistintamente, pero su precisión ha sido cuestionada por su sensibilidad a las diferencias en la extensión de los textos, pues al aumentar la cantidad de palabras el TTR disminuye proporcionalmente, y viceversa (Johansson, 2008; McCarthy & Jarvis, 2010). En los estudios que han servido de antecedentes para la presente investigación no se encontró un parámetro de diversidad léxica en TTR para la evaluación de los textos académicos; sin embargo, se puede tomar como contraste la escritura infantil temprana ($TTR \geq 0.4$), la de adultos jóvenes ($TTR \geq 0.5$) y la que se presenta en textos profesionales como la prensa ($TTR \geq 0.6$), según un trabajo de Ávila (1991).

Vocabulary diversity (VocD) significa “diversidad de vocabulario”; representa la variedad de palabras diferentes que hay en un texto y es el resultado de una serie de cálculos hechos sobre muestras aleatorias de una cantidad determinada de tokens, sobre las

²³ En adelante se referirán las mediciones de densidad y variedad léxica y nominal por medio de sus siglas en inglés, ya que se trata de unidades de nombre estable y de uso muy frecuente en estudios de desarrollo y lingüística aplicada en el ámbito académico anglosajón, aunque se reportan también en publicaciones en español, como se vio en el capítulo dedicado a los Antecedentes (ver, por ejemplo, Figueroa 2018).

cuales se mide el TTR. Luego de una serie de muestreos de diversas extensiones, los analizadores automatizados como CLAN, Coh-Metrix o TextInspector²⁴ calculan una curva empírica de TTR sobre las medias almacenadas en los muestreos; por ello se considera que el índice VocD resuelve la inconsistencia del TTR ante textos de extensión desigual (McCarthy & Jarvis, 2010). Los valores van de 10 a 100 –aunque en la práctica no existe un límite determinado–; los valores más altos se consideran indicativos de una mayor diversidad.

Measure of textual lexical diversity (MTLD) significa “medida de diversidad léxica textual” (McCarthy & Jarvis, 2010). Se trata de un índice de diversidad léxica calculado de manera secuencial para salvar el sesgo producido por las diferentes extensiones de los textos. Los analizadores automatizados calculan la extensión media de cadenas secuenciales de palabras que mantienen un determinado valor de TTR (convencionalmente es .720); esto es, el índice MTLD resulta del cálculo de la media de todos esos segmentos, por lo cual su precisión baja al evaluar textos cortos (tokens<100).

Al igual que con VocD, los valores de MTLD cubren un rango del 10 al 100 sin que este sea un límite real. Los parámetros para la escritura académica en MTLD o en VocD son aproximados (>80), pero pueden entenderse mejor si se contrastan con otros datos disponibles, por ejemplo con la escritura de un adolescente aprendiz de segunda lengua (VocD=35-75), o con la de un adulto aprendiz de ESL (VocD=40-70), según un estudio de Malvern, Richards, Chipere y Durán (2004). En el presente trabajo se hicieron dos mediciones de diversidad léxica (TTR y VocD) pero se usó VocD en la estadística por dos

²⁴ TextInspector es un analizador automatizado de textos que hace diversas mediciones de extensión y diversidad léxica, además de etiquetar sintácticamente y extraer el vocabulario con base en corpus en inglés. Fue desarrollado por Stephen Bax y se localiza en línea en <http://textinspector.com>. Coh-Metrix es una herramienta computacional que produce índices de las representaciones lingüísticas y discursivas de un texto. Desarrollado por Arthur C. Graesser y Danielle McNamara, Coh-Metrix calcula la coherencia de los textos en muchas medidas diferentes. Se halla disponible bajo suscripción en <http://www.cohmetrix.com>. Por su parte, CLAN es un editor y analizador de interacciones orales que forma parte del proyecto CHILDES desarrollado por Brian MacWhinney, disponible en <https://childes.talkbank.org>.

razones: la primera por su estabilidad ante la variabilidad de tamaño de los textos y la segunda porque reportó resultados de alta confiabilidad en las pruebas de varianza.²⁵

5.3.2.3. *Densidad léxica*

La densidad léxica es otra característica del texto académico señalada en el marco teórico (véase Capítulo 3). Se trata de la proporción de ítems léxicos (palabras de contenido) sobre el total del discurso (O'Loughlin, 1995; Ure, 1971) y puede ser medida de varias maneras, ya sea sobre el total de palabras o sobre la unidad gramatical mayor, como la cláusula (Halliday, 1990); bien sea separando los ítems de alta y baja frecuencia lexical, o considerándolos todos sin importar su frecuencia.

Según el enfoque, la lengua en estudio y los objetivos de cada investigación, se consideran palabras léxicas todas las que portan contenido referencial y semántico –de clase abierta–, a diferencia de las categorías funcionales, que tienen contenido principalmente gramatical –de clase cerrada–, aunque entre las primeras, a menudo es excluido parcial o totalmente el adverbio (Berman & Nir, 2009, 2010).

La definición de densidad léxica que opera en este trabajo fue adaptada de Johansson (2008), quien a su vez se basa en Halliday (1990). Johansson propone como *densidad léxica* el total de palabras con propiedades léxicas dividido por el total de palabras ortográficas del texto. Acorde con dicha definición, fueron consideradas palabras léxicas los sustantivos, los verbos, los adjetivos y los adverbios, y el resultado se expresó como porcentaje. Como parámetros de comparación se pueden citar los trabajos de Ure (1971),

²⁵ De hecho, la confiabilidad de la medición VocD –observable en las constantes de los intervalos de confianza– en las pruebas estadísticas se debe precisamente a que el tipo de cálculo del que se desprende el VocD hace a esta medición de diversidad léxica una dimensión independiente del tamaño del texto. Es decir, se obtienen resultados fiables al comparar la diversidad léxica de textos de extensiones diferentes, lo cual no sucede con la medición TTR, que aumenta sus valores en textos cortos y viceversa, lo cual falsea los resultados de análisis de varianza respecto de cualquier otra variable.

quien encontró una densidad mayor en textos escritos (más de 40%) en comparación con interacciones orales (menos de 40%), y de Johansson (2008), quien identificó proporciones de 31% a 39% de densidad léxica en textos expositivos escritos de individuos de 13 a 17 años.

5.3.2.4. Diversidad nominal

La *diversidad nominal* en este trabajo se definió, con base en Ravid (2006), como la proporción de sustantivos diferentes presentes en el texto. Se calcula dividiendo la cantidad de tipos diferentes de sustantivos sobre el total de sustantivos en el texto y se expresa como porcentaje. Esta unidad se incluye entre los predictores léxicos de dominio de la escritura académica en razón de que es una variante de la diversidad léxica (Johansson, 2008) que puede interactuar con la densidad nominal en el fenómeno del encapsulamiento de información en el texto (Hess, 2013; Ravid, 2006), aunque en este trabajo solo se exploren sus efectos cuantitativos y no sus manifestaciones en el manejo del discurso por parte de los estudiantes.

5.3.2.5. Densidad nominal

Como se explicó en el marco teórico de este trabajo, la densidad nominal también ha sido documentada como rasgo característico del texto académico; esta es indicativa de la densidad informativa del texto, y es un recurso para la expresión de los participantes en el discurso, que en este tipo de escritos se presentan preferentemente de manera explícita y lexicalizada (Schleppegrell, 2001; Snow & Uccelli, 2009).

Para este estudio la *densidad nominal* se define como la proporción de sustantivos en un texto, y se midió de dos maneras: a) como un porcentaje con respecto al total de palabras de contenido semántico en el texto, y b) como proporción respecto del total de cláusulas, expresada a partir de un cero teórico hasta un infinito también teórico. Esta última medición

se basa en la definición de densidad nominal de Hess (2013, p. 120): “la cantidad de sustantivos totales presentes por cláusula”, según la cual una mayor densidad nominal es indicativa de que el texto escrito posee una mayor cantidad de información en menos cláusulas, es decir, un mayor encapsulamiento de la información. Respecto de esta variable, como de la anterior, se aclara que en este trabajo son relevantes sus efectos cuantitativos y no sus manifestaciones cualitativas. Como parámetros se cuenta, por ejemplo, con corpus de narraciones escritas por adolescentes, con una densidad de 1.09-1.19 sustantivos por cláusula (Hess, 2013), y textos expositivos escritos también por adolescentes con 2.05 sustantivos por cláusula (Ravid, 2006).

5.3.2.6. Categorías de sustantivos

Existen varias corrientes lingüísticas con distintas posturas teóricas que se ocupan de la semántica de los sustantivos.²⁶ En el presente estudio se sigue una perspectiva funcional. Por ello se asume que los grupos nominales expresan los participantes gramaticales en cualquier tipo de discurso: los actores, enunciadores, pensadores y cualquier otro rol implicado en los procesos construidos en las cláusulas. Las estructuras nominales contribuyen a elaborar diferentes significados: unos específicos, otros genéricos, también abstractos y técnicos; pueden referir a una realidad del contexto de la enunciación o a algo construido solo en el texto (Fang, Schleppegrell & Cox, 2006).

Al respecto, como se describió en el capítulo dedicado al marco teórico, una de las particularidades del texto académico escrito es el empaquetamiento denso de la información por medio de grupos nominales abstractos, cláusulas nominales complejas y nominalizaciones. Este fenómeno ha sido suficientemente documentado debido a que

²⁶ Ejemplos de esas posturas son la tipología funcional de base sintáctica de Givón (1990), la gramática cognitiva de Langacker (1987) o la gramática hispánica, de corte generativista (Alarcos, 1995; Bosque, 2011; Gili Gaya, 2002; Munguía, 2017, entre otros).

reviste especial interés en el campo del discurso académico (Bello, 2016; García, Hall & Marín, 2005; Halliday & Martin, 1993; Fang, Schleppegrell & Cox, 2006).

Para los fines del presente estudio, se seleccionó como predictor de dominio de la escritura académica la incorporación de sustantivos abstractos en los textos, precisamente por la capacidad que le atribuye la teoría de condensar información, construir nuevos referentes (de tipo disciplinar o conceptual) y de cohesionar el texto. Vale aclarar que, por los alcances de esta investigación, los sustantivos abstractos se analizaron cualitativamente solo para categorizarlos, pero su tratamiento principal fue cuantitativo: se contabilizaron por clases como unidades mensurables de un léxico abstracto empleado por los estudiantes en la construcción del texto académico.

La categorización de los sustantivos con que se operó en esta investigación es de base semántica; recupera la ontología de los referentes de Lyons (1977) y la lexicología de los sustantivos de Schmid (1999b),²⁷ porque las conceptualizaciones de estos autores ayudan a establecer un criterio útil de diferenciación *concreto/abstracto* según ciertas características de los referentes, como la individuación y la imaginabilidad, que aluden a operaciones cognitivas que pueden o no ser realizadas sobre las entidades a las que refieren los sustantivos (Schmid, 1999a).

En este sentido, todos los sustantivos que etiquetan, categorizan, destacan, recortan y ponen en perspectiva referentes de la realidad extralingüística a los cuales la conciencia pueda imaginar e individuar son considerados concretos; es decir, se trata de sustantivos que pueden atribuirse a una clase, a un individuo, a un grupo o colectivo, o bien a un rol (por ejemplo, *silla*, *Virgilio*, *rebaño*, *docente*). En contraparte, los sustantivos abstractos construyen sus referentes al asignarles un nombre, los concretizan y objetivan como si fueran entidades existentes, para poder operar con esas realidades complejas al

²⁷ El Anexo G (Tabla 20) presenta un resumen de la propuesta de lexicología de los sustantivos de Schmid (1999b) que se usó como fuente.

condensarlas en una forma nominal (son ejemplos de sustantivos abstractos *grandeza, imaginación, fuerza, intelecto, velocidad*; pero también *mitosis, fisión, isomorfismo* o *connotación*, que además pertenecen al lenguaje de las ciencias).

Este criterio básico permite, para los fines del presente trabajo, separar entre sustantivos abstractos *stricto sensu* y otros sustantivos que, aunque también hacen un recorte del flujo de la experiencia, pueden remitir a realidades concretas, fácilmente imaginables por estar sujetas al espacio y al tiempo, y permite también poner entre paréntesis cualquier criterio morfológico en la clasificación, dando por sentada la posibilidad de que se incluyan sustantivos derivados (nominalizaciones) entre los sustantivos abstractos, pero, al no ser estos objeto de interés específico, los abstractos en su calidad de derivados no forman una clase particular en el análisis.

De este modo, metodológicamente, la propuesta semántica de este trabajo distingue dos categorías de sustantivos: ‘concretos’ y ‘abstractos’, más cuatro clases de sustantivos abstractos. Los sustantivos concretos quedan definidos como sustantivos que nombran entidades de primer orden, que en la escala de Lyons (1977) son los referidos a clases generales como personas, animales, organismos y objetos físicos localizados en el espacio, con propiedades perceptuales constantes que permiten determinarlos como existentes en un tiempo dado. Se incluyen algunas entidades de segundo orden como las unidades temporales (*minuto, hora, día, año*), espaciales (*milla, tramo, palmo, kilómetro*) y los roles sociales (*autor, policía, esposa*), que por su carácter partitivo y clasificatorio permiten individualizar lo que nombran).

Los *sustantivos abstractos* se han definido como sustantivos que nombran entidades de segundo orden (Lyons, 1977) no individuados, constructos perceptuales que son concretizados como eventos o resultados de acciones por medio de la partición intencionada de sucesos continuos que por naturaleza no poseen bordes o límites. También

incluyen entidades de tercer orden, constructos conceptuales que son objetivados por medio de un proceso de hipostatización constructora de referentes cargados ideológicamente, y cuyos contenidos tienen una realización más o menos duradera y estable, o bien efímera, según el contexto de la enunciación (por ejemplo, los significados de *verdad*, *racionalidad*, *pecado*, *vida* o *sentido* pueden cambiar según si los emplea un filósofo, un sociólogo o un teólogo, aunque mantengan una cierta base estable; mientras que otros sustantivos abstractos concretan, clasifican o evalúan fragmentos del texto específico en que aparecen, sin que esa concretización tenga necesariamente que ser la misma en otro texto: *hecho*, *ejemplo*, *dato*, *situación*, *contexto*, *asunto* o *problema*).

Con base en esta primera categorización, se delimitaron cuatro clases de sustantivos abstractos que responden adecuadamente a las necesidades de análisis de los sustantivos del corpus. Estas son: *Eventos del mundo físico y estados de cosas*, *Procesos del mundo mental y estados afectivos*, *Constructos teóricos y conceptos* y *Encapsuladores*. Enseguida se explica cada clase, con ejemplos tomados del corpus.

Eventos del mundo físico y estados de cosas. En esta clase entran dos tipos de referentes que tienen en común la misma fuente, que es el mundo de la experiencia perceptible por los sentidos. El primero, *eventos del mundo físico*, se relaciona con alguna actividad, por lo que se le puede pensar como un evento con una duración, es decir, la conciencia puede recortar o idear sus bordes a partir de un proceso con un inicio y un fin hipotéticos, como sucede con un *accidente*, un *ascenso* o un *desarrollo*, o bien, la conciencia puede situarlos puntualmente en una circunstancia en la que son relevantes, como son los casos de una *solución*, una *respuesta* o una *reacción*. Los eventos también se pueden derivar de una acción realizada por algún agente, de la cual resulta, por ejemplo, un *golpe*, una *vuelta* o un *ajuste*. Una clave para diferenciarlos es pensar en estos referentes como ‘cosas que suceden o se hacen’. Por otra parte están los *estados de cosas*, que tienen referentes menos dinámicos y por lo tanto no se conciben como resultado de acciones sino

más bien como rasgos, cualidades o disposiciones que particularizan algo o a alguien en un determinado momento; dichos atributos se pueden inferir a través de la observación física de entes, personas o colectivos que los poseen de un modo más o menos estable o que los experimentan de manera transitoria. La clave es pensarlos como ‘algo que se tiene o se experimenta’ en un horizonte temporal dado, como puede ser la *naturalidad*, el *racismo*, la *oportunidad*, el *placer*, una *alternativa* o la *pobreza*.

Procesos del mundo mental y estados afectivos. La diferencia central entre este tipo de referentes y los de la clase anterior es que, aunque se deriven también de procesos o estados, estos no se pueden observar directamente, pues conforman la vida interior anímica e intelectual de las personas u otros seres a los que atribuimos animacidad. Por ello, la manifestación de dichos procesos o estados se infiere de ciertas conductas o de su efecto sobre otros.

Una clave para distinguir los *estados afectivos* de los *procesos del mundo mental* de las cualidades o atributos transitorios del carácter (como la *naturalidad* o la *sencillez*) es su mayor grado de agentividad: de los sujetos a los que se les atribuyen esas ‘pasiones’ las sufren o las viven; en tanto que ese ‘trabajo cognitivo’ lo promueven o lo movilizan en su interior. Este esfuerzo o aquella experiencia pueden ser inferidos por otros a partir de las conductas externas, como sucede con el *miedo*, la *frustración*, la *inspiración*, el *asombro*, la *fascinación* o el *desencanto*; una *interpretación*, una *clasificación* o un *análisis*.

Constructos teóricos y conceptos. El sustantivo que da nombre a una construcción teórica se caracteriza por su saturación ideológica, lo cual quiere decir que en lugar de tener un referente verificable en el mundo de la experiencia, es construido y está cargado de significados que pueden ser debatidos, sujetos a falsación, matizados o reconceptualizados según el horizonte histórico, cultural y la postura disciplinar desde los que se les examine y use. Una clave para distinguirlos es que funcionan como etiquetas que permiten a sus

usuarios llegar a un acuerdo convencional sobre sus contenidos y operar con ellos a través de las disciplinas, como sucede con *verdad, vida, clase (social), esencia, despotismo o destino*, nociones a las cuales diversas ciencias conceden, agregan o matizan cargas de sentido determinadas.

Encapsuladores. Este tipo de sustantivo abstracto funciona como encapsulador cuando marca un ‘vacío’ que se llena con la información del contexto discursivo inmediato, lo cual quiere decir que –en principio– es un sustantivo que no está saturado de sentido. De este modo, su referente es transitorio y situado en el contexto, razón por la cual se dice que ‘encapsula’ o ‘empaqueta’ información que está presente en el discurso de manera más o menos adyacente al sustantivo encapsulador. Algunos ejemplos de ello son *evidencia, ejemplo, argumento, factor o elemento*, cuya realización léxica se debe buscar en el entorno discursivo inmediato, independientemente de los referentes extra-lingüísticos que podrían o no ajustarse al uso actual dentro del texto; es decir, la definición de diccionario de la palabra *ejemplo*, aunque puede guiar la identificación de uno cuando lo encontramos en contexto, actualiza un contenido efectivo especificado en las cláusulas circundantes.

Para esclarecer lo anterior, véase el siguiente caso: “El extranjero nos da un claro *ejemplo* de lo que es la indiferencia ante la vida, Mersault es un tipo completamente indiferente ante situaciones extraordinarias, como Ø el asesinato de un hombre” (15H20); aquí la primera de las cláusulas subrayadas concretiza *ejemplo* (“Mersault es un tipo completamente indiferente ante situaciones extraordinarias”), y a su vez *situaciones extraordinarias* se completa con la especificación “como el asesinato de un hombre”. En algunos casos el encapsulador porta información modalizadora que puede ser graduada, como en *asunto, inconveniente, problema, desgracia*, sustantivos que evidencian matices graduales en el punto de vista del hablante, desde una cierta neutralidad hacia la evaluación negativa, y este mismo hablante, al usarlos, trata de influir en su interlocutor para que

clasifique de ese modo los contenidos a los que esos sustantivos hacen referencia en el texto. Es el caso en: “Mersault no se entristece con su partida de este mundo, la acepta, no ve a la muerte como algo trascendente, simplemente la ve como algo inevitable, un *suceso* que estaba escrito desde que llegó a este mundo” (14H19), donde la muerte del personaje, metaforizada como la “partida de este mundo” es calificada como *suceso*, sin un matiz fatalista.

Paralelamente, los sustantivos abstractos encapsuladores pueden servir al escritor para cohesionar anafóricamente: “En la época, la censura era duramente aplicada a los escritores, *razón* por la cual Camus recurre a la manifestación discreta en sus obras” (14M15), donde *razón* encapsula y recupera la información de la cláusula precedente; o catafóricamente, al conectar un segmento textual sucesivo que funge como complemento: “El *propósito* de este ensayo es comparar El Viejo y el Mar y El Extranjero” (15M01).

En la siguiente página, la Tabla 3 resume la propuesta de categorización de los sustantivos con estas definiciones, la clasificación de los abstractos y algunos ejemplos.

5.3.2.7. *Léxico especializado*

En el planteamiento de esta tesis se incluye el supuesto de que los estudiantes adoptan, para la composición de su discurso, el léxico de las materias escolares que forman parte de la situación de escritura, y por tanto el vocabulario usado por ellos en los textos puede ser categorizado semánticamente como un repertorio que tiende hacia la especialización disciplinar, siendo este uno de los rasgos más sobresalientes de la escritura académica (Coxhead, 2017; Halliday & Martin, 1993; Schleppegrell, 2004; Zwiers, 2008).

Tabla 3. *Categorización semántica de los sustantivos empleada en esta investigación.*

Categorías	Clases	Ejemplos
<p><i>Concretos.</i> Sustantivos que nombran entidades de primer orden en la escala de Lyons (1977): clases generales como personas, animales, organismos y objetos físicos localizados en el espacio, con propiedades perceptuales constantes que permiten determinarlos como existentes en un tiempo dado. Se incluyen algunas entidades de segundo orden como las unidades temporales, espaciales y los roles sociales, que por su carácter partitivo y clasificatorio permiten individualizar lo que nombran.</p>	<p><i>No aplica</i></p>	<p>aire, Alemania, día, escritor, gente, Hemingway, hombre, mundo, persona, pez, siglo.</p>
<p><i>Abstractos.</i> Sustantivos que nombran entidades de segundo orden (Lyons, 1977) no individuados, constructos perceptuales que son concretizados como eventos o resultados de acciones por medio de la partición intencionada de sucesos continuos que por naturaleza no poseen bordes o límites. También entidades de tercer orden, constructos conceptuales que son objetivados por medio de un proceso de hipostatización constructora de referentes cargados ideológicamente, y cuyos contenidos tienen una realización más o menos duradera y estable, o bien efímera, según el contexto de la enunciación (Schmid, 1999b).</p>	<p><i>Eventos del mundo físico y estados de cosas.</i> Sustantivos que refieren procesos, actividades y sus resultados; y también cualidades y estados de cosas perceptibles y localizados en el tiempo, con diversos grados de dinamismo y agentividad, de los cuales se dice que ocurren o suceden, más que existir.</p>	<p>ambigüedad, apogeo, aportación, argumentación, ascenso, audiencia, combate, debate, dramatismo, enseñanza, esfuerzo, liberación, respaldo, salud, tono, trivialidad, unión, vitalidad, vuelta.</p>
	<p><i>Procesos del mundo mental y estados afectivos.</i> Sustantivos que refieren procesos mentales y estados anímicos de un carácter más subjetivo que físico, los cuales pueden ser inferidos por las conductas de sujetos que los experimentan o que les son atribuidos.</p>	<p>aborrecimiento, amor, apuro, autoevaluación, cariño, comparación, desprecio, enjuiciamiento, luto, percepción, reconocimiento, recuerdo, terror, vergüenza.</p>

Categorías	Clases	Ejemplos
	<i>Constructos teóricos y conceptos.</i> Dan nombre a entidades de tercer orden, proposiciones e ideas que no están sujetas al espacio y al tiempo, de las cuales se dice que pueden ser afirmadas, negadas, recordadas u olvidadas. Son categorías del pensamiento susceptibles de falsación; relaciones abstractas o haces de ideas con las que se opera tanto en la vida cotidiana como en las ciencias.	comunismo, cultura, diálogo, didactismo, espacio, espíritu, existencialismo, hipertextualidad, personaje, poder, razón, realidad, sentido, ser, teoría, valor, verdad, virtud.
	<i>Encapsuladores (shell nouns).</i> Expresiones lingüísticas que requieren una realización léxica en su co-texto, por un proceso de especificación efímera. La porción de experiencia singularizada por ellos depende del contexto lingüístico y situacional. Vehiculan una representación mental modalizadora del hablante/escritor. Tienen funciones anafóricas y complementantes.	antecedente, aspecto, asunto, caso, clase, cuestión, ejemplo, evidencia, factor, meta, opción, propósito, punto, rasgo, tipo.

Fuente: Elaboración propia con base en la escala nominal de Lyons (1977), la lexicología de los sustantivos y el modelo de *shell nouns* de Schmid (1999b, 2000, 2018).

Chung y Nation (2004) señalan que hay varios métodos para identificar el vocabulario técnico en los textos: usar una escala de aproximación a una esfera de discurso especializado, usar diccionarios y opiniones de expertos, observar las pistas del escritor en el texto (etiquetas en diagramas, marcación tipográfica y definiciones), o bien el análisis comparativo de frecuencias en corpus de distinto contexto y sus colocaciones.

Estos mismos autores señalan que el carácter técnico o especializado de una palabra es un aspecto funcional de la misma (Chung & Nation, 2004), y por ello tiene que estudiarse el uso para determinar qué función está cumpliendo la palabra en el contexto del discurso. En razón de ello, en este trabajo se hace una aproximación mixta entre el uso de diccionarios y el análisis del uso en líneas de concordancia obtenidas por medios automatizados.

Se considera *léxico especializado* a los elementos nominales de carácter científico, técnico o específico de una materia o disciplina, que son a menudo –pero no exclusivamente– abstractos, y que configuran un vocabulario o léxico especial, en el sentido de que los ítems léxicos tienden a ocurrir principalmente en una disciplina y no en otras (Biber, 2006; Coxhead, 2017; Halliday & Martin, 1993; Schleppegrell, 2001; Snow & Ucelli, 2009; Zwiers, 2008). Algunos autores le llaman léxico especializado, vocabulario técnico o vocabulario tecno-científico (Martín, 2004), y se puede organizar por disciplinas o campos conceptuales.²⁸

A cada ítem léxico especializado se le identifica en este estudio como *término especializado*. Un *término* es “toda palabra perteneciente a un lenguaje de especialidad” (Martín, 2004, p. 15); concepto amplio que incluye y se distingue de *tecnicismo* y de *nomenclatura* en que el primero se refiere a la vertiente práctica de la ciencia exclusivamente, y del segundo porque es una clasificación convencionalizada para los objetos de estudio de las ciencias exactas.

Con el fin de agrupar los términos que componen el léxico especializado, se usa la noción de *campo disciplinar* para referirnos a un conjunto identificable de ramas del saber, de los cuales en el corpus se observaron dos: el campo de la lengua y la literatura y otras disciplinas humanísticas. Estos campos disciplinares se operativizaron así: *Lengua y literatura*: referentes del hecho literario, de la poética, la narrativa y la lingüística.²⁹ *Otras*

²⁸ Martín (2004) utiliza la noción de *campo conceptual* en lugar de *campo semántico*, para organizar el vocabulario tecno-científico porque sirve para identificar y reunir unidades léxicas dentro de una misma categoría sobre una base conceptual y no por afinidad de significado. La ciencia y la tecnología, según este autor, abarcan ramas diversas –las disciplinas–, pero también actividades, procesos, métodos propios y maneras de presentar sus resultados que generan un metalenguaje particular que puede agruparse en función de su utilidad para referir todos esos aspectos, y que necesariamente se adaptan a la naturaleza y contenidos de cada corpus analizado. En nuestro caso usaremos la noción de *campo disciplinar* por ser la que más conviene a nuestro trabajo.

²⁹ Para constatar el uso disciplinar de una palabra, se empleó el *Diccionario de la Lengua Española* de la RAE, el *Diccionario Español de Términos Literarios Internacionales* (Garrido, 2017) y el libro de texto de la clase en que produjeron los textos (Betancourt, 2009).

disciplinas humanísticas: términos con que operan otras disciplinas como la filosofía, la sociología, la psicología o la economía.

La Tabla 4 define estas categorías y ofrece ejemplos por campo disciplinar.

Tabla 4. *Categorización de los sustantivos como léxico especializado.*

Sustantivos	Campo disciplinar	Ejemplos
<i>Léxico no especializado.</i> Sustantivos generales, de uso común, no relacionados con una disciplina, sean concretos o abstractos.		adversidad, ajuste, anzuelo, cicatriz, compañero, condena, desesperanza, foso, noche, reaparición, terror.
<i>Léxico especializado.</i> Elementos nominales de carácter científico, técnico o específico de una materia o disciplina, que son a menudo –pero no exclusivamente– abstractos.	<i>Lengua y literatura.</i> Referentes del hecho literario, de la poética, la narrativa y la lingüística.	adjetivo, analogía, ensayo, metáfora, novela, párrafo, relato, símil, título, tragedia, vocabulario.
	<i>Otras disciplinas humanísticas.</i> Términos con que operan otras disciplinas como la filosofía, la sociología, la psicología o la economía.	alter ego, arquetipo, fascismo, marxismo, posmodernidad, potencia, radicalismo, sociedad, superhombre.

5.3.2.8. *Léxico abstracto especializado*

En la definición misma de léxico especializado se aludió a la posibilidad de que un elemento nominal abstracto o concreto funcione como término especializado. Para cruzar las categorías *concreto/abstracto* y *especializado/no especializado* con que se trabaja sobre el corpus, se definió el siguiente criterio: De acuerdo con la lexicología de los sustantivos de Schmid (1999b), los sustantivos concretos sirven para realizar las operaciones cognitivas de nombrar, categorizar y poner en perspectiva entidades, objetos, personas, etc. Dentro de este caso entra, por ejemplo, etiquetar una ciencia para distinguirla de otras (*lingüística*); clasificar un fenómeno como objeto o unidad de análisis (*sustantivo, cláusula*), o

categorizar el rol social o la función de una persona en un campo de acción determinado (*lingüista*).

En esa misma línea argumentativa, este tipo de sustantivos ayudan a los hablantes a formarse representaciones ‘concretas’ de las realidades que esos sustantivos refieren, y en ese sentido se clasifican como entidades de primer orden (Lyons, 1977), individualizables e imaginables. No son, pues, consideradas entidades abstractas y por ende tampoco sus significantes son considerados sustantivos abstractos, de modo que estos términos disciplinares (*lingüística, sustantivo, cláusula, lingüista*) son, para los fines de este análisis, sustantivos concretos que se categorizan como *sustantivos concretos especializados*.

Otros sustantivos, en cambio, refieren entidades que son producto de actos cognitivos de mayor complejidad, como la partición conceptual y la reificación o hipostatización, actos intencionados que ‘construyen’ por medio del significante su referente, mismo que no es una entidad palpable sino algo que se experimenta, se recorta de un proceso temporal o se hace ‘real’ como objeto mental para poder operar con él en el discurso (Schmid, 1999a). Estas últimas entidades –objetos mentales, ideas, proposiciones y eventos–, de segundo y tercer orden en la tipología de Lyons (1977), son las que Schmid (1999b, p. 221) reconoce como sustantivos abstractos “genuinos”.

A partir de esta distinción, podemos introducir la división entre lo que el propio Schmid (1999b) llama “sustantivos abstractos de todos los días” y que aquí son identificados como *sustantivos abstractos comunes*, y los sustantivos abstractos que sirven para estructurar e individuar las realidades con las que operan las ciencias y la tecnología en todos los campos disciplinares, y que él llama “términos técnicos” y en el presente estudio llamamos *sustantivos abstractos especializados*, aquellos cuyo referente cumpla esa doble condición: estar relacionados con algún campo disciplinar y ser construcciones conceptuales (Ejemplos: *léxico, modernismo, personificación, prolepsis, retórica* o

vocabulario, en el campo de la Lengua y literatura; humanidad, ideología, idiosincrasia, individualismo, paradigma o nihilismo, en el campo de Otras). Por último, todos los sustantivos concretos que no tengan relación con algún campo disciplinar se categorizan como *sustantivos concretos comunes*.

La Figura 2 presenta sintéticamente la conceptualización del cruce entre las categorías *concreto/abstracto* y *especializado/no especializado* que opera en esta investigación, como una función que pueden cumplir las palabras en determinados contextos.

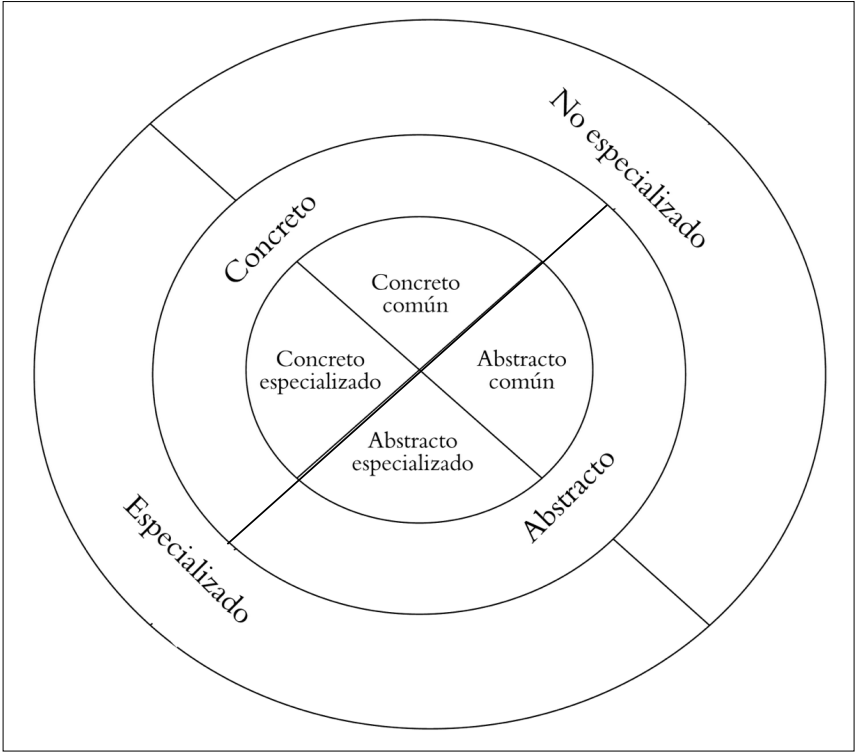


Figura 2. Cruce de las categorías de sustantivos.

En seguida, la Tabla 5 resume los criterios de identificación de los sustantivos abstractos en función de término, con definiciones y algunos ejemplos.

Tabla 5. *Categorización de los sustantivos como léxico abstracto especializado.*

Cruce de categorías	Campo disciplinar	Ejemplos
<i>Sustantivos concretos comunes.</i> Sustantivos concretos, según la Tabla 3, que no tengan relación con algún campo disciplinar.	<i>No aplica</i>	adulto, alarma, filósofo, mancha, revólver, Troya, turco, vecino
<i>Sustantivos abstractos comunes.</i> Sustantivos abstractos, según la Tabla 3, que no tengan relación con algún campo disciplinar.	<i>No aplica</i>	aborrecimiento, acción, angustia, búsqueda, castigo, conexión, desarrollo
<i>Sustantivos concretos especializados.</i> Sustantivos concretos, según la Tabla 3, que tengan relación con algún campo disciplinar.	<i>Lengua y literatura.</i> Referentes del hecho literario, de la poética, la narrativa y la lingüística.	adjetivo, antihéroe, autor, cuentista, ensayo, obra, personaje, texto, verbo
	<i>Otras.</i> Términos con que operan otras disciplinas como la filosofía, la sociología, la psicología o la economía.	depredador, historiador
<i>Sustantivos abstractos especializados.</i> Sustantivos abstractos, según la Tabla 3, que tengan relación con algún campo disciplinar.	<i>Lengua y literatura</i>	absurdismo, intertextualidad, personificación, significado, vanguardismo, vocabulario
	<i>Otras</i>	despotismo, esencia, individuo, régimen, subconsciente, superhombre, teología

En la siguiente página se presenta, en la Tabla 6, un resumen de todas las variables antes descritas para su visualización en conjunto.

Tabla 6. *Concentrado de variables.*

Variable	Medida	Definición	Fuentes
Dominio de la Escritura Académica	PromEA	Puntaje (promedio o nivel) obtenido en una valoración criterial realizada por profesores experimentados, con base en una rúbrica, acerca del dominio de la escritura académica manifiesta en un texto producido en el ámbito escolar.	McNamara, Crossley y McCarthy, 2009
Extensión en palabras	Palabras	Tamaño del texto en palabras ortográficas (tokens) o en	Johansson, 2008
Extensión en cláusulas	Cláusulas	cláusulas (predicado que expresa una situación simple).	Berman y Slobin, 1994
Diversidad léxica	VocD	Proporción de palabras diferentes (tipos) respecto del total de palabras ortográficas (tokens) que constituyen un texto.	Johansson, 2008; Malvern <i>et al.</i> , 2004; McCarthy y Jarvis, 2010
Densidad léxica	LexD	Proporción de palabras de contenido semántico (S, V, A, Adv) sobre el total de palabras ortográficas (tokens), sin distinción de frecuencia.	Johansson, 2008; Ure, 1971
Densidad nominal en el léxico	SustLex	Proporción de sustantivos en un texto, medida sobre el total de palabras de contenido semántico o	Hess, 2013
Densidad nominal en la cláusula	SustCl	por cláusula.	Hess, 2013; Ravid, 2006
Diversidad nominal	SustDif	Proporción de sustantivos diferentes (tipos) sobre el total de sustantivos en el texto.	Ravid, 2006
Categorías de sustantivos	Conc/Abs	Sustantivos concretos y sustantivos abstractos presentes en el texto.	
Clases de sustantivos	Fís/Ment/Const/Enc	Clases de abstractos: Eventos del mundo físico y estados de cosas, Procesos del mundo mental y estados afectivos, Constructos teóricos y conceptos, Encapsuladores.	Lyons, 1977; Schmid, 1999b, 2000, 2018
Léxico especializado	No Esp/Esp LeLit/Otras	Elementos nominales de carácter científico, técnico o específico de una materia o disciplina, que son a menudo, pero no exclusivamente, abstractos.	Halliday y Martin, 1993; Schmid, 1999b; Zwiers, 2008
Léxico abstracto especializado	ConCo/ConEs/AbsCo/AbsEsp	Elementos nominales de carácter científico, técnico o específico de una materia o disciplina, que cumplen la condición de ser también abstractos.	

5.4. Procedimientos del análisis I: Indicadores léxicos

En esta primera etapa de la investigación las variables predictoras fueron la diversidad y la densidad léxica, así como la diversidad y la densidad nominal, todas ellas expresadas como variables continuas. La variable dependiente fue el nivel de dominio de la escritura académica (PromEA) expresado como variable continua con un decimal. A continuación se desarrollan los procedimientos de obtención de los datos para cada variable.

En cuanto al PromEA, variable que ha sido explicada en el apartado 5.3.1, las puntuaciones de los cinco jueces expertos devolvieron un 72% de confiabilidad (ICC2k=0.72, $p < 0.0001$). Con este porcentaje se alcanzó la expectativa planteada para una estrategia de evaluación con piloto pero sin entrenamiento. Tomando además en cuenta que el 80% de confiabilidad quedó dentro del intervalo de confianza (0.59 a 0.82), se consideraron válidas y confiables las puntuaciones emitidas por los jueces para construir el indicador PromEA. Dicha construcción fue detalladamente explicada en el apartado 5.2.2.

Por su parte, la medición de la extensión del texto se hizo por el conteo de los tokens de palabras ortográficas de manera automatizada con el software AntConc,³⁰ y de cláusulas, por segmentación manual y conteo automatizado en Microsoft Excel.

Los coeficientes de diversidad léxica se calcularon con ayuda del software libre TextInspector. La densidad léxica se obtuvo etiquetando la lista completa de palabras del corpus con el software Freeling,³¹ y aplicando luego un filtro en Microsoft Excel para el conteo por clase sintáctica, y una fórmula de proporción que dividió la suma de palabras de contenido de cada texto entre el total de palabras ortográficas (tokens).

³⁰ AntConc es una herramienta de uso libre desarrollada por el lingüista Laurence Anthony para el análisis de textos y la obtención de concordancias. Disponible en <http://www.laurenceanthony.net/software.html>

³¹ El etiquetador sintáctico Freeling pertenece a una suite de herramientas para análisis del lenguaje desarrollada en la Universitat Politècnica de Catalunya. En México, el Grupo de Ingeniería Lingüística de la UNAM desarrolló una interfaz para usar dicha herramienta vía web: <http://www.corpus.unam.mx/servicio-freeling/>

La densidad nominal se obtuvo igualmente con fórmulas en el entorno de Excel dividiendo la cantidad de sustantivos sobre el total de palabras de contenido en cada texto, y luego sobre la cantidad de cláusulas. La diversidad nominal se obtuvo al dividir los tipos de sustantivos entre el total de instancias de sustantivos en cada texto.

Para contestar las preguntas de investigación de esta etapa del trabajo se emplearon dos métodos de análisis estadístico de los datos: 1) el método correlacional, que calcula la fuerza y la probabilidad de asociación entre dos variables (correlación de Pearson), 2) el predictivo, que relaciona una variable dependiente con un conjunto de variables explicativas como herramienta para predecir efectos (regresión lineal múltiple). Las pruebas y los gráficos se realizaron en el entorno de R.

5.5. Procedimientos del análisis II: Sustantivos abstractos

La categorización y cuantificación de los sustantivos se hizo en dos pasos: primero, con ayuda de AntConc se extrajeron las líneas de concordancia de cada sustantivo del corpus; enseguida se analizó el contexto discursivo de cada sustantivo para desambiguar su función sintáctica³² y determinar su estatus dentro de las categorías ‘concreto’ o ‘abstracto’, luego, se clasificaron dentro de las cuatro clases de abstractos: ‘eventos del mundo físico y estados de cosas’, ‘procesos del mundo mental y estados afectivos’, ‘constructos teóricos y conceptos’ y ‘encapsuladores’; tal como se definieron dichas categorías y clases en el Apartado 5.3 y se puntualizaron en la Tabla 3.

Esta información se organizó en tres archivos: dos con datos categóricos, el primero con 22,225 sustantivos del corpus completo categorizados; el segundo con una cala sobre la muestra de 60 textos (los más altos y los más bajos en el indicador PromEA) con 11,170

³² La clase funcional de las palabras etiquetadas automatizadamente como sustantivos se desambiguó manualmente (contra adjetivos o verbos) con ayuda del Diccionario de la Lengua Española de la RAE. También se descartó su posible función como marcadores discursivos. Un resumen de los criterios de inclusión/exclusión y conteo puede consultarse en el Anexo H.

sustantivos igualmente categorizados, con la identificación del texto en el que aparece cada uno y el nivel alto o bajo del promedio. El tercer archivo, numérico, recogió las frecuencias simples de cada categoría y clase de sustantivos en cada texto de la muestra, y el nivel de PromEA correspondiente (alto/bajo).

En una exploración cuantitativa, con la base de datos categóricos se obtuvieron las frecuencias simples y las proporciones de cada categoría y clase de sustantivos en el corpus completo y se graficaron los resultados. Con la segunda base de datos se obtuvieron las frecuencias y proporciones de estas categorías en los textos de la muestra, agrupados por su ubicación en los niveles alto o bajo de PromEA.

El análisis estadístico puso a prueba la contingencia de las variables para verificar si las distintas frecuencias de las categorías y clases de sustantivos se relacionaban o no con el nivel alto o bajo de promedio de un texto; para ello se aplicó una prueba de independencia Chi cuadrado y se corroboró el resultado con la prueba exacta de Fisher.

Un segundo análisis, de tipo discriminante,³³ consistió en usar la base de datos numéricos para calcular la media de cada categoría y clase de sustantivos en los textos agrupados por su ubicación alta o baja en PromEA. Se aplicó una prueba t de Student para grupos independientes para verificar si las medias de cada categoría y clase de sustantivos en los dos niveles de promedio eran estadísticamente diferentes, y se graficaron los resultados.

Con el fin de ofrecer algunos ejemplos de uso de los sustantivos abstractos en contexto, se seleccionaron casos del corpus, dejando fuera los sustantivos abstractos

³³ “El análisis discriminante es una técnica estadística multivariante cuya finalidad es analizar si existen diferencias significativas entre grupos de objetos respecto a un conjunto de variables medidas sobre los mismos para, en el caso de que existan, explicar en qué sentido se dan y facilitar procedimientos de clasificación sistemática de nuevas observaciones de origen desconocido en uno de los grupos analizados” (De la Fuente, 2011).

especializados para hacer su descripción de manera separada,³⁴ y la descripción de los ejemplos elegidos se enfocó aspectos semánticos, no morfológicos.

5.6. Procedimientos del análisis III: Léxico especializado

Con todos los sustantivos ya etiquetados por su categoría y su clase, se procedió a extraer datos respecto del léxico especializado que emplearon los estudiantes en sus textos, el campo disciplinar al que pertenecía el término, y su frecuencia de uso, lo cual se cuantificó en primer lugar sobre el corpus completo.

Para relacionar el uso de léxico especializado con el dominio de la escritura académica percibido por los jueces expertos, se trabajó sobre la muestra de 60 textos, los 30 de nivel más alto y los 30 de nivel más bajo en el dominio. Los sustantivos de esta muestra fueron caracterizados mediante el filtrado de sus líneas de concordancia, cuyas instancias nominales fueron previamente etiquetadas. De cada uno se obtuvo la cantidad total (n) y la frecuencia total de uso (f) de cada término especializado, y se asentó en la base de datos junto con el nivel obtenido en PromEA (alto/bajo).

El criterio básico para considerar un sustantivo como término especializado fue el uso dado por el escritor a esa palabra, mismo que se hizo observable mediante el análisis de su contexto discursivo en las líneas de concordancia, que permitió deslindar si se trataba de un uso general o disciplinar, y por tanto especializado, como se describió en el Apartado 5.3.2.7 (ver Tabla 4). De este modo los sustantivos se agruparon en dos categorías: especializados y no especializados, y luego se agruparon los especializados en las dos grandes clases que se pudieron observar en el corpus por su relación con alguna disciplina: *Términos de Lengua y literatura*, y *Otras disciplinas humanísticas*. Como se explica en el Anexo H de este trabajo, un término disciplinar también puede ser sustantivo abstracto; lo es mientras no se refiera a algo físico e individuado como una *biografía* o una *obra*, o

³⁴ El análisis del cruce ‘abstracto-especializado’ es descrito en el Apartado 5.7.

configure un rol como un *escritor*, un *antagonista* o un *antihéroe* (sustantivos que son contados entre los concretos por referir objetos físicos o individuos, de acuerdo con lo establecido para esta investigación). Por lo demás, se categorizaron y clasificaron con base en los mismos criterios generales que aplicaron para todos los sustantivos.

Los análisis estadísticos siguieron el mismo método descrito para los sustantivos abstractos en el apartado inmediato anterior. Para ilustrar la forma en que el léxico especializado fue empleado en la construcción de los textos, se describieron algunos ejemplos en su contexto de uso.

5.7. Procedimientos del análisis IV: Léxico abstracto especializado

El análisis de los sustantivos abstractos y de los términos especializados en el corpus dejó ver la posibilidad de un cruce entre estas dos categorías, que por método se realizó aparte, a la manera de cala. Para obtener los datos, se etiquetó manualmente la lista completa de sustantivos de la muestra de 60 textos, de acuerdo con las categorías ‘concreto’ o ‘abstracto’ y ‘especializado’ o ‘común’ que fueron definidas en el Apartado 5.3.2. Luego estas etiquetas se concatenaron para generar los cruces ‘concreto común’, ‘concreto especializado’, ‘abstracto común’ y ‘abstracto especializado’. Una vez re-etiquetados todos los sustantivos de esta manera, se extrajeron las listas de cada nueva categoría con el método de filtrado en Excel.

Estos datos se organizaron en dos archivos: uno con variables categóricas, con cada sustantivo etiquetado y la identificación del texto en el que aparece cada etiqueta; y otro con datos categóricos y numéricos: la frecuencia de cada categoría de sustantivos por texto, la identificación del texto y su nivel de PromEA.

En un primer análisis de contingencia de variables, con la base de datos categóricas se obtuvieron las proporciones de cada cruce de categorías en la muestra y se graficó el

resultado. Del mismo modo se obtuvieron las proporciones de estas categorías en los textos agrupados por su ubicación alta o baja en el promedio (PromEA).

De manera semejante a como se trataron los datos de sustantivos y de términos especializados, se realizó una prueba de independencia Chi cuadrado y se corroboró el resultado con la prueba exacta de Fisher para verificar si las distintas proporciones de las categorías dependían o no del grupo al que pertenecían los textos.

Luego, el análisis discriminante se realizó con los datos numéricos para calcular la media de uso de cada cruce de categorías en los textos agrupados por su ubicación alta o baja en PromEA. A continuación se corrió una prueba t para grupos independientes para verificar si las medias de cada cruce de categorías en los grupos eran estadísticamente diferentes, y se graficaron los resultados.

Como corolario, se seleccionó una sub-muestra de la categoría de mayor interés para este trabajo, o sea, los sustantivos abstractos especializados, y se describieron algunos ejemplos descolantes en función de dos criterios, que la vez sirvieron para asegurar la representatividad de esta sub-muestra: 1) su frecuencia y 2) su campo y procedencia probable, como se explica a continuación.

En primer lugar, se da por sentado que la frecuencia de una palabra en un corpus es la medición de las instancias u ocurrencias de un constructo o fenómeno cuya relevancia siempre es relativa, según la ley de Zipf (McEnery & Hardie, 2011), pero que en ciertas condiciones puede ser elocuente acerca de la penetración de algún elemento léxico en un determinado tipo de discurso (Sinclair, 2006, pp. 28-29). Así, la ocurrencia de una palabra puede ser vista como un *cuanto* –un monto, un valor– de información, pero también como la aparición representativa de un fenómeno. Sobre esta base, se eligieron los primeros 20 sustantivos abstractos especializados de la lista ordenada por frecuencia.

En segundo lugar, por lo que toca al campo y procedencia de los sustantivos abstractos especializados, se tomaron en cuenta dos cosas: a) que la mayoría de los términos especializados proviene del campo de la lengua y la literatura ('LyL') en razón de que los textos que forman el corpus fueron producidos en una clase de Español y literatura, mientras que la minoría proviene de otras áreas de las ciencias sociales y las humanidades ('Otras'); b) que entre ellos se halla una lista de términos del análisis comparativo literario que los estudiantes podían tomar directamente de su libro de texto (en razón de lo cual fue etiquetado como 'Libro de texto'); mientras que los demás podían tener procedencias diversas, desde la clase expositiva con medios escritos y visuales, junto con sus interacciones orales en la forma de discusiones de clase, hasta las propias obras literarias que fueron objeto de estudio en el aula (todo lo cual quedó englobado en la categoría 'Aula'). De este modo, se ponderaron esas dos circunstancias del contexto –campo y procedencia– en los sustantivos abstractos especializados de la lista de mayor frecuencia.

6. RESULTADOS

En este capítulo se presentan los resultados de cada etapa del análisis con referencia al objetivo central y las preguntas específicas de investigación. Se contrastan los datos de las variables en el corpus y en la muestra con la hipótesis establecida acerca de su relación con el indicador PromEA; además, se ofrece una descripción breve, de carácter ilustrativo, de algunas manifestaciones de las principales variables de uso del léxico.

El objetivo general de este trabajo fue determinar qué características del léxico pueden ser indicadores efectivos del nivel de dominio de la escritura académica, y en qué orden se puede jerarquizar su importancia, mediante el análisis del léxico en textos que fueron producidos por estudiantes de bachillerato y evaluados por jueces expertos. La pregunta central de la tesis fue: ¿Qué características del léxico en los textos escritos por estudiantes son indicadores efectivos del nivel de dominio de la escritura académica y en qué orden de importancia? Para ello se hicieron cuatro abordajes al corpus, cuyos resultados se reportan en los siguientes apartados.

6.1. Análisis I. Indicadores léxicos

El propósito de este primer acercamiento a los datos fue identificar qué características en el uso del léxico en los textos muestran tener algún efecto sobre la calificación otorgada por jueces expertos, y en qué condiciones se da ese efecto. En concreto, se determina si las diferentes mediciones de diversidad léxica, densidad léxica, densidad nominal, diversidad nominal y extensión en cantidad de palabras en los textos suponen una variación en la calificación; si algunas de esas variables combinan sus efectos sobre el promedio, y si estas observaciones permiten hacer un deslinde de dichos aspectos del léxico por su viabilidad como indicadores de dominio de la escritura académica.

La pregunta específica a contestar en esta etapa fue: ¿Cómo interactúan las mediciones de extensión de los textos, de diversidad y densidad léxica, de diversidad y densidad nominal, y cuáles son sus efectos sobre el promedio de calificación otorgada a los textos por jueces expertos?

En los siguientes apartados se exponen las observaciones de esas variables sobre los textos y su comportamiento como predictores de la calificación.

6.1.1. Resultados de las mediciones de las unidades de léxico en el corpus

Para dar contexto a los resultados, se presenta un panorama general del léxico desplegado en el corpus. En conjunto, este se compone de 99,973 palabras ortográficas. De ellas, casi la mitad son palabras de contenido semántico, cuyas frecuencias se distribuyeron así: 22,225 sustantivos, 16,526 verbos, 6,601 adjetivos y 3,541 adverbios. La Tabla 7 muestra estos datos y los porcentajes de cada clase funcional respecto del total de palabras.

Tabla 7. *Léxico empleado en los textos del corpus.*

Clase	Total de instancias	%
Sustantivos	22,225	22.2
Verbos	16,526	16.5
Adjetivos	6601	6.6
Adverbios	3541	3.5
Palabras léxicas	48,893	48.9
Palabras ortográficas	99,973	

Nota: El porcentaje se calculó en relación con el total de palabras ortográficas.

El modo en que se usaron estas informaciones para convertirlas en indicadores de uso del léxico se detalló en la sección metodológica; ahora bien, el resultado de esos

procedimientos se sintetiza en la Tabla 8, cuyos referentes –obtenidos en estudios previos– pueden ser consultados en el apartado 5.3.2.

Tabla 8. *Mediciones de cada variable en el conjunto de los textos.*

Variable	Unidad	Media (DE)
Dominio de la Escritura Académica	PromEA	3.77 (0.65)
Extensión en palabras	Palabras	807.1 (255.76)
Extensión en cláusulas	Cláusulas	112.37 (37.33)
Diversidad léxica	VocD	73.14 (10.66)
Diversidad nominal	SustDif	57.62 (8.27)
Densidad léxica	LexD	49.61 (1.83)
Densidad nominal en el léxico	SustLex	45.60 (4.20)
Densidad nominal en la cláusula	SustCl	1.65 (0.36)

Nota: DE=desviación estándar. Las fuentes y los parámetros se citaron en la Tabla 6.

Con estas características de los textos en cada uno de los indicadores léxicos, se procedió a analizar los efectos de estas mediciones sobre la variación de PromEA.

6.1.2.Efectos del léxico sobre PromEA

En un primer paso se buscaron correlaciones de cada variable respecto de PromEA. La prueba Pearson arrojó nula significancia en todas las mediciones excepto la extensión de los textos (que no es un indicador léxico). Este hallazgo condujo a la hipótesis de que los indicadores léxicos no tienen un efecto simple (uno a uno) con la calificación, sino que interactúan; esto significa que al modificarse uno de ellos afecta la manera en que otro repercute en el sentido de la evaluación. Lo anterior se comprobó tras ajustar e interpretar

una serie de modelos de regresión lineal múltiple entre las variables predictoras y PromEA como variable numérica continua, incluyendo la extensión de los textos. Los efectos que se describen a continuación dan cuenta de un equilibrio –al parecer necesario– entre esos aspectos del léxico en la construcción de un texto para que este sea bien valorado.

La Tabla 9 presenta un resumen de estas estimaciones, comenzando por el tamaño del efecto de cada combinación de indicadores sobre PromEA; seguida por el poder explicativo de cada una, que se interpreta como un porcentaje de los datos que pueden ser relacionados con esa interacción de factores, y finalmente la probabilidad estadística de cada interacción.

Tabla 9. *Estimación de los efectos de las interacciones entre variables sobre PromEA.*

VARIABLES	INTERACCIONES	F	R ²	p<
Densidad nominal en el léxico	VocD:Sust.Lex	5.47	0.20	0.05
Densidad léxica	Tokens:LexD	4.18	0.15	0.05
Densidad nominal en la cláusula	SustLex:Sust.Cl	3.97	0.14	0.01
Diversidad nominal	Sust.Lex:Sust.Dif	3.97	0.14	0.001
Diversidad léxica	Tokens:VocD	3.77	0.14	0.05

Notas: Modelos de regresión lineal múltiple con 7 y 112 grados de libertad, $\alpha=0.05$. Ordenados por el tamaño del efecto (F). Se reportan los valores de p de las interacciones únicamente; todos los modelos reportaron significancia estadística.

Los resultados de los modelos predictivos muestran que las medidas de densidad en los textos son las que tienen mayores efectos sobre la calificación. En primer lugar, la densidad nominal medida sobre el léxico, seguida por la densidad léxica y por la densidad nominal en relación con la cláusula. La diversidad nominal tiene un efecto semejante al de la densidad nominal, mientras que la diversidad léxica mostró ser el indicador de menor peso sobre la calificación. Todas estas interacciones demostraron tener poder estadístico (p

siempre menor a 0.05); un considerable efecto sobre la calificación (F superior a 3), y explican en buena medida la variación de PromEA (R^2 de 0.14 a 0.20).

Enseguida se explican uno a uno estos efectos con apoyo gráfico, considerando la variable con la que interactúa cada indicador.

1. Una densidad nominal alta (40%-50% de sustantivos en el léxico) tiene efecto positivo sobre la calificación en textos de alta diversidad léxica (VocD 80-100). Es decir, en textos de léxico variado, es bien evaluado que haya muchos sustantivos. Los escenarios de exceso (muy superiores a la media) y defecto (muy por debajo de la media) pierden precisión en la predicción de la calificación (ver Figura 3).

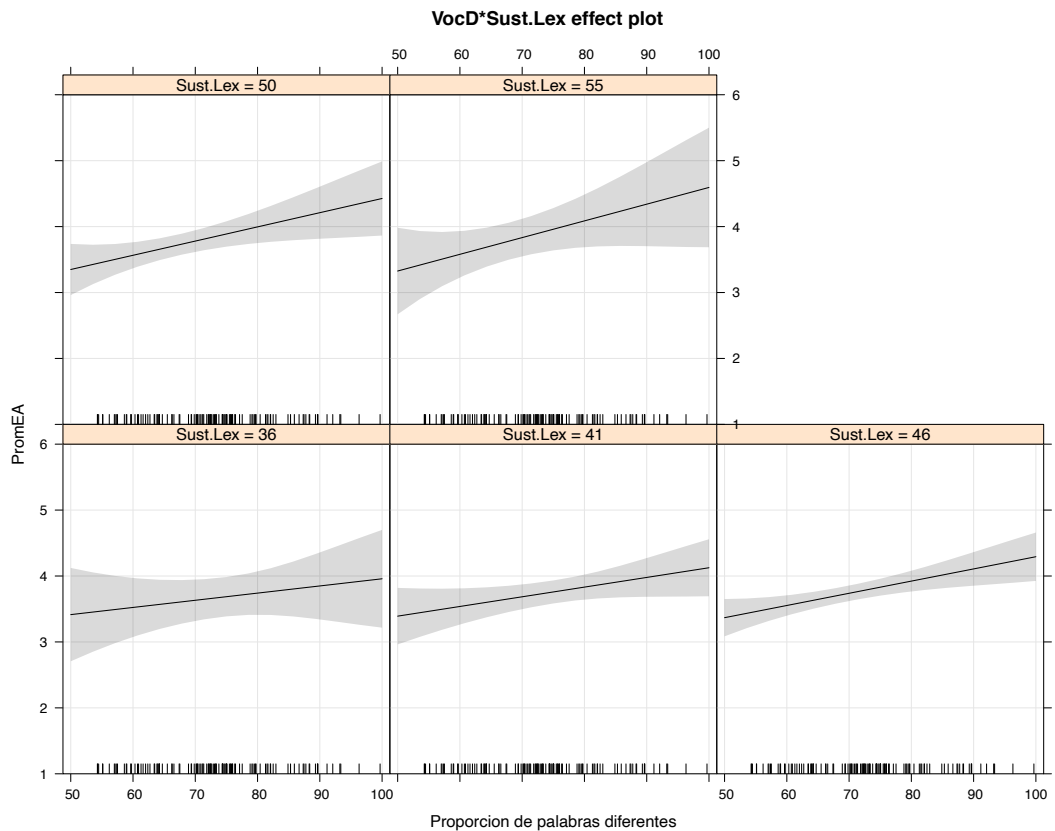


Figura 3. Efecto de la interacción entre la densidad nominal y la diversidad léxica sobre la calificación.

2. En cuanto a la densidad léxica, las proporciones altas de palabras de contenido semántico (47%-52% del total de palabras) son bien valoradas en los textos largos (>800 palabras), pero no en los cortos, lo cual indica un equilibrio entre la densidad léxica y la extensión de los textos. Los escenarios extremos, con muy pocas palabras de contenido semántico (36%) o demasiadas (>55%) restan precisión al modelo porque hay más variabilidad en la calificación (ver Figura 4).

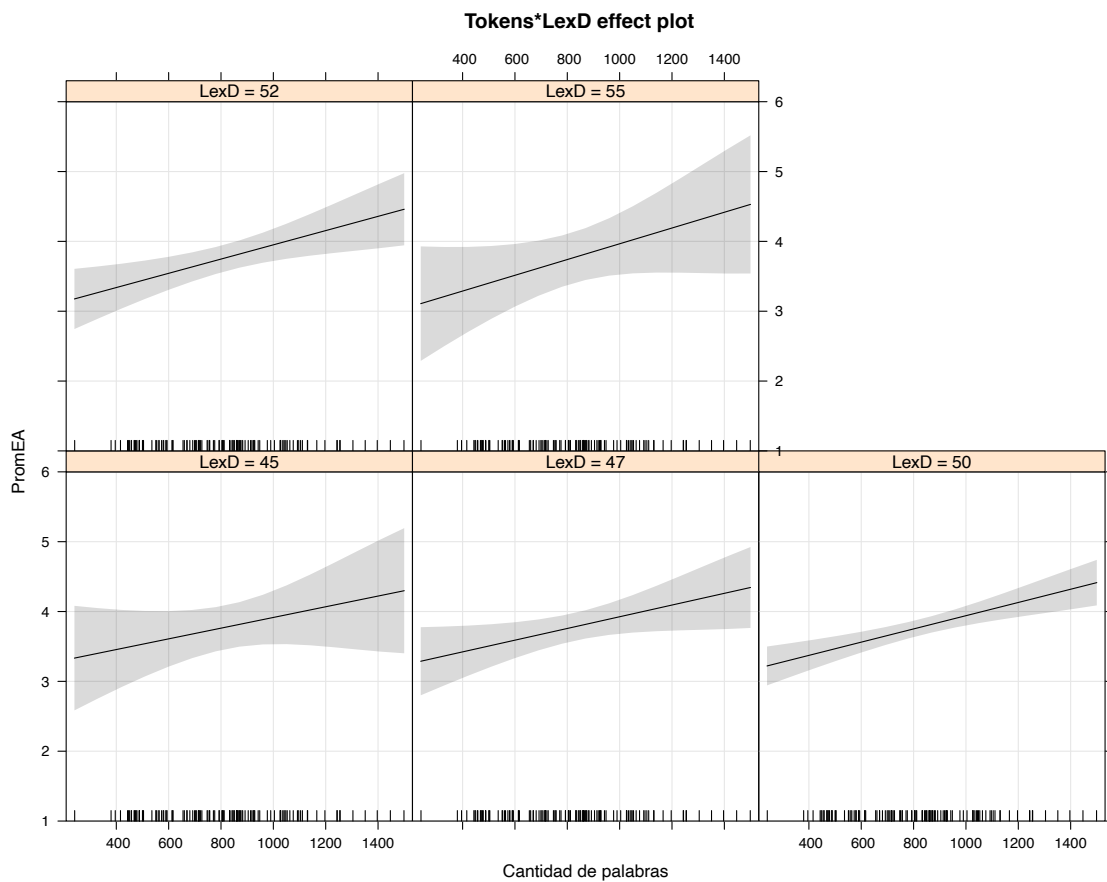


Figura 4. Efecto de la interacción entre la densidad léxica y la cantidad de palabras en el texto sobre la calificación.

3. La densidad nominal en el léxico y al interior de la cláusula revela también un equilibrio: una proporción alta de sustantivos sobre el total de las palabras de contenido (>45%) es bien evaluada, siempre y cuando se distribuyan como máximo dos por cláusula. Un escenario con más de dos o menos de un sustantivo por cláusula resta precisión al modelo predictivo, es decir, hay más variabilidad en la calificación (ver Figura 5).

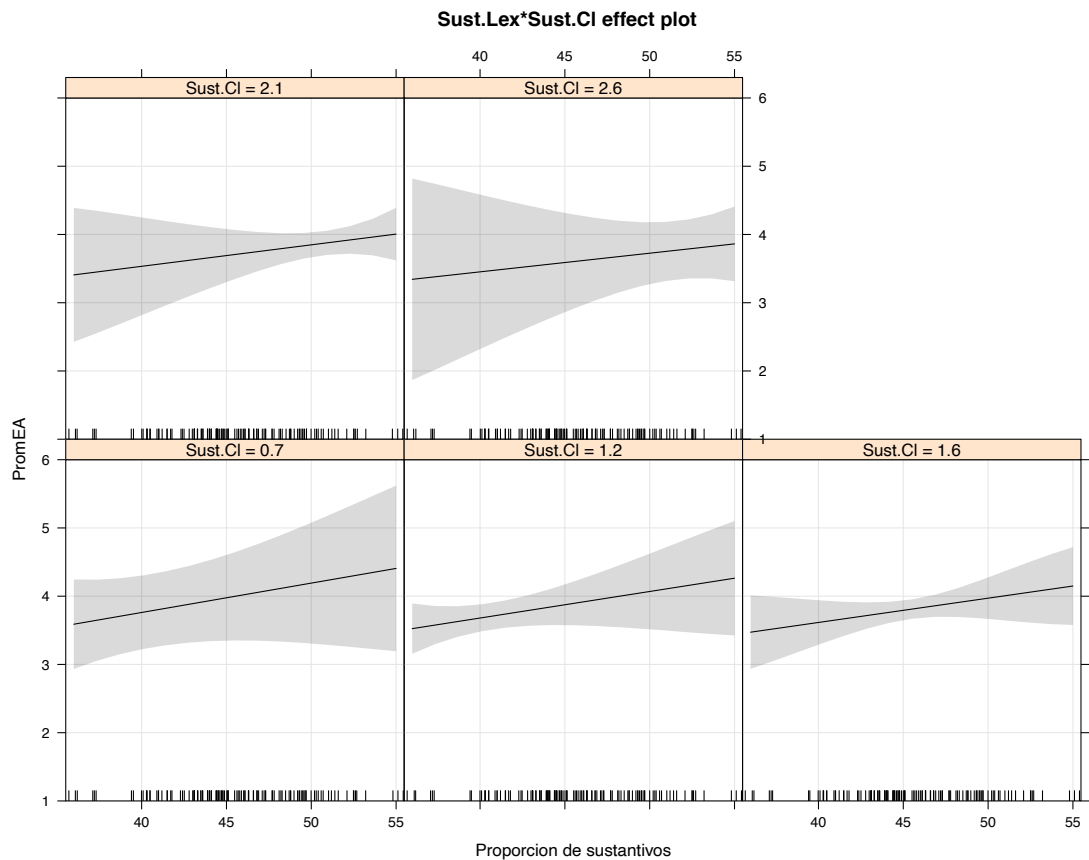


Figura 5. Efecto de la interacción entre la densidad nominal en el léxico y en la cláusula sobre la calificación.

4. Con respecto a la diversidad nominal, esta entra en interacción con la densidad nominal descrita en el párrafo anterior. Una proporción superior al 45% de sustantivos en el léxico es bien valorada, siempre y cuando la mitad de ellos, al menos, sean diferentes. El que haya menos tipos diferentes de sustantivos (41%-49%) parece ser penalizado, como se ve en la variabilidad de la calificación en ese escenario. Esto significa que la frecuente repetición de sustantivos no es bien vista por los jueces (ver Figura 6).

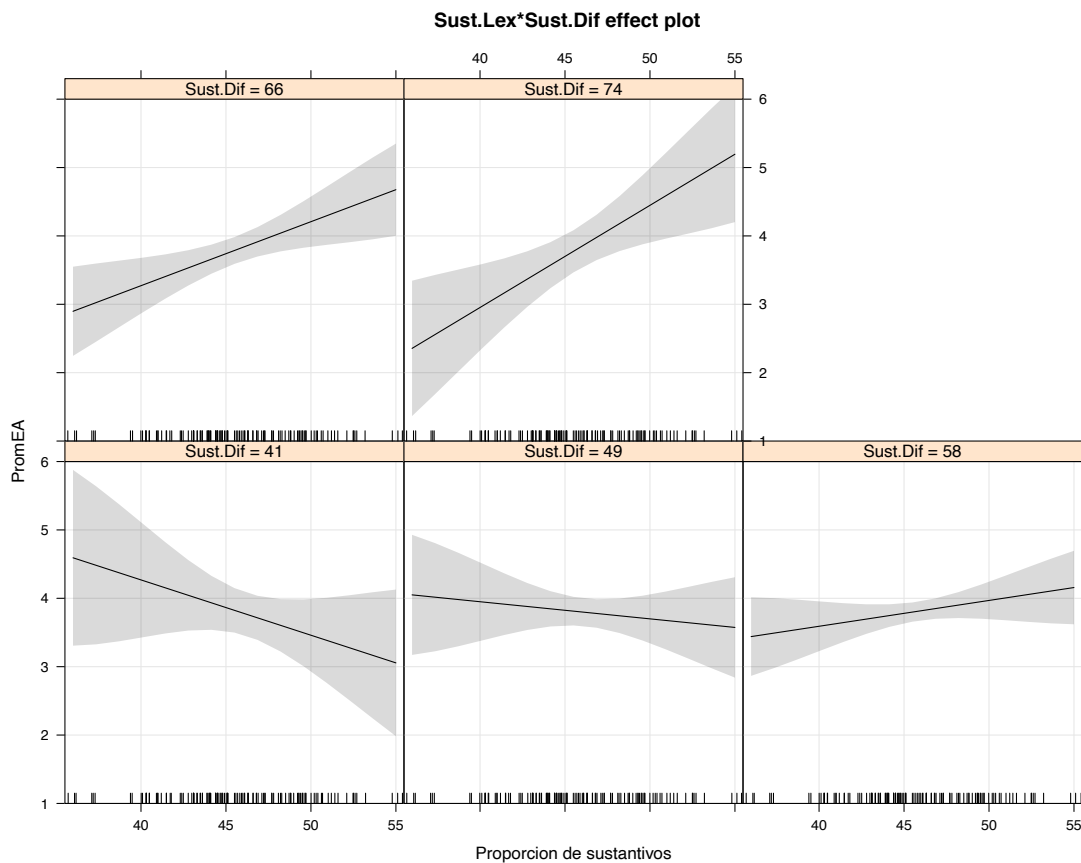


Figura 6. Efecto de la interacción entre la densidad y la diversidad nominal sobre la calificación.

5. Finalmente, una diversidad de vocabulario media a alta (VocD 50-80) es bien valorada en textos de más de 800 palabras; es decir, por arriba de la media. En contraste,

una diversidad demasiado alta ($VocD > 90$) repercute en una disminución de la calificación, incluso si los textos son largos (ver Figura 7).

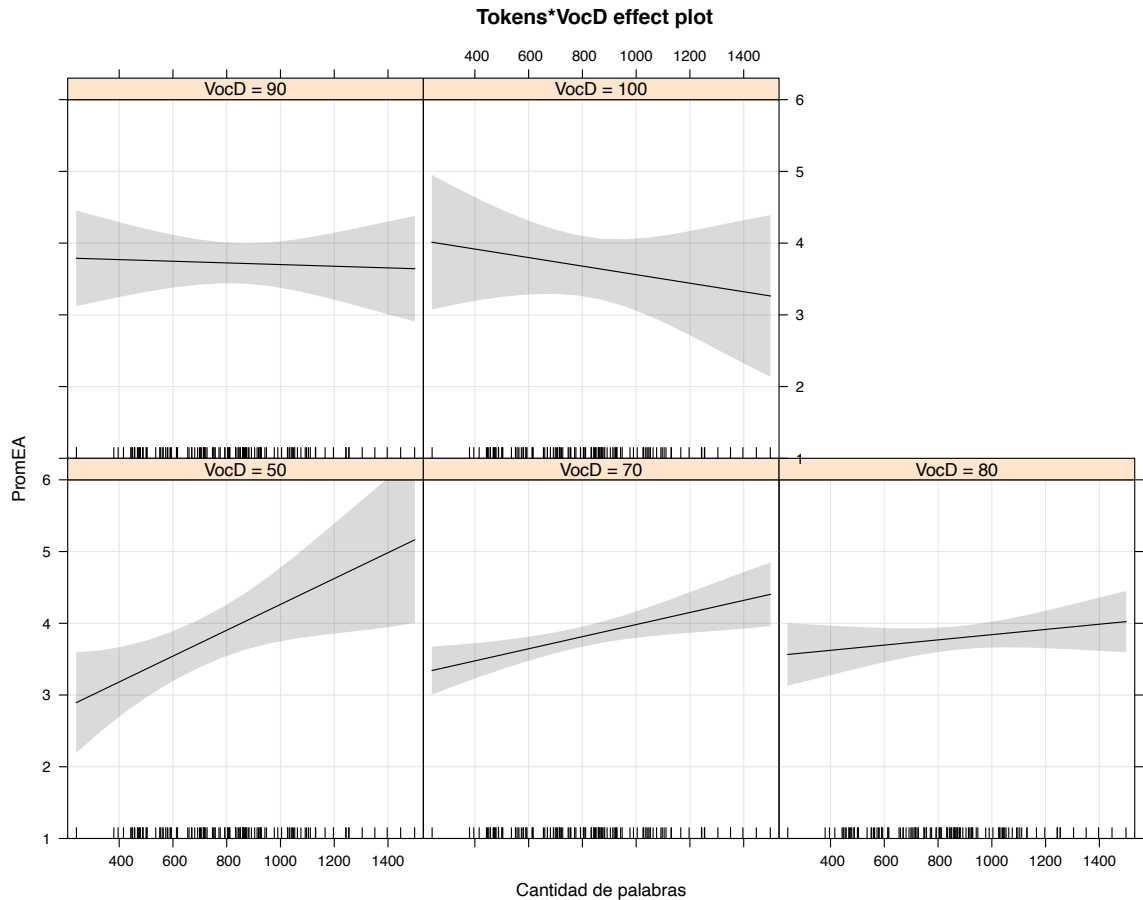


Figura 7. Efecto de la interacción entre la diversidad léxica y la extensión de los textos sobre la calificación.

En resumen, los cinco indicadores analizados –extensión de los textos y características de uso del léxico– demostraron tener efecto sobre la calificación otorgada por los jueces, pero no de manera independiente. Solo su interacción con algún otro

componente permite predecir la ubicación del texto en un rango más o menos preciso de calificación. De manera general, se hace visible un cierto patrón de equilibrio en los usos del léxico que predice valoraciones positivas, evidente en la penalización de los extremos por carencia o por exceso en cada escenario de interacción analizado. Vale la pena, sin embargo, notar que no fue la diversidad léxica (entendida como el uso de un léxico variado en los textos) lo que tuvo más peso en la calificación, sino que fueron las mediciones de densidad las que ofrecieron una mayor precisión como predictores de dominio de la escritura académica.

6.2. Análisis II. Sustantivos abstractos

El propósito de este segundo acercamiento a los datos fue observar las diferencias de uso de los sustantivos en los textos, específicamente de los sustantivos abstractos y sus clases, y si estas diferencias permiten hacer inferencias sobre su impacto en la calificación otorgada a los textos. En otras palabras, se trató de determinar la viabilidad de los sustantivos abstractos empleados en la construcción del texto como indicadores de dominio de la escritura académica. La pregunta a contestar en esta etapa fue: ¿En qué proporciones se encuentran los sustantivos abstractos y sus clases en el corpus y qué efecto tiene el uso de este léxico abstracto en los textos sobre la variación en la calificación de una muestra agrupada por su nivel de dominio de la escritura académica?

6.2.1. Los sustantivos abstractos en el corpus

La exploración sobre el corpus completo muestra un uso amplio de sustantivos en los textos. En total se contabilizaron 22,225 sustantivos.³⁵ De esta frecuencia acumulada, el 40% cumple con la caracterización semántica de la abstracción. La Figura 8 presenta las

³⁵ En otra parte de este trabajo se abordó la densidad y la diversidad nominal; solo como memorando, de las 48,893 palabras de contenido en el corpus, 22,225 (45.5%) son sustantivos. Entre estos se identificaron 1,768 tipos diferentes, de los cuales 921 son sustantivos abstractos y 847 son sustantivos concretos.

frecuencias simples (8,872 sustantivos abstractos y 13,353 sustantivos concretos), lo mismo que su expresión como proporciones (40% y 60%, respectivamente).

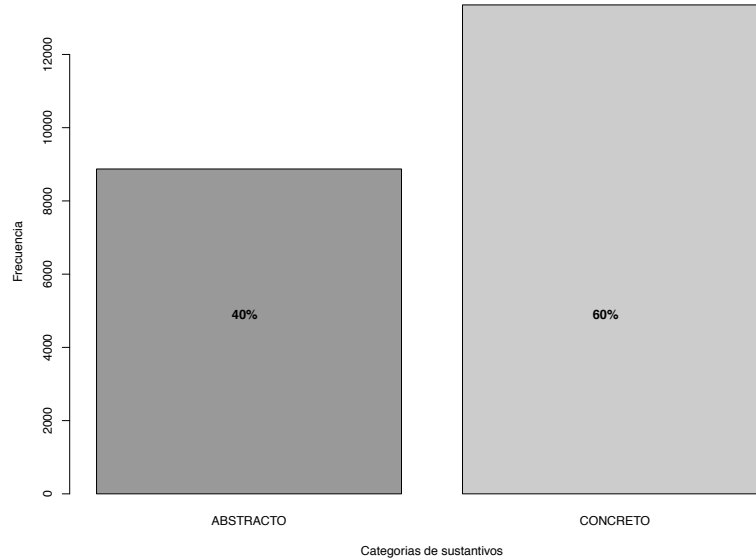


Figura 8. Frecuencias y proporciones de las categorías de sustantivos en el corpus.

Por lo que toca al uso de las distintas clases de sustantivos abstractos, en el corpus se encontró que la clase más frecuente (con 4,889 instancias) es la de ‘eventos del mundo físico y estados de cosas’, muchos de los cuales son sustantivos derivados de otras clases funcionales (nominalizaciones de verbos o adjetivos como diferente>*diferencia*, generar>*generación*, mover>*movimiento*, solo>*soledad*), mientras que la base o el origen de algunos otros es difícil de determinar.³⁶ Hay menores cantidades de las clases restantes,

³⁶ Aun se podría abundar más en una caracterización morfológica de esta clase de sustantivos abstractos. En este trabajo se emplearon criterios semánticos para categorizarlos, pues el cuerpo central de la teoría consultada admite que hay nominalizaciones en distintas clases de abstractos, tanto los que refieren eventos, procesos y resultados como en los que señalan cualidades y estados, y que la derivación nominal es un recurso abundantemente empleado en el discurso académico y científico (Bello, 2016; Bosque, 2011; Cademártori, Parodi & Venegas, 2006; Halliday & Martin, 1983, entre otros); asimismo, Bosque (2011) resalta que hay sustantivos abstractos no derivados como *acción*, *ambición*, *emoción*, *tensión*, etc. (todos estos del corpus) cuya etimología no puede ser objeto de este estudio. Para centrarnos en las características del léxico académico optamos por no sobre-analizar los sustantivos abstractos.

en este orden: ‘procesos del mundo mental y estados afectivos’ (1,579 instancias), muchos de ellos derivados de adjetivos como indiferente>*indiferencia*, perseverar/perseverante>*perseverancia*, triste>*tristeza*, o de verbos cognitivos como pensamiento>*pensar*, comparar>*comparación*, reflexionar>*reflexión*; en el tercer sitio están los ‘constructos teóricos y conceptos’ (1,305 instancias), como *sentido*, *ser*, *tiempo*, *realidad* o *valor*, y en último lugar los ‘encapsuladores’ (1,099 instancias), como *tema*, *caso*, *hecho*, *tipo* o *meta*. Estas proporciones se muestran en la Figura 9.

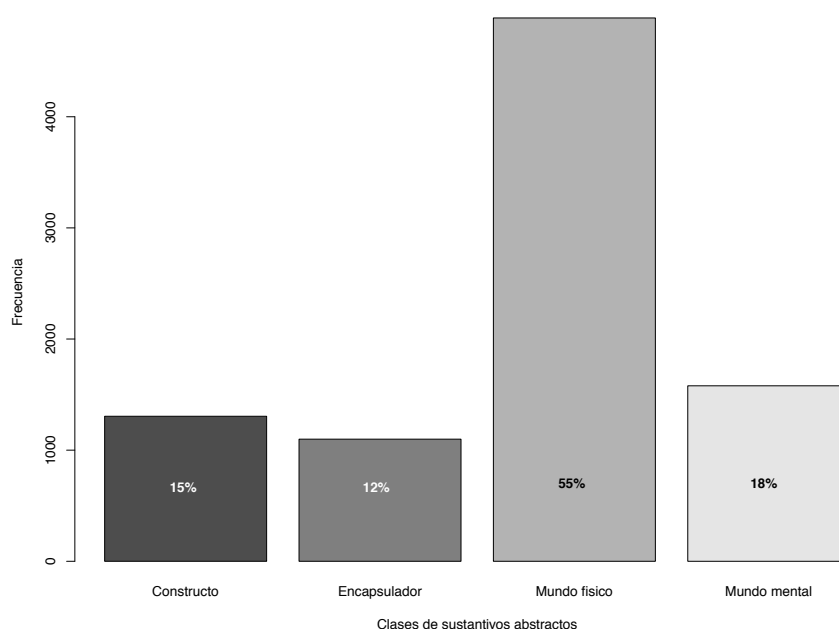


Figura 9. Frecuencias y proporciones de las clases de sustantivos abstractos en el corpus.

6.2.2. Los sustantivos abstractos en la muestra

En la muestra de 60 textos de los niveles alto y bajo del indicador de dominio PromEA las proporciones de las categorías de sustantivos reflejan las del corpus (concretos 61%, abstractos 39%, aproximadamente) sin importar el nivel de PromEA; pero la frecuencia de

uso es más alta en los textos mejor evaluados: 4,135 concretos (60.83%) y 2,663 abstractos (39.17%) en el grupo de textos de nivel alto, y 2,653 concretos (60.68%) y 1,719 abstractos (39.32%) en el grupo de textos de nivel bajo, lo cual se concentra en la Figura 10.

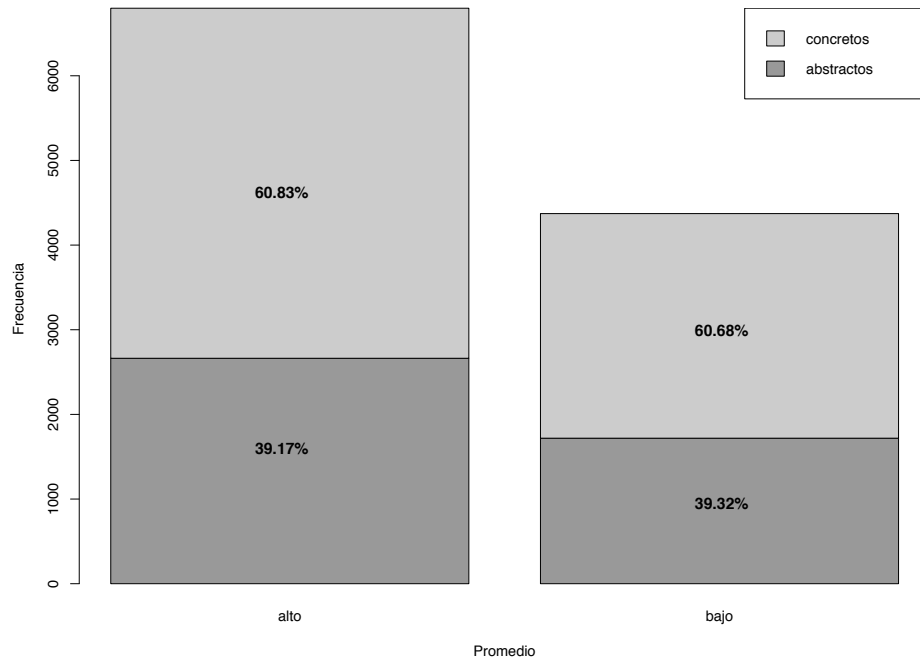


Figura 10. Frecuencias y proporciones de las categorías de sustantivos en la muestra de textos.

Por su parte, el uso de las distintas clases de sustantivos abstractos en la muestra varía en función del nivel de PromEA: si bien se mantiene en primer lugar de frecuencia la clase de ‘eventos del mundo físico y estados de cosas’, en los textos mejor evaluados las clases ‘constructos teóricos y conceptos’ y ‘procesos del mundo mental y estados afectivos’ se igualan en el segundo lugar, en tanto que en los textos peor evaluados bajan las instancias de ‘constructos teóricos y conceptos’ en relación con el resto. Como se observa en la Figura 11, nuevamente es la frecuencia de cada clase la que permite precisar

diferencias entre los dos niveles del indicador PromEA, aunque las proporciones al interior de los grupos se asemejan.

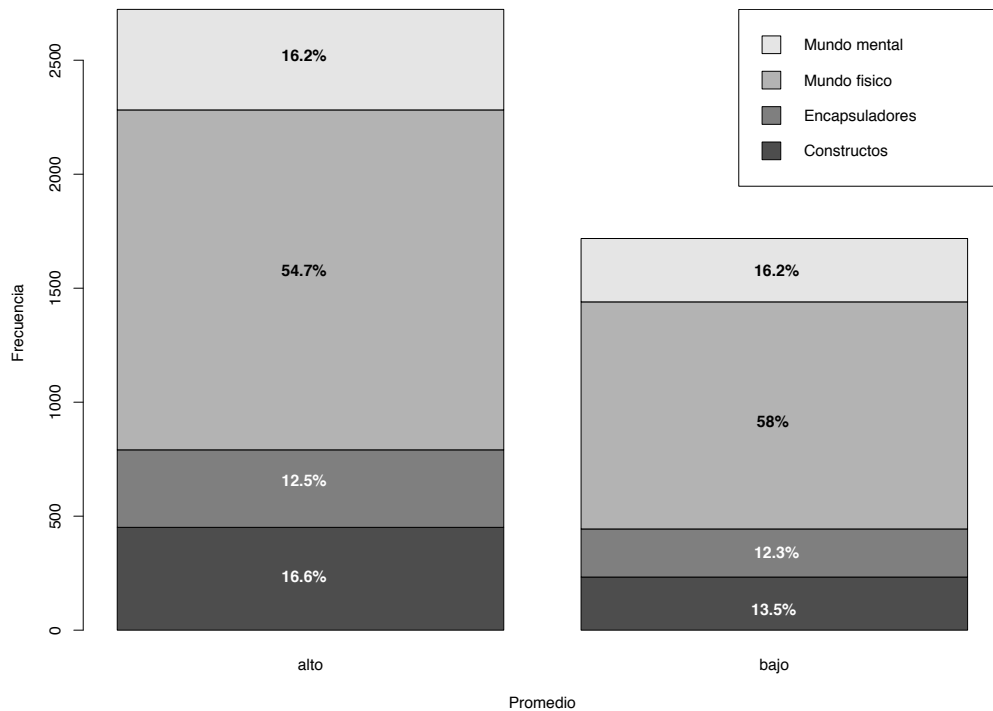


Figura 11. Frecuencias y proporciones de las clases de sustantivos abstractos en la muestra de 60 textos.

6.2.3.Efectos sobre el nivel de PromEA

Con respecto al potencial de los sustantivos abstractos como indicadores de dominio de la escritura académica, se planteó la pregunta sobre si un uso diferente de las categorías y clases de sustantivos en los textos podría predecir la ubicación de los mismos en los niveles alto o bajo del indicador PromEA. Para ese fin, en primer lugar se corrió una prueba Chi cuadrado con las proporciones de cada categoría según el nivel en que se sitúan los textos. Como se preveía, dado que la distribución de las dos categorías ‘concreto’ y ‘abstracto’ en cada nivel de la muestra reproduce casi sin cambio la distribución de las mismas en el

corpus, la prueba confirmó que el promedio no depende de la proporción de concretos y abstractos ($X^2=0.017$, $gl=1$, $p=0.89$).

Es diferente lo que sucede con las proporciones de las distintas clases de sustantivos abstractos: una segunda prueba de independencia Chi cuadrado demostró que el promedio sí tiene relación con la forma diferenciada en que se presenta cada clase de abstractos en la muestra ($X^2=8.14$, $gl=1$, $p<0.05$). Para profundizar en esta relación, y verificar el aporte de cada categoría y clase en la predicción de los niveles alto o bajo del promedio, se realizó una serie de pruebas t de Student para comparar las medias aritméticas de cada medición.³⁷ Para ello, sirven como referente los datos de la Tabla 10, que resume los estadísticos de los textos de la muestra agrupados por su nivel en el indicador PromEA.

Tabla 10. *Medias de uso de las categorías y clases de sustantivos según el nivel de PromEA.*

Categoría/clase	PromEA	
	Alto Media (DE)	Bajo Media (DE)
Sustantivos	214.8 (67.0)	145.7 (48.1)
Concretos	129.5 (49.4)	88.4 (30.5)
Abstractos	88.8 (23.5)	57.3 (22.0)
Mundo físico	49.7 (15.4)	33.2 (13.7)
Constructos	15.0 (9.5)	8.2 (6.0)
Mundo mental	14.4 (8.1)	9.3 (5.0)
Encapsuladores	11.3 (7.6)	6.6 (4.0)

Nota: DE=Desviación estándar.

³⁷ Todas las variables bajo análisis (frecuencia de uso de sustantivos, de sustantivos abstractos y de cada clase de sustantivos abstractos) cumplieron las condiciones de normalidad en la distribución y de homocedasticidad (pruebas Levene y Fligner-Killeen).

El análisis discriminante arrojó los resultados que se resumen en la Tabla 11.

Tabla 11. *Estimación de los efectos de las categorías y clases de sustantivos sobre el nivel de PromEA.*

Variable (frecuencia)	t	p <
Sustantivos abstractos	5.35	0.001
Sustantivos concretos	4.69	0.001
Sustantivos totales	4.58	0.001
Eventos del mundo físico y estados de cosas	4.36	0.001
Constructos teóricos y conceptos	3.35	0.01
Encapsuladores	3.00	0.01
Procesos del mundo mental y estados afectivos	2.91	0.01

Notas: Pruebas t de Student para grupos independientes con 58 grados de libertad, $\alpha=0.05$. Ordenadas por el tamaño del efecto (t).

Con estos resultados, es posible jerarquizar las variables por su potencial para discriminar los dos niveles de la variable dependiente PromEA. De entrada, tanto por la significancia estadística como por el valor del coeficiente de la prueba, la presencia misma de sustantivos resulta ser un factor de diferenciación entre los textos mejor y peor evaluados. Con una media de 214.8 sustantivos en los primeros y de 145.7 en los segundos, esta brecha predice con alta precisión el nivel de dominio de la escritura académica percibido por los jueces, como se verificó en el análisis previo (ver apartado 6.1.2). En la Figura 12 se patentiza que el rango alrededor de la mediana en cada grupo de textos es diferente, más bajo en los textos de promedio bajo.

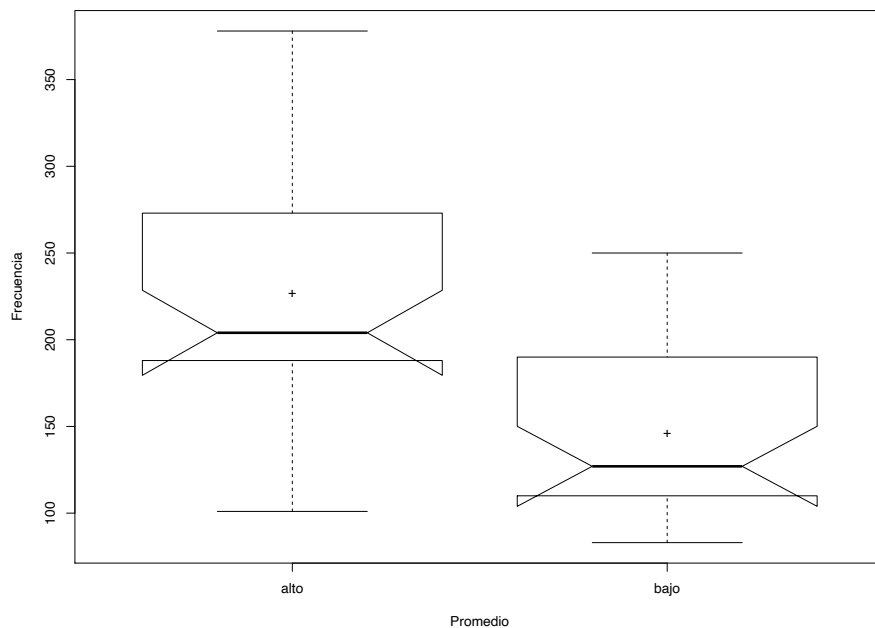


Figura 12. Uso de sustantivos en la muestra, según el nivel de PromEA.

Enseguida, la media de uso de los sustantivos abstractos predice también con alta precisión la ubicación de un texto en uno de los dos niveles de dominio, de lo cual se puede inferir que los jueces son sensibles al empleo de este tipo de sustantivos en la construcción del discurso, que al parecer destacan de los concretos aun cuando estos son empleados con mayor abundancia. El hecho de que la diferencia en las medias de uso de los sustantivos concretos (129.5 frente a 88.4) sea también estadísticamente significativa probablemente entre dentro del mismo fenómeno expuesto antes: que la frecuencia de los sustantivos en general es bien valorado en este tipo de textos. La Figura 13 deja clara la diferencia de uso de los sustantivos abstractos en los niveles de PromEA (88.8 en los textos de promedio alto frente a 57.3 en los de promedio bajo), dato que se ajusta al carácter distintivo de la abstracción en el lenguaje académico, tal como se describió en el marco teórico de este trabajo.

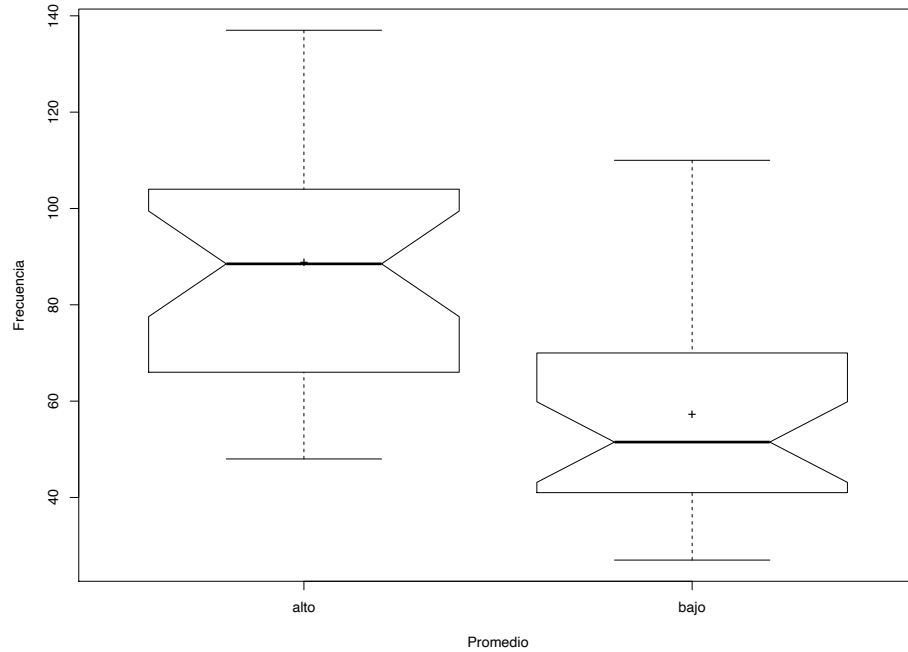


Figura 13. Uso de sustantivos abstractos en la muestra, según el nivel de PromEA.

Por lo que toca a las clases de sustantivos abstractos, estas se jerarquizan, según su poder estadístico como variables discriminantes del nivel de dominio adjudicado a un texto, de la siguiente manera: primero, los abstractos de ‘Eventos del mundo físico y estados de cosas’, seguidos de ‘Constructos teóricos y conceptos’, luego los ‘Procesos del mundo mental y estados afectivos’, y finalmente los ‘Encapsuladores’.

La primera clase de sustantivos abstractos es muy frecuente en ambos grupos de textos (medias de 49.7 en los de promedio alto frente a 33.2 en los de promedio bajo), pero los rangos son diferentes, como se patentiza en la Figura 14.

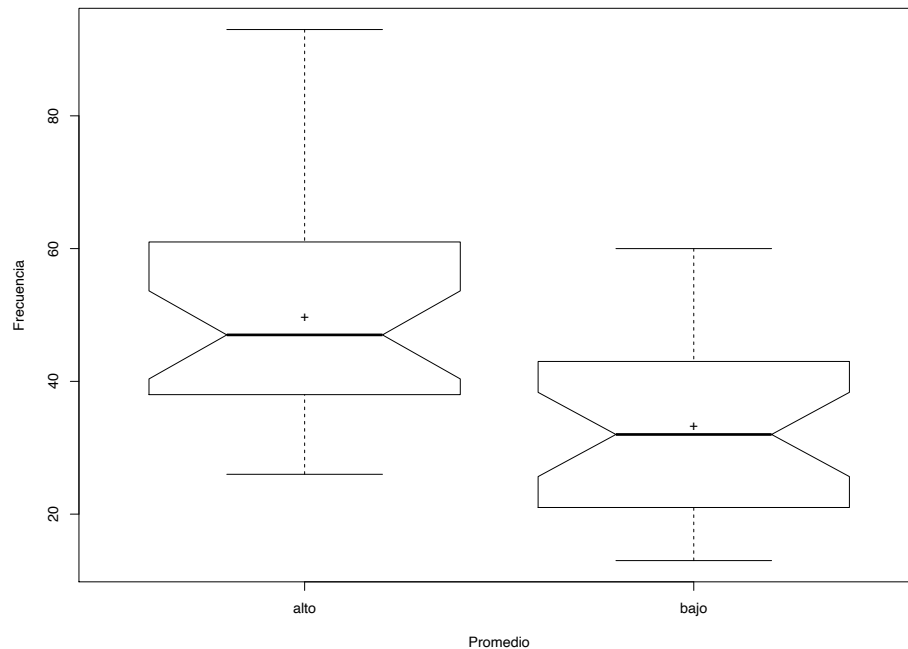


Figura 14. Uso de ‘Eventos del mundo físico y estados de cosas’ en la muestra, según el nivel de PromEA.

A partir de la clase ‘Constructos teóricos y conceptos’ la media de uso disminuye en ambos grupos de textos (15 frente a 8.2 en el mismo ordenamiento), pero nuevamente la diferencia en los rangos es el factor discriminante del nivel de dominio, como se patentiza en la Figura 15, según la cual en los textos de nivel alto la distribución va de 10 a 20 de estos abstractos, con al menos un caso extremo más allá de las 30 instancias, mientras que el uso en los textos de nivel bajo no rebasa los 12, con pocas excepciones.

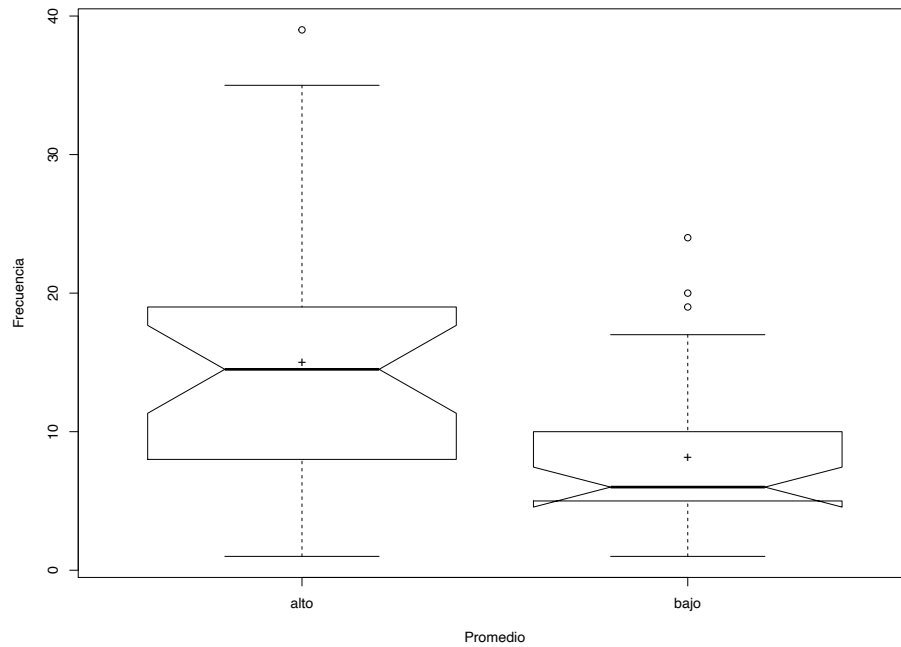


Figura 15. Uso de ‘Constructos teóricos y conceptos’ en la muestra, según el nivel de PromEA.

La penúltima clase en cuanto a su poder estadístico para discriminar un texto bien evaluado de uno mal evaluado es la que refiere ‘Procesos del mundo mental y estados afectivos’, con medias de 14.4 y 9.3 respectivamente. La Figura 16 muestra rangos alrededor de la mediana en cada agrupación a punto de solaparse, lo cual se refleja en el valor de la probabilidad (ver también Tabla 11).

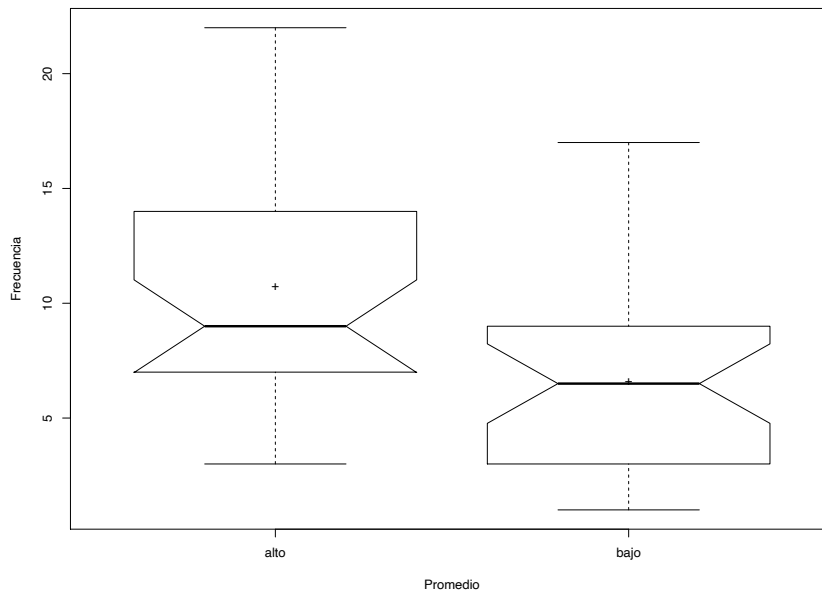


Figura 16. Uso de 'Procesos del mundo mental y estados afectivos' en la muestra, según el nivel de PromEA.

Finalmente, en la Figura 17 es evidente que un mayor uso de sustantivos abstractos encapsuladores también contribuye a diferenciar la ubicación de un texto entre los mejor o peor evaluados (11.3 frente a 6.6).

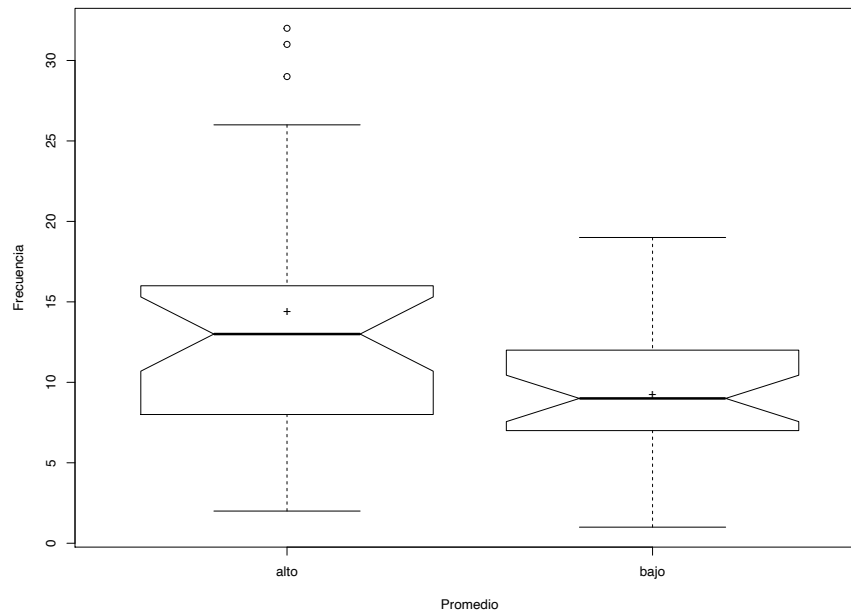


Figura 17. Uso de 'Encapsuladores' en la muestra, según el nivel de PromEA.

Vistos en conjunto, los sustantivos resultan influyentes en la puntuación de los jueces, como se había verificado en el análisis de la densidad nominal (ver apartado 6.1.2); empero, la frecuencia con la que se emplean las categorías ‘concreto’/‘abstracto’, y las clases de abstractos, muestra tener un peso diferenciado en esa valoración.

6.2.4. Ejemplos de los sustantivos abstractos en el corpus

A continuación se muestran algunos ejemplos de las cuatro clases de sustantivos abstractos usados por los estudiantes en la construcción de sus textos, tomados de la lista completa del corpus.

1) “(...) la *sociedad* no tiene un título de *propiedad* sobre las personas para decidir quién es apto para vivir o no, eso lo decidirá una *fuerza* superior” (14M13).

2) “El viejo en el mar siente *dolores* físicos pero también tiene *sentimientos* y *pensamientos* que no lo son, que evocan *alegrías*, *tristezas*, *esperanzas*, *lástima*” (14M11).

En (1) se tienen tres sustantivos abstractos de la clase ‘Eventos del mundo físico y estados de cosas’: *sociedad* refiere a un colectivo, *propiedad* a un estado, mientras que *fuerza* es una cualidad; todos los sustantivos del fragmento en (2) son de la clase ‘Procesos del mundo mental y estados afectivos’, aunque *dolor* y *pensamiento* tienen la particularidad, el primero de ser una sensación física, y el segundo el resultado de un proceso mental: *pensamientos*, de la misma clase que *razonamiento* en el texto de (3).

3) “Ambos autores aportaron *teorías* sobre el *razonamiento* humano en *condiciones* complejas, como lo es estar solo en altamar o frente a un tribunal” (14M16).

4) “(...) los *temas* principales que reflejaba en la mayoría de sus obras eran el *sentido del honor*, *la acción*, *lucha* y *entrega* de los personajes, *coraje*, entre muchos otros” (15M08).

En (3) y en (4) se presentan ‘Constructos teóricos y conceptos’ como *teorías, sentido, honor*; ‘Eventos del mundo físico y estados de cosas’ como *condiciones, acción, lucha*, y estados afectivos como *entrega y coraje*.

5) [“La situación social de esa época no permitía a los escritores expresarse con libertad, la censura actuaba de forma despiadada y radical, así que recurrían a otro tipo de métodos literarios para expresar su posición, contraria a lo que se vivía y su afán de cambiarlo. Camus fue uno de estos artistas que crispados con la <=situación decidió manifestarse discretamente a través de sus obras, en este caso,=> a través de El Extranjero.”] Este es un <=dato importante para comprender el significado dentro de esta obra” (15H30).

En (4) y (5) se muestran ejemplos de encapsuladores cumpliendo dos de las principales funciones que se han documentado para esta clase de sustantivos abstractos en el español: primero, la recuperación de un elemento textual, bien sea que haya sido referido previamente (en cuyo caso se realiza una función cohesiva anafórica), o de manera pospuesta al encapsulador (en cuyo caso la función cohesiva es catafórica), y segundo, la categorización, recurso de carácter más metadiscursivo que sintáctico, pues otorga una etiqueta al segmento textual que refleja una cierta posición del enunciador ante su dicho (Francis, 1994; López, A., 2011; Schmid, 2000).

El sustantivo abstracto *temas* del ejemplo (4) encapsula, catafóricamente, todos los elementos que complementan al verbo en la cláusula nominal subrayada; mientras que el fragmento de (5) presenta tres casos de encapsulación: *situación*, que recupera la información de la cláusula que abre el párrafo; *caso*, que anuncia y categoriza el ejemplo que le sigue, y finalmente *dato*, el cual encapsula la totalidad de la información precedente y la categoriza para fines de la atención del lector. Estas funciones de los encapsuladores han sido documentadas de manera amplia en otros estudios sobre lenguaje académico

escrito en diversos contextos (Aktas, 2005; Aktas y Cortes, 2005; Benítez-Castro, 2015; Borreguero, 2018; Díez, 2018; Flowerdew, 2003, 2006, 2015; Flowerdew & Forest, 2015; López, A., 2011, 2013, 2015, 2018; Pecorari, 2014; Tárnayková, 2018; Schmid, 1997, 1999a, 1999b, 2000, 2018), por lo que en esta investigación solo se contabilizan y ejemplifican.

6.3. Análisis III. Léxico especializado

El propósito de este tercer acercamiento a los textos del corpus fue observar los usos disciplinares de los sustantivos, es decir, identificar los sustantivos funcionando como términos especializados: *elementos nominales de carácter científico, técnico o específico de una materia o disciplina*, tal como se definieron en el apartado 5.3.2, y que configuran un léxico especializado. También, se propuso determinar si hay diferencias en el empleo del léxico especializado que permitieran hacer inferencias sobre su impacto en la calificación otorgada a los textos, para poder determinar su viabilidad como indicador de dominio de la escritura académica.

La pregunta a contestar en esta etapa fue: ¿Qué proporción de los sustantivos en el corpus son términos especializados de algún campo disciplinar, y qué efecto tiene el uso de este léxico especializado en los textos sobre la variación en la calificación de una muestra agrupada por su nivel de dominio de la escritura académica?

6.3.1. El léxico especializado en el corpus

En el corpus se identificaron 1,779 sustantivos diferentes (tipos); de estos, 226 se consideraron términos especializados (12.7% de los tipos) de acuerdo con los criterios expuestos en el Apartado 5.3.2.7; los tipos restantes fueron considerados sustantivos de uso común, es decir, no especializados. En cuanto a sus frecuencias, los términos especializados

reúnen 6,433 instancias, lo que constituye casi un tercio de los 22,225 sustantivos del corpus. La Figura 18 muestra esta última proporción.

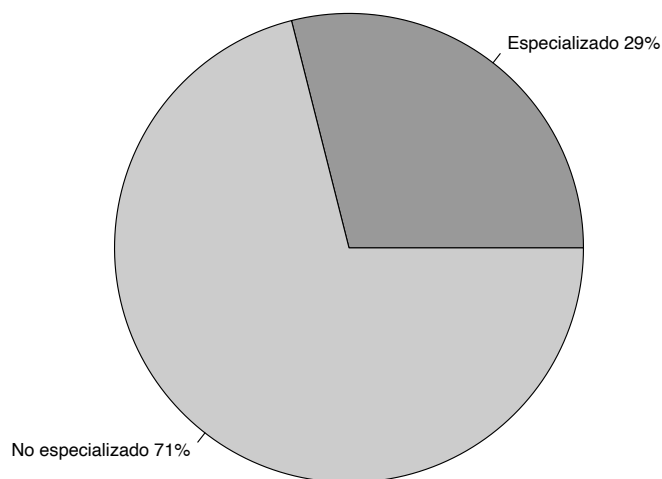


Figura 18. Proporciones del léxico especializado y no especializado en el corpus.

El gráfico permite valorar la penetración del léxico especializado en el discurso construido por los estudiantes, pues aunque el repertorio de términos especializados sea una proporción modesta del listado de sustantivos en el corpus, su uso, en cambio, es muy frecuente: 29 de cada 100 instancias de sustantivos en el corpus están funcionando como términos especializados.

El campo disciplinar de mayor presencia en el léxico especializado empleado por los estudiantes está claramente influenciado por la materia escolar en que se contextualizó la situación de escritura, pues 161 de los términos pertenecen al campo disciplinar de ‘Lengua y literatura’, y son también los más frecuentemente empleados; los restantes 65 términos especializados pertenecen la categoría ‘Otras’, o sea, a otras disciplinas del campo de las humanidades y las ciencias sociales, como la economía, la sociología o la psicología, etc., y se encuentran menos abundantemente en el corpus, como se observa en la Figura 19.

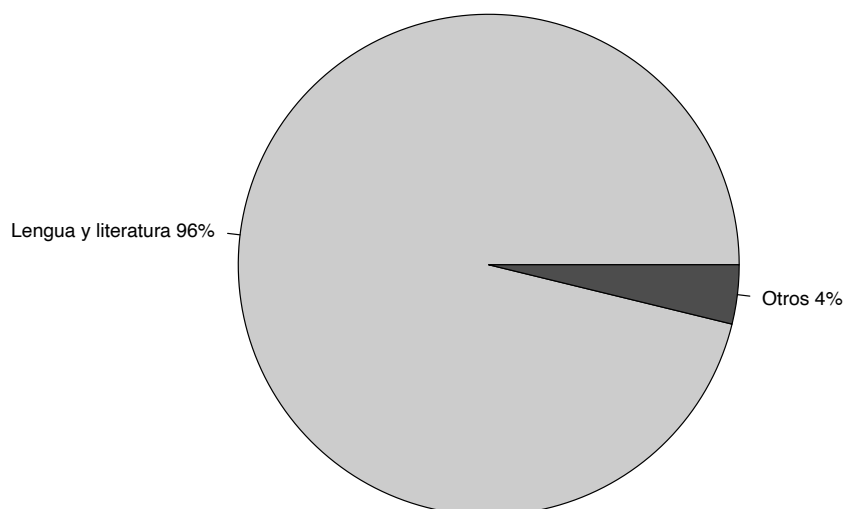


Figura 19. Proporciones del léxico especializado por campo disciplinar en el corpus.

6.3.2. El léxico especializado en la muestra

Para examinar la relación entre el uso de léxico especializado y las variaciones en la calificación, se obtuvieron las medias de términos especializados en la muestra de 60 textos agrupados por su nivel de PromEA, información que se resume en la Tabla 12.

Tabla 12. Medias de uso de léxico no especializado y especializado por campo disciplinar, según el nivel de PromEA.

Categoría/campo	PromEA	
	Alto Media (DE)	Bajo Media (DE)
No especializado	151.9 (61.2)	101.2 (35.6)
Especializado	62.9 (18.8)	44.6 (18.5)
Lengua y literatura	60.1 (19.1)	43.0 (17.7)
Otras	2.8 (4.0)	1.6 (2.1)

Nota: DE=Desviación estándar.

La Tabla 12 muestra que las medias de léxico especializado, tanto de *Lengua y literatura* como de *Otras disciplinas humanísticas*, son más altas en los textos mejor evaluados.

Por su parte, las categorías ‘no especializado’ y ‘especializado’ se distribuyen en la muestra de una manera muy semejante: 72% y 28% respectivamente, en los textos de nivel alto, y 70% y 30% en los textos de nivel bajo en PromEA (ver Figura 20), en concordancia con lo encontrado en estudios previos (Chung & Nation, 2004).

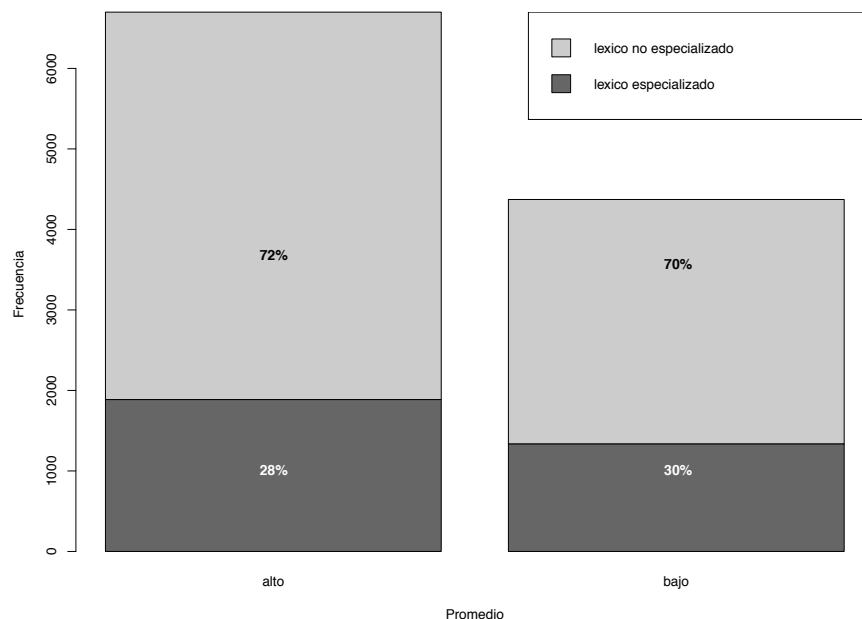
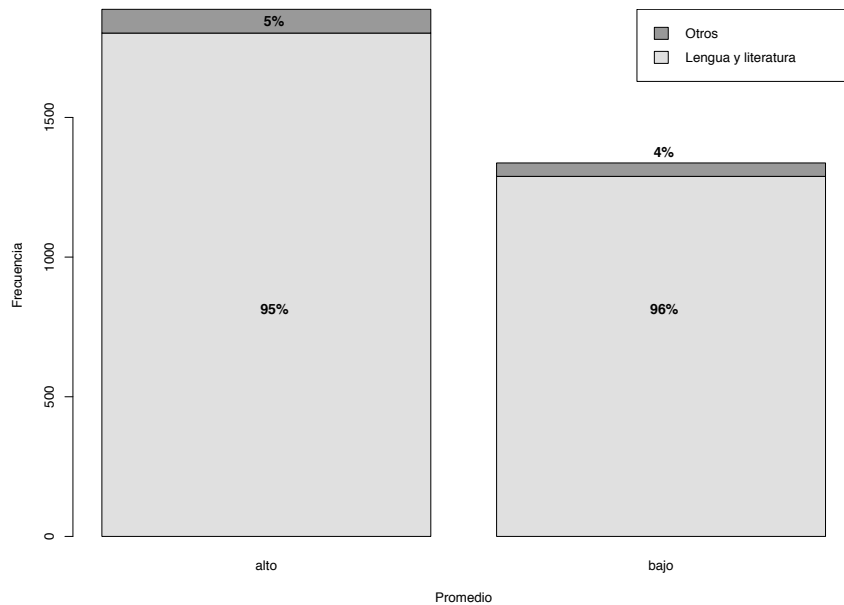


Figura 20. Proporciones del léxico especializado en la muestra.

Lo mismo sucede con las proporciones del léxico especializado de los campos disciplinares identificados. En cada grupo de textos, el léxico especializado de lengua y literatura es el más empleado, como se muestra en la Figura 21.

Figura 21. Proporciones del léxico especializado por campo disciplinar en la muestra.



Estas semejanzas en la distribución relativa parecerían suponer que estas variables no tendrían un impacto en la calificación; sin embargo, es la frecuencia absoluta en cada grupo de textos la que ofrece información más precisa sobre la varianza en PromEA. No solo la frecuencia de uso de este tipo de términos es diferente entre los textos de nivel alto y bajo de PromEA, sino también los repertorios de términos especializados empleados en cada grupo de textos (cantidad de tipos distintos): en los textos mejor evaluados hubo 23 términos diferentes en promedio, frente a 17.6 de los textos de nivel bajo; el repertorio medio de términos de lengua y literatura en los mejor evaluados fue de 21.3, frente a 16.7 del grupo de nivel bajo. El efecto diferenciador de estos datos se puso a prueba estadísticamente.

6.3.3.Efectos sobre el nivel de PromEA

Para comprobar si estas medias de repertorio y de uso son significativamente diferentes, y por tanto representan un criterio para discriminar entre los dos niveles de calificación de los

textos, se corrieron sendas pruebas t de Student,³⁸ en primer lugar con los términos especializados sin subclasificar, y luego con los de lengua y literatura. La Tabla 13 presenta los coeficientes obtenidos, ordenados por el tamaño del efecto.

Tabla 13. *Estimación de los efectos del léxico especializado sobre el nivel de PromEA.*

Variable	t	p <
Términos especializados (frecuencia)	3.80	0.001
Términos especializados (repertorio)	3.73	0.001
Términos de Lengua y literatura (frecuencia)	3.59	0.001
Términos de Lengua y literatura (repertorio)	3.40	0.01

Notas: Pruebas t de Student para grupos independientes con 58 grados de libertad, $\alpha=0.05$. Ordenadas por el tamaño del efecto (t).

Como puede apreciarse, la frecuencia de uso de términos especializados en general, y la frecuencia de uso de los términos específicos del campo disciplinar de lengua y literatura son más influyentes sobre la ubicación de los textos en un nivel de PromEA, que sus respectivos repertorios. A reserva de discutirlo más adelante, este hallazgo se asemeja a lo encontrado respecto de los sustantivos abstractos, cuya frecuencia de uso reveló tener más peso que el empleo de un determinado repertorio.³⁹

La Figura 22 muestra que en los textos con alto promedio aparecen con mayor frecuencia términos especializados, agrupados en su mayoría entre las 45 y las 75 apariciones, con un caso excepcional que rebasa ese límite superior, y con una media de uso de 62.9 términos. Los textos con promedio bajo, por su parte, alcanzan una media de 44.5

³⁸ Todas las variables bajo análisis (cantidad de tipos de términos, su frecuencia de uso, cantidad de tipos de lengua y literatura y su frecuencia de uso) cumplieron las condiciones de normalidad en la distribución y de homocedasticidad (pruebas Levene y Fligner-Killeen).

³⁹ Una serie de pruebas no paramétricas Wilcoxon sobre los repertorios de sustantivos abstractos (tipos diferentes en cada categoría y clase), cuyos resultados ya no se presentaron por economía, verificaron que no es el repertorio, sino la frecuencia de uso de estos sustantivos lo que afecta la calificación.

términos empleados, agrupados en su mayoría entre los 25 y los 60 términos. Salta a la vista que los intervalos de confianza alrededor de la mediana en ambos grupos se alejan uno del otro, lo que sostiene la hipótesis de que existe una diferencia estadísticamente significativa en el uso de léxico especializado según el nivel de PromEA alcanzado por los textos, lo cual se verificó con la prueba t ($t=3.80$, 58 gl, $p<0.001$).

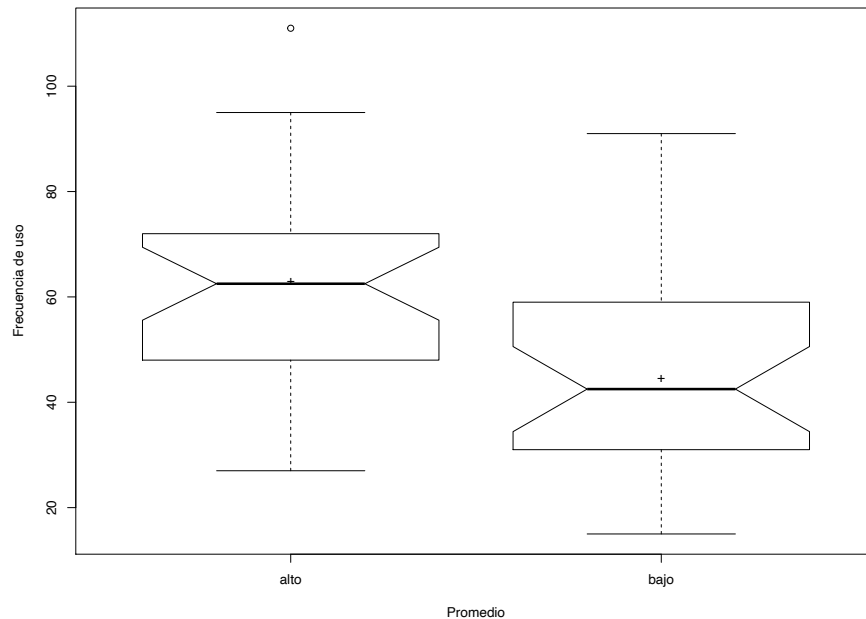


Figura 22. Uso de términos especializados según el nivel de PromEA.

Finalmente, en cuanto al uso de léxico especializado en el campo disciplinar de lengua y la literatura, se observó que en los textos con promedio alto, la frecuencia media de esta clase de términos es de 60.1, en tanto que en los de promedio bajo la media es de 43.0, de modo tal que, al parecer, la incorporación de términos de este campo disciplinar⁴⁰ repercute en la evaluación, como se muestra en la Figura 23.

⁴⁰ No se extendió el análisis al campo disciplinar restante de términos especializados, pues su repertorio y su frecuencia de uso son demasiado bajos para esperar diferencias estadísticamente significativas.

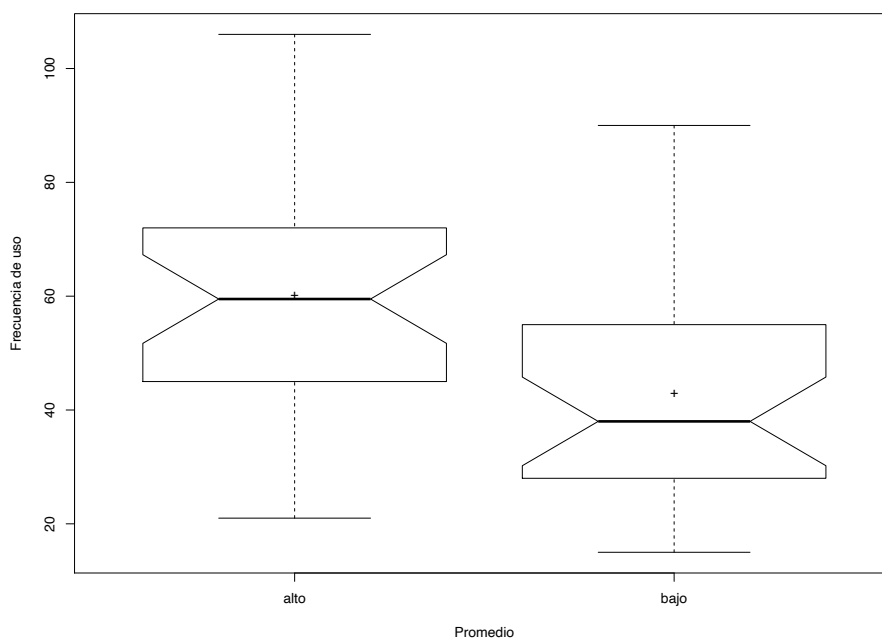


Figura 23. Uso de términos especializados de Lengua y literatura según el nivel de PromEA.

En la Figura 23 es visible una distribución mucho más homogénea en los textos mejor evaluados, mientras que en el grupo de bajo PromEA hay algunos textos con tendencia a usar más de 40 veces los términos del repertorio disponible, lo cual eleva un tanto la media sin que los intervalos de confianza lleguen a solaparse. Estas diferencias se pusieron a prueba por medio del análisis discriminante, que confirmó la hipótesis con una alta significancia estadística ($t=3.59$, 58 gl, $p<0.001$).

6.3.4. Ejemplos del léxico especializado en el corpus

Enseguida se presentan algunos ejemplos de los términos especializados empleados en la construcción de los textos, en los dos campos disciplinares identificados en el corpus:

Lengua y literatura

- 6) “Ernest Hemingway es un *autor* caracterizado por innovar en un *estilo* literario (...) este utilizaba una *sintaxis* no muy compleja, la mayoría de las *oraciones* de las que hacía uso eran sencillas así como su *léxico*” (15H27).
- 7) “Beckett utiliza este *soliloquio* para dar vida a su segunda relación enferma, la relación entre Pozzo y Lucky” (15H21).
- 8) “(...) el colombiano García Márquez se encarga de crear una *atmósfera* llena de suspenso y misterio (...) a lo largo del *relato* dentro de la cual se distinguen la *analepsis* y la *prolepsis*” (14M07).
- 9) “Esperando a Godot es una *tragicomedia* perteneciente al *teatro* absurdo escrita por el renombrado *autor* irlandés Samuel Beckett en 1952” (14M15).

Otras disciplinas

- 10) “El *existencialismo* surge de Nietzsche cuya *filosofía* se basaba en la ausencia de un dios que temer” (14M17).
- 11) “Su pensamiento se veía influenciado por el *modernismo*, el *pesimismo* y el *cosmopolitismo*” (15H17).
- 12) “Albert Camus es claro: Ni siquiera el *superhombre* puede vencer a la *posmodernidad*.” (14H19).

Se resaltan en los fragmentos de (6) a (9) palabras que emplearon los estudiantes para nombrar elementos de la narrativa y cualidades estilísticas del lenguaje literario (*estilo*, *oraciones*, *lenguaje*, *obra*), las cuales pueden encontrarse en el discurso cotidiano, pero otras no tanto, como *analepsis*, *analogía*, *sintaxis*, *soliloquio* o *tragicomedia*. Las viñetas de (10) a (12) incluyen nombres de corrientes o doctrinas (*modernismo*, *pesimismo*, *cosmopolitismo* y *posmodernidad*) y una categoría del *existencialismo* (*superhombre*). En

otros casos ya no citados aquí, se encontraron términos de las humanidades o las ciencias sociales empleados para contextualizar y explicar opiniones sobre las obras literarias, como *alexitimia, alter ego, capitalismo, depresión, economía, ética o fascismo*.

6.4. Análisis IV. Léxico abstracto especializado

El propósito de este último acercamiento a los datos fue realizar una cala en una muestra del corpus para determinar si el cruce de las categorías ‘abstracto’ y ‘especializado’ en los sustantivos empleados en la construcción de los textos configura algún repertorio particular, y si las diferencias en el uso de este léxico abstracto especializado repercuten en la calificación otorgada a los textos, vale decir, determinar su viabilidad como indicador de dominio de la escritura académica.

La pregunta que dirigió esta etapa final del trabajo de investigación fue: ¿Qué proporción de los sustantivos abstractos son a la vez términos especializados en la muestra, y qué efecto tiene el uso de este léxico abstracto especializado en los textos sobre la variación en la calificación?

6.4.1. El léxico abstracto especializado en la muestra

En la muestra de 60 textos, 30 del nivel más alto y 30 del nivel más bajo en el promedio de escritura académica PromEA, se encontró que los más frecuentes son los sustantivos concretos comunes (40%); un tercio son sustantivos abstractos comunes (31%); les siguen en frecuencia los sustantivos concretos especializados (21%), y el resto son sustantivos abstractos especializados (8%), como se ve en la Figura 24.

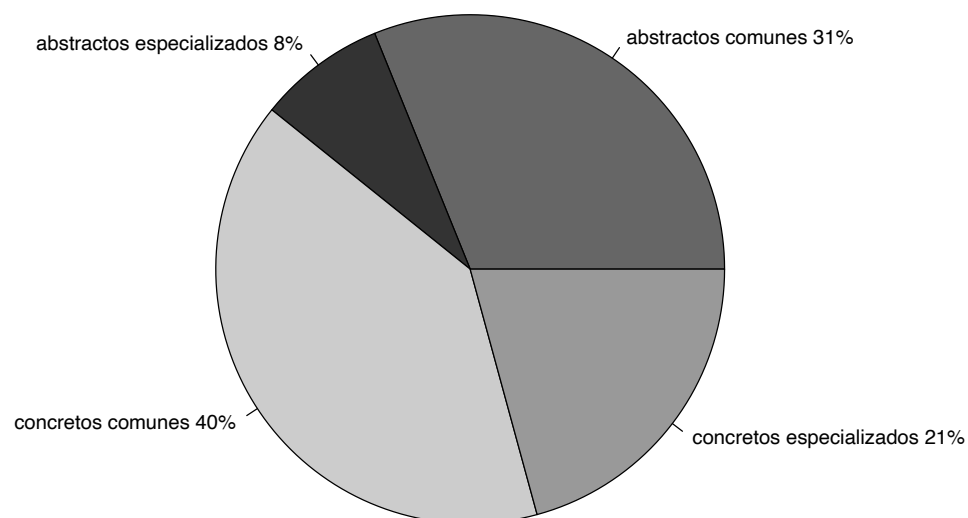


Figura 24. Proporciones de los cruces de categorías de sustantivos en la muestra.

Las frecuencias absolutas de cada cruce de las categorías de sustantivos en la muestra se presentan en la Tabla 14.

Tabla 14. Frecuencia de uso de los cruces de categorías de sustantivos en la muestra.

Categorías	Total de instancias
Concretos comunes	4469
Concretos especializados	2319
Abstractos comunes	3477
Abstractos especializados	905
Total	11,170

En los textos agrupados por promedio, la frecuencia de todos los cruces de categorías de sustantivos resultó más alta en los textos de alto promedio y más baja en los de promedio menor, con distintas proporciones (ver Figura 25).

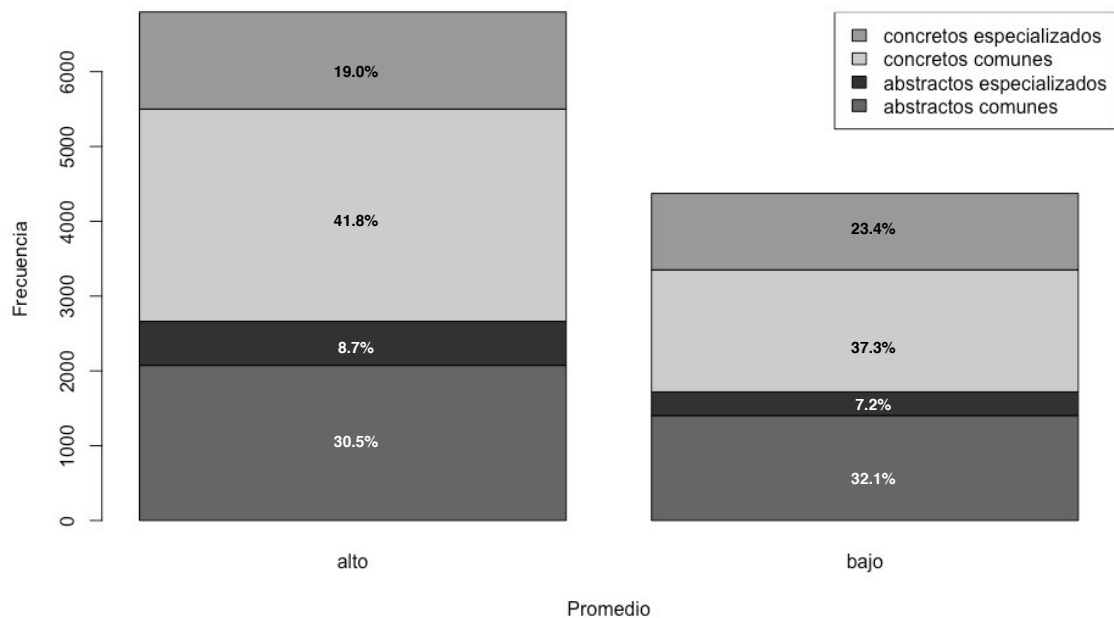


Figura 25. Proporciones de los cruces de categorías de sustantivos en la muestra, según el nivel de PromEA.

Los números absolutos de la Tabla 15 revelan que se usaron de manera diferente las distintas categorías de sustantivos, con amplios márgenes entre los grupos de textos de alto y bajo promedio (PromEA), exceptuando una de ellas, la de concretos especializados.

Tabla 15. Frecuencia de uso de los cruces de categorías de sustantivos según el nivel de PromEA.

Categorías	PromEA	
	Alto	Bajo
Concretos comunes	2837	1632
Concretos especializados	1298	1021
Abstractos comunes	2074	1403
Abstractos especializados	589	316
Totales	6798	4372

Para corroborar si estas diferencias en las frecuencias de las categorías de sustantivos en los dos grupos de textos es significativa en términos estadísticos, se hizo una prueba de independencia Chi cuadrado, con la hipótesis de que las variables Promedio y Categoría no serían independientes una de la otra, lo cual se pudo verificar positivamente ($\chi^2=45.067$, 3 grados de libertad, $p<0.0001$). Esta relación entre las variables sugiere que la ubicación del texto entre los mejor o peor evaluados depende de cuánto se use cada categoría de sustantivos: en el conjunto de textos con promedio alto las dos categorías de sustantivos comunes son las más frecuentes (los concretos suman 2,837 y le siguen los abstractos con 2,074). Esta relación entre la frecuencia de uso de las categorías de sustantivos y el nivel de promedio obtenido por los textos es explorada con mayor detalle en la siguiente sección, pues no se puede descartar la influencia sobre la calificación de las categorías de sustantivos que son empleadas con menos frecuencia, es decir, las dos categorías de sustantivos especializados, tanto concretos como abstractos.

6.4.2.Efectos sobre el nivel de PromEA

Un examen discriminante mediante la prueba t para grupos independientes aplicado a cada cruce de las categorías de sustantivos permitió verificar si la ubicación del texto entre los mejor o peor evaluados depende de cuánto se use cada categoría de sustantivos se cumple caso por caso. Enseguida se presentan los resultados.

En primer lugar, se analizaron los sustantivos concretos comunes, cuyas medias de uso en cada grupo son diferentes (alto=94.56, bajo=54.40), con alta significancia estadística ($t=4.73$, 58 grados de libertad, $p<0.0001$).

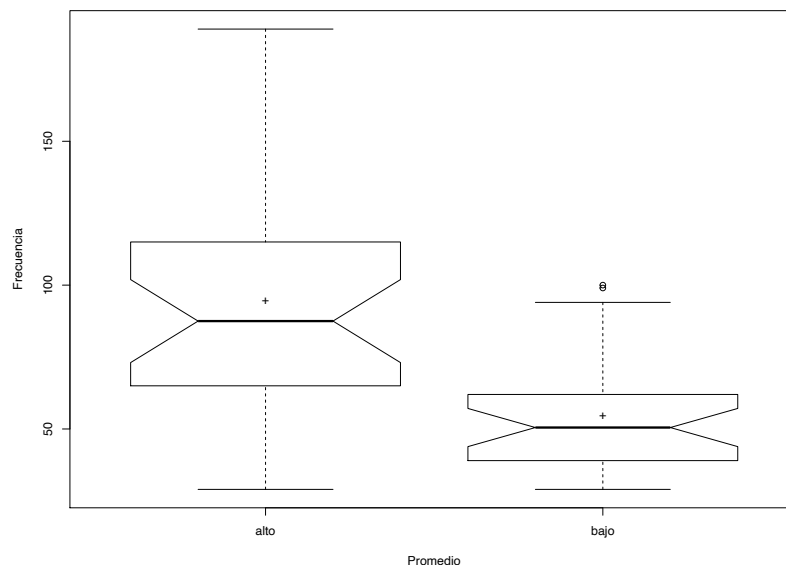


Figura 26. Uso de los sustantivos concretos comunes según el nivel de PromEA.

La Figura 26 muestra cómo los intervalos de confianza alrededor de la mediana en ambos grupos no se solapa, lo cual es indicativo de una diferencia estadísticamente significativa en la frecuencia de esta categoría de sustantivos. Este nuevo dato se apoya en las observaciones acerca del efecto de los sustantivos en general sobre la calificación, fruto del análisis correlacional que se presentó en el apartado 6.1.2 de este trabajo.

En segundo lugar, al calcular las medias de uso de los sustantivos concretos especializados en cada grupo, estas resultaron cercanas (alto=43.26, bajo=34.03), aunque la prueba t mostró que entre ellas sí hay diferencia estadísticamente significativa ($t=2.28$, 58 grados de libertad, $p<0.05$). Este hallazgo sugiere que a pesar de que los términos especializados con referente concreto fueron empleados con relativa frecuencia en ambos grupos de textos, sus rangos de uso son diferentes (ver Figura 27), por lo que esta categoría de sustantivos sí marca una diferencia en la forma en que el texto es evaluado, es decir, a mayor uso de sustantivos concretos especializados, mejor puntuación.

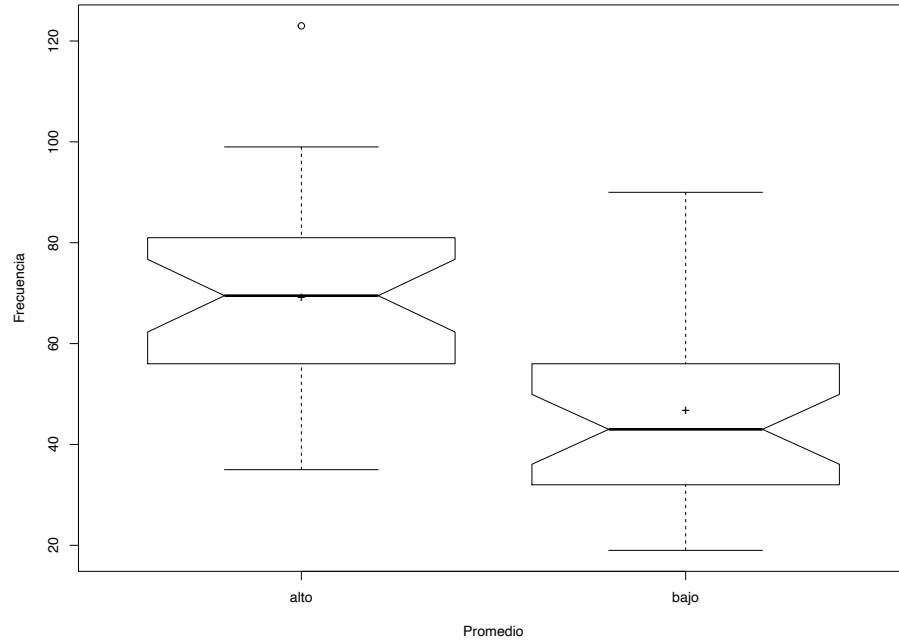


Figura 27. Uso de los sustantivos concretos especializados según el nivel de PromEA.

En tercer lugar, se comparó el uso de los sustantivos abstractos comunes en los dos grupos de textos. La diferencia en las medias de esta categoría en cada uno (alto=69.13, bajo=46.76) se verificó positivamente con la prueba t ($t=4.40$, 58 grados de libertad, $p<0.0001$). La Figura 28 muestra cómo los intervalos de confianza alrededor de la mediana en ambos grupos quedan en rangos distintos, indicando una diferencia estadísticamente significativa en la frecuencia de esta categoría de sustantivos en relación con el nivel de PromEA. Este nuevo dato se agrega a los hallazgos del apartado 6.2.3 de este trabajo, en que se verificó el efecto de los sustantivos abstractos en general sobre la calificación obtenida por los textos, antes de realizar el cruce de categorías.

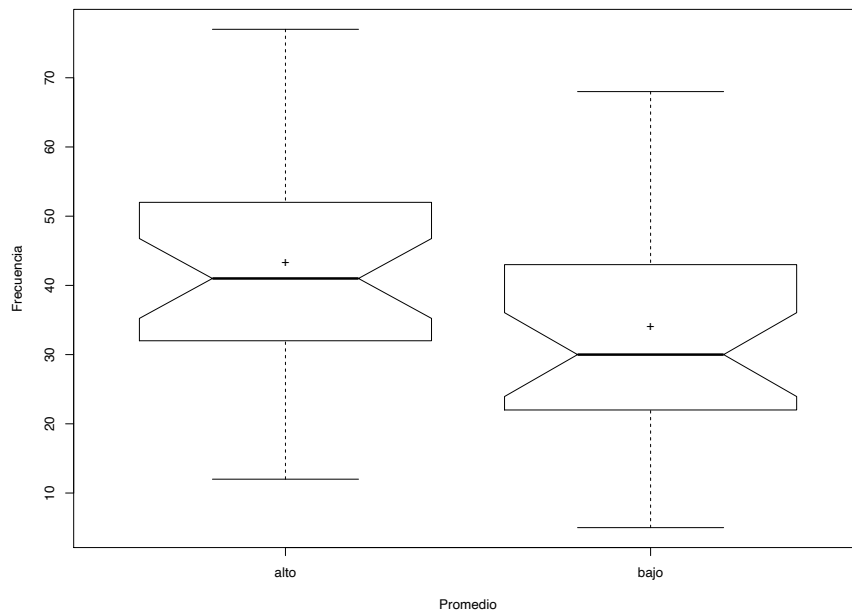


Figura 28. Uso de los sustantivos abstractos comunes según el nivel de PromEA.

Por último, en relación con la variable de mayor interés para esta etapa de la investigación, se calcularon las medias del cruce de sustantivos abstractos especializados en los dos grupos de textos (alto=19.63, bajo=10.53) y se corrió una prueba t para grupos independientes para determinar si la doble característica de estos sustantivos hace alguna diferencia en la evaluación y, por lo tanto, pueden ser considerados entre los predictores de dominio de la escritura académica. Esta prueba estimó que las medias de sustantivos abstractos especializados en los textos agrupados por su promedio son estadísticamente diferentes ($t=4.94$, 58 grados de libertad, $p<0.0001$), por lo que puede decirse que un mayor uso de esta categoría de sustantivos es bien puntuado por los jueces. La Figura 29 muestra cómo los intervalos de confianza claramente se ubican en diferentes rangos.

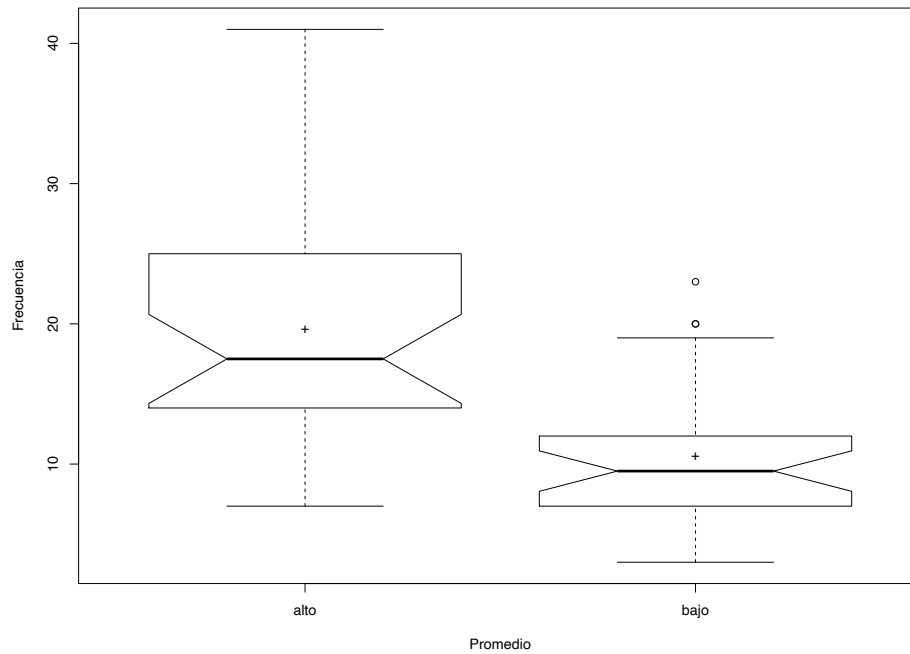


Figura 29. Uso de los sustantivos abstractos especializados según el nivel de PromEA.

En cuanto al orden en que estas variables afectan el nivel de PromEA de los textos, la Tabla 16 muestra en resumen todos resultados de la prueba, jerarquizados por el valor de probabilidad, aunque también se observó el tamaño del efecto.

Tabla 16. *Estimación de los efectos del léxico abstracto especializado sobre el nivel de PromEA.*

Variable (frecuencia)	t	$p <$
Sustantivos concretos comunes	4.73	0.0001
Sustantivos abstractos comunes	4.40	0.0001
Sustantivos abstractos especializados	4.94	0.001
Sustantivos concretos especializados	2.28	0.05

Nota: Pruebas t de Student para grupos independientes con 58 grados de libertad, $\alpha=0.05$.

Los resultados de esta última sección del trabajo subrayan, primero, la importancia de los sustantivos en general en la evaluación que hacen los jueces de los textos. Luego, estos valores también destacan la importancia de los sustantivos abstractos específicamente, en dicha evaluación, de acuerdo con lo que se había demostrado para este corpus en el apartado 6.2. Enseguida, tal como se preveía por los resultados del apartado 6.3, se pudo confirmar que la frecuencia con que se usen los sustantivos abstractos especializados tiene gran influencia sobre la variación en la calificación; mientras que la frecuencia de los sustantivos concretos especializados tiene una menor repercusión sobre la varianza.

6.4.3. Ejemplos del léxico abstracto especializado más frecuente en la muestra

Con la intención de ejemplificar el uso del léxico abstracto especializado, se extrajeron los casos más frecuentes en la muestra analizada. Los ejemplos fueron descritos en función de su campo (*Lengua y literatura* u *Otras disciplinas humanísticas*) y su procedencia probable (la clase en el aula o el libro de texto), para ilustrar el postulado (verificado por otros autores, como Chung & Nation, 2004; Coxhead, 2017, o Torner & Battaner, 2002) acerca de la estrecha relación de este tipo de léxico con las disciplinas de la ciencia (y, en este caso, las materias de la escolaridad y el discurso pedagógico).

Ordenados por frecuencia, los primeros 20 sustantivos abstractos especializados resultan ser *efecto*, *contexto*, *estilo*, *interpretación*, *influencia*, *existencialismo*, *género*, *intertextualidad*, *existencia*, *corriente*, *individuo*, *movimiento*, *descripción*, *significado*, *diálogo*, *perspectiva*, *recepción*, *superhombre*, *lectura* y *moral*.

La Tabla 17 los presenta en ese orden, junto con su frecuencia de uso, campo y procedencia.

Tabla 17. *Posición, campo y procedencia de los sustantivos abstractos especializados más frecuentes en la muestra.*

#	Sustantivo	f	Campo	Procedencia
1	efecto	266	LyL	Libro de texto
2	contexto	178	LyL	Libro de texto
3	estilo	152	LyL	Libro de texto
4	interpretación	107	LyL	Libro de texto
5	influencia	78	LyL	Libro de texto
6	existencialismo	76	Otros	Aula
7	género	69	LyL	Libro de texto
8	intertextualidad	56	LyL	Libro de texto
9	existencia	54	Otros	Aula
10	corriente	51	LyL	Aula
11	individuo	40	Otros	Aula
12	movimiento	40	LyL	Aula
13	descripción	33	LyL	Aula
14	significado	30	LyL	Aula
15	diálogo	24	LyL	Aula
16	perspectiva	24	LyL	Aula
17	recepción	24	LyL	Libro de texto
18	superhombre	24	Otros	Aula
19	lectura	20	LyL	Aula
20	moral	20	Otros	Aula

Una descripción más detallada de las observaciones de la Tabla 17 muestra que 8 términos del análisis comparativo literario se ubicaron entre los sustantivos abstractos especializados más frecuentes: *efecto*, *contexto*, *estilo*, *interpretación*, *influencia*, *género*, *intertextualidad* y *recepción*. El abundante uso de los primeros términos en esa lista evidencia los enfoques de comparación preferidos por los estudiantes. Enseguida se ofrecen ejemplos de tres de ellos.

13) “Se da a entender que el *efecto* es lo que la obra produce en el lector, podría ser un sentimiento o simplemente una sensación. El *efecto* que causa El viejo y el mar es algo de compasión y respeto por el viejo, ya que es una persona tan dedicada” (14H04).

14) “Si se analiza el *contexto* se puede encontrar que Joyce escribió Retrato de un artista adolescente para poder probar la teoría de William James del Flujo de conciencia” (14H30).

15) “En el texto El viejo y el mar se puede apreciar el *estilo* de Ernest Hemingway, el cual detallaba de una manera ruda y realista los acontecimientos” (15H07).

El fragmento citado en (13) muestra que el término referido a las sensaciones que genera una obra en su lector (*efecto*) fue con mucho el enfoque preferido de los estudiantes, probablemente porque les permite expresar lo que, en palabras coloquiales, les gusta o les disgusta de las obras, pero usando para ello un lenguaje disciplinar.

El ejemplo de (14) muestra que el enfoque basado en el análisis del *contexto* en que se producen las obras literarias (social, cultural, geopolítico) se ubica en segundo lugar de preferencia entre los términos del análisis comparativo. Esto se debe, muy probablemente, a que ofrece a los estudiantes la posibilidad de documentarse en fuentes externas, una habilidad a la que acuden por su familiaridad, dado que la consigna dirigida a ellos de hacerse con información externa a la clase o a los textos que se leen en el aula está entre las más socorridas por los profesores de enseñanza básica y media. Mayor interés reviste el hecho de que en el tercer puesto, y con alta frecuencia, se encuentre representado un análisis más complejo, el del *estilo*, pues implica que los estudiantes recurran a lo aprendido en clase para localizar y describir alguna característica que distinga el lenguaje del escritor, que puede ser una figura retórica, su sintaxis, sus personajes, el tono general de la obra o las cualidades de la narración, como en el ejemplo (15), que alude al realismo

descriptivo de la novela. Entre los primeros 10 sustantivos de la lista se ubicaron también dos conceptos clave para el tratamiento de un tópico común a muchos de los textos: se trata de uno de los conceptos filosóficos de mayor abstracción, *existencia* (16), y un término derivado de él, en el sentido de movimiento o corriente: *existencialismo* (17).⁴¹ Los siguientes son ejemplos de su uso en la muestra:

16) “Por otra parte Camus expone la vida cruda del ser humano, la *existencia* misma, sin preocupaciones, ni miedos” (14H16).

17) “...el libro hace referencia a la ‘condición humana’ a la que este personaje está sometido. De igual manera muestra el *existencialismo* absurdo en el que vive el protagonista” (14M09).

A partir del puesto 10 en frecuencia se ubican sustantivos abstractos altamente especializados: *corriente* y *movimiento* son constructos de la sociología de la literatura que también sirven a la vertiente histórica de esta disciplina; *descripción*, *diálogo*, *perspectiva* y *lectura* son términos narratológicos, mientras que *significado* es un constructo empleado en varias disciplinas, como la filosofía y la lingüística, y adoptado por la semiótica literaria.

En seguida se ejemplifican *diálogo* (18) y *perspectiva* (19), términos del análisis literario que refieren, el primero, a un acto verbal que a la vez categoriza una estrategia de discurso, y el segundo a un constructo conceptual empleado en la teoría y la crítica de varias disciplinas artísticas, pero adoptado también por la narratología.

18) “Es posible notar que el narrador es omnisciente, aunque el texto está interrumpido por *diálogos* y narraciones del protagonista” (14M01).

19) “El escritor Albert Camus, que participó en el drástico cambio de la narrativa contemporánea, representaba un pensamiento influenciado por Schopenhauer,

⁴¹ Una de las novelas leídas en el contexto de la clase, y que fue objeto de análisis en muchos de los textos que forman el corpus, fue *El extranjero*, del filósofo y literato existencialista francés Albert Camus.

Nietzsche y el existencialismo alemán, caracterizados por una *perspectiva* existencialista y desde un pesimismo extremo” (15M08).

Finalmente, en esta lista aparecen otros conceptos interesantes: *individuo* (20), *superhombre* (21) y *moral* (22). Los tres son constructos de la filosofía que permean el discurso literario, y especialmente ciertos tópicos de literatura existencialista y posmoderna provenientes de la clase.

20) “Camus está representando al hombre en el momento en el que este está ‘siendo creado’ nuevamente, es un ‘reclamo’ hacia la sociedad que olvida al *individuo* y lo priva de sentimientos de pertenencia ante la comunidad” (14M09).

21) “Según Nietzsche, el *superhombre* es aquel que es capaz de generar su propio sistema de valores identificando como bueno, todo aquello que provenga de su genuina voluntad” (14M05).

22) “mientras parece que el haber agotado fuerzas por el objetivo deseado es válido en *El viejo y el mar* de Ernest Hemingway, en *el extranjero* de Albert Camus conlleva a preguntarse, a dudar de su propia *moral* y ver si el objetivo que perseguía hace tiempo tiene interés o relevancia ahora que nada tiene sentido” (15H07).

El caso de *moral* merece atención porque, aunque pertenece igualmente a la filosofía, específicamente a la ética, es más difícil establecer su procedencia, pues no necesariamente aparece en el libro de texto o en las obras comparadas por los estudiantes. En resumen, este tipo de sustantivos de referente disciplinar identificable en los campos de la lengua y la literatura, la filosofía y otras materias humanísticas pudo ser adoptado por los estudiantes por contacto con tres fuentes principales: el libro de texto, las obras literarias que leyeron y compararon, y la clase, aunque al menos uno de ellos (*moral*) tiene una procedencia menos clara.

7. DISCUSIÓN DE LOS RESULTADOS

Para dar inicio a la discusión de los resultados del presente trabajo de tesis, es conveniente reiterar sus ejes vertebradores: el objetivo general, la hipótesis, las acciones que se llevaron a cabo para verificarla y el marco para interpretar los resultados.

El objetivo general fue determinar qué características del léxico empleado por estudiantes en la escritura de textos pueden ser indicadores efectivos del nivel de dominio de la escritura académica, y establecer una jerarquía de su importancia que oriente futuros tratamientos didácticos y de evaluación.

Como se recordará, se analizaron 120 textos producidos por jóvenes estudiantes de bachillerato en una clase de literatura, los cuales fueron evaluados con base en una rúbrica por un conjunto de jueces expertos para obtener un índice del constructo *dominio de la escritura académica*, al que se denominó PromEA. Además, sobre esos textos se hicieron mediciones de extensión, diversidad léxica, diversidad nominal, densidad léxica, densidad nominal, de frecuencia de uso de sustantivos abstractos, de léxico especializado y de léxico abstracto especializado, y estos datos se correlacionaron estadísticamente con el índice PromEA para calcular el efecto de esas mediciones con la variación en la calificación.

La hipótesis de trabajo fue que los textos mejor evaluados serían aquellos que presentaran un léxico más variado (más palabras diferentes), más denso léxicamente (más palabras de contenido semántico), más denso nominalmente (alta proporción de sustantivos), y un mayor empleo de sustantivos abstractos, de léxico especializado y de léxico abstracto especializado, a diferencia de los textos con evaluaciones bajas.

Este razonamiento se basó dos supuestos: a) entre las consignas de la pauta de escritura que usaron los estudiantes para producir sus textos, se les solicitaba usar un

lenguaje académico variado y preciso, evitando repeticiones y ambigüedades; en razón de ello se podía suponer, con mucha probabilidad, que los estudiantes se habrían visto motivados a emplear los recursos lingüísticos a su alcance para adecuar su discurso a las exigencias del lenguaje académico y cumplir los requisitos de la tarea; b) las descripciones del discurso académico y del léxico en la escritura académica asentadas en la literatura previa permitían suponer que entre esos recursos se encontrarían los siguientes: variedad en el uso del léxico, concentración informativa a través del contenido léxico y nominal, uso de sustantivos abstractos para construir conceptos, nombrar procesos y dar valor de término disciplinar a ciertas palabras, que podían ser de referente concreto o abstracto.

Todos estos conocimientos de partida se expusieron en el capítulo dedicado al marco teórico, por lo que en este espacio solo se sintetizan para contrastarlos con los hallazgos de la presente investigación: el discurso escrito académico se caracteriza por la diversidad y la densidad léxica (Chafe & Danielewicz, 1987; Snow & Uccelli, 2009); las elecciones léxicas precisas, el uso de términos técnicos y conceptos abstractos (Coxhead, 2017; Halliday & Martin, 1993; Snow & Uccelli, 2009; Zwiers, 2008); las entidades y los participantes del discurso explícitos por medio de nombres y frases nominales (Schleppegrell, 2001); las nominalizaciones (Bello, 2016; Halliday & Martin, 1993; García, Hall & Marín, 2005), las expresiones formulaicas y el metalenguaje con funciones encapsuladoras y cohesivas (Fang, Schleppegrell & Cox, 2006; Jiang & Hyland, 2015; Hyland & Tse, 2004). Por lo tanto, encontrar variaciones en el empleo de estos recursos entre los textos mejor y peor evaluados daría sustento a su potencialidad como indicadores de un cierto nivel dominio de la escritura académica en este tipo particular de textos (textos académicos de formación ligados a una materia escolar, con fines de evaluación) y en este nivel educativo (bachillerato).

Enseguida se discutirán los resultados de cada etapa del análisis a la luz de este marco teórico y de lo verificado por estudios previos, según se vio en el apartado de antecedentes,

para describir y valorar el cumplimiento de los objetivos específicos y la forma en que se contestaron las preguntas de investigación.

Objetivo específico I. Identificar en un corpus de textos producidos por estudiantes de bachillerato cuáles características en el uso del léxico (extensión, diversidad léxica, densidad léxica, diversidad nominal y densidad nominal) tienen potencial para predecir una variación en la calificación otorgada por jueces expertos, y en qué condiciones se observa esa relación.

Este objetivo se logró contestando la *Pregunta específica 1*. ¿Cómo interactúan las mediciones de extensión de los textos, de diversidad y densidad léxica, de diversidad y densidad nominal, y cuáles son sus efectos sobre el promedio de calificación otorgada a los textos por jueces expertos?

El principal hallazgo de esta primera etapa del trabajo es que la realización de la consigna de emplear un lenguaje académico diverso no fue uniforme en todos los textos, y esa variabilidad tuvo su correlato en la evaluación, pues la interacción de algunos pares de indicadores léxicos afectó la calificación en los modelos estadísticos. Esto dio sustento a la idea de que, en efecto, existen configuraciones del léxico que pueden predecir una mejor o peor calificación para los textos académicos, lo cual tiene lógica porque el nivel léxico-gramatical es un componente básico de la unidad discursiva mayor que es el texto.

Los modelos de regresión lineal múltiple calcularon las variaciones de calificación en todos los escenarios hipotéticos de cada medición de las variables del léxico y asentaron las combinaciones con mayor poder para predecir un rango de calificación del texto. Los escenarios limítrofes por exceso o por defecto demostraron tener un efecto negativo en la puntuación de los jueces, y, por otra parte, los coeficientes de tamaño del efecto de esos modelos demostraron con muy buena probabilidad que son las mediciones de densidad –no las de variedad– las que más pesan sobre la calificación.

Dicho de otro modo, un texto con muchas palabras diferentes no necesariamente es percibido por los jueces como un texto académico bien logrado, por más que los profesores de lengua o de cualquier otra asignatura prioricen ese aspecto (tan visible en el discurso como la ortografía) en sus pautas de escritura, y tampoco puede considerarse un rasgo con valor académico de manera aislada al resto de los componentes de la escritura (Atorresi, 2005; Martínez, 2010; Morales, 2004; Navarro, Ávila & Gómez, 2019).

En cuanto a la necesidad de establecer jerarquías, señalada como parte de la propuesta de este trabajo de tesis, se puede decir que, en este corpus, la *densidad nominal* (la proporción de sustantivos en el texto y por cláusula) y la *densidad léxica* (la proporción de palabras de contenido en el total de palabras del texto) fueron mejores indicadores de variación en las calificaciones que las medidas de diversidad, lo cual puede estar relacionado con las expectativas de calidad del contenido temático de la selección de palabras y de sofisticación en el lenguaje empleado en los textos por encima de la evitación mecánica de repeticiones, como lo habían constatado algunos estudios previos (McNamara, Crossley & McCarthy, 2009; Stömqvist, Johansson, Kriz, Ragnarsdóttir, Aisenman & Ravid, 2002). Aunque es necesario aclarar que el papel de la diversidad léxica en la valoración de los textos no debe descartarse, por sus estrechos vínculos con la superposición de palabras y la cohesión (Crossley & McNamara, 2011; Uccelli, Dobbs & Scott, 2012).

Esta jerarquía es relevante porque actualiza los hallazgos de algunos trabajos previos como los de Ure (1971), Chafe y Danielewicz (1987) y Johansson (2008), que establecieron la densidad léxica como un medio válido para medir el grado de ‘oralidad’ frente a la ‘literacidad’ en un texto. Como quedó establecido en el apartado 4.1.1, una de las expectativas que los miembros de una comunidad académica tienen hacia los estudiantes es que se adapten al registro formal del discurso académico y aprendan a diferenciarlo de las interacciones coloquiales (Schleppegrell, 2004; Uccelli, Barr, Dobbs, Phillips Galloway,

Meneses & Sánchez, 2014). Además, por estudios previos se sabe que una mayor proporción de palabras de contenido, especialmente sustantivos y adjetivos, puede ser indicativa de una mayor complejidad al nivel de la cláusula, la cual es también un rasgo diferenciador del discurso académico (Scheleppegrell, 2001; Uccelli, Dobbs & Scott, 2012); también puede reflejar la adquisición de algunas habilidades cruciales para participar en prácticas letradas en el ámbito académico, como empaquetar y desempaquetar información densa y cohesionar el texto (Uccelli, Phillips Galloway, Barr, Meneses y Dobbs, 2015), aunque el enfoque dirigido al nivel léxico y el carácter cuantitativo del presente estudio no hayan permitido corroborar estas implicaciones.

Los hallazgos de esta primera etapa del trabajo dirigieron el interés de esta investigación hacia las características y el comportamiento de los sustantivos en el texto académico, pues estas unidades léxicas mostraron ser un aspecto muy relevante en la valoración de la escritura académica por parte de los expertos.

Objetivo específico II. Determinar si las diferencias en el uso de sustantivos abstractos, de términos especializados y de sustantivos abstractos especializados en los textos tienen potencial para discriminar el nivel de calificación otorgada a los textos por los jueces.

Este objetivo se logró contestando tres preguntas de investigación, que enseguida se contestan una a una. *Pregunta específica 2.* ¿En qué proporciones se encuentran los sustantivos abstractos y sus clases en el corpus y qué efecto tiene el uso de este léxico abstracto en los textos sobre la variación en la calificación de una muestra agrupada por su nivel de dominio de la escritura académica?

En esta segunda etapa de la investigación se constató que cuatro de cada diez sustantivos en el corpus son abstractos, y esta proporción se repite en la muestra de textos. Pero en la frecuencia de uso se encontraron diferencias entre los textos mejor evaluados y

los de calificaciones más bajas; asimismo, hubo un uso diferenciado de las cuatro clases de sustantivos abstractos en los dos grupos de textos. Las pruebas estadísticas verificaron la influencia de este tipo de léxico en la variación en la calificación, lo cual permite, por una parte, incluir los sustantivos abstractos entre los indicadores de dominio de la escritura académica, y, por otra parte, demuestra que su aprendizaje debe ser modelado y sostenido didácticamente, pues no todos los estudiantes pudieron incorporarlos a sus textos de la misma manera, ni en la misma jerarquía.

Estos hallazgos confirman la valoración positiva que tiene el manejo de léxico abstracto en el ámbito académico (Bello, 2016), sin perder de vista que también se le atribuye una mayor complejidad conceptual (García, Hall & Marín, 2005).

La opción de emplear recursos como la nominalización para empaquetar conceptos en la escritura escolar revela, por parte de los estudiantes, una intención de adscribir su texto a las convenciones del discurso académico, diferenciándolo de los usos lingüísticos de las interacciones cotidianas que son más familiares para ellos (Fang, Schleppegrell & Cox, 2006; Schleppegrell, 2004). La incorporación de léxico abstracto en el discurso propio puede ser considerado como uno de los accesos a la cultura y las prácticas académicas que faciliten su ingreso a estudios superiores, pues se espera que el discurso universitario, como se dijo en el marco teórico, aumente el nivel de especialización disciplinar en los contenidos curriculares, lo que puede aparejar un mayor grado de abstracción en las demandas de escritura.

Por lo que toca a la posibilidad de inferir una jerarquía en el tratamiento didáctico de los sustantivos abstractos, se puede decir, con base en los resultados de las pruebas discriminantes, que las nominalizaciones de procesos, acciones y resultados, y todas las abstracciones que caben en la clase ‘Eventos del mundo físico y estados de cosas’ son el tipo de sustantivos que mayor efecto tienen sobre la calificación, lo cual confirma dos

supuestos previos sobre el discurso al que los estudiantes se acercan: uno, la importancia del recurso gramatical de la nominalización para empaquetar información densa en una palabra (Halliday & Martin, 1993), y dos, el alto grado de generalización y abstracción semántica de este tipo de discurso, que permite clasificar, etiquetar y hacer taxonomías de los fenómenos del mundo social y natural (Zwiers, 2008).

Le siguen los ‘Constructos teóricos y conceptos’, cuyo empleo en los textos también permite predecir con alto efecto variaciones en la calificación. Este tipo de elecciones léxicas son bien valoradas en el ámbito académico porque sirven para reducir la ambigüedad, confieren formalidad y precisión disciplinar al texto, a diferencia del léxico más concreto y del sentido común que caracteriza las interacciones informales y coloquiales (Monippally & Shankar, 2010; Snow & Uccelli, 2009).

Tanto los ‘Constructos teóricos y conceptos’ y los ‘encapsuladores’, no obstante su menor frecuencia en el corpus, son dos clases de sustantivos abstractos cuyas funciones en el discurso académico merecen un análisis de mayor profundidad y de carácter cualitativo, pues no es su abundancia lo que se espera en este tipo de textos, sino la precisión de su empleo en la construcción del discurso (López, A., 2013, 2018). Además, se sabe que el uso de estos tipos de sustantivos abstractos trae consigo una serie de habilidades cognitivas de alta complejidad, como la selección y activación de información en el texto, el empaquetamiento temporal de conceptos, la reificación, la puesta en perspectiva de entidades, etcétera (Schmid, 1999a), por lo que el estudio de su impacto en el aprendizaje resulta relevante.

Asimismo, los usos de la clase de sustantivos abstractos ‘Procesos del mundo mental y estados afectivos’ entre los textos mejor y peor evaluados tuvieron poco efecto como diferenciadores. Probablemente la baja precisión en el efecto se deba a la decisión metodológica de introducir en una misma clase los sustantivos que aluden a actos mentales

como pensar, inteligir, analizar e interpretar (pensamiento, intelección, análisis, interpretación) y los que refieren a emociones personales (alegría, tristeza, esperanza, lástima). Habría que hacer un análisis por separado del efecto de este tipo de sustantivos abstractos, pues probablemente los sustantivos usados para expresar sentimientos y emociones no sean el tipo de léxico que se evalúe positivamente en un texto académico de las características de los textos analizados en esta investigación.

Como se aclaró antes, el segundo objetivo de esta tesis comprendió tres preguntas. A continuación se atiende la *Pregunta específica 3*. ¿Qué proporción de los sustantivos en el corpus son términos especializados de algún campo disciplinar, y qué efecto tiene el uso de este léxico especializado en los textos sobre la variación en la calificación de una muestra agrupada por su nivel de dominio de la escritura académica?

En los textos producidos por los estudiantes de bachillerato en su clase de literatura se pudo identificar un repertorio de términos especializados, y estos términos se pudieron distribuir en dos grandes campos disciplinares (*Lengua y literatura* y *Otras disciplinas humanísticas*), lo cual concuerda con el carácter altamente específico de este tipo de lenguaje, que depende estrechamente de la materia, el campo o la rama de actividad científica a la que refiere (Chung & Nation, 2004; Coxhead, 2017; Martín, 2004). En concordancia con lo encontrado por Chung y Nation (2004), respecto de la proporción de términos técnicos en el total de palabras en corpus disciplinares (20% a 30%), el léxico especializado analizado en el corpus de esta tesis comprende casi un tercio de los sustantivos, lo cual es indicativo de la medida en que el léxico disciplinar se extiende en el lenguaje empleado por los estudiantes.

Con base en la teoría revisada para esta investigación, se puede decir que la adopción de léxico especializado en los textos académicos es una importante habilidad discursiva con efectos en el éxito académico (Halliday & Martin, 1993; Nippold & Scott,

2010; Schleppegrell, 2001, 2004; Snow & Uccelli, 2009; Uccelli, Dobbs & Scott, 2012), lo cual se pudo constatar al relacionar el uso de términos especializados con la calificación otorgada a los textos por los jueces. El análisis discriminante entre los textos mejor y peor evaluados reveló que, en efecto, la percepción de su calidad se vio influida por la presencia de terminología especializada, por la pertenencia de este léxico al campo disciplinar que fue tema u objeto de la escritura, y por la frecuencia con que se empleó este tipo de léxico.

Ahora bien, comparando estos resultados con los del análisis anterior, los sustantivos abstractos constituyen aproximadamente el 40% de los sustantivos, tanto en el corpus como en la muestra de textos agrupados por promedio, y la terminología especializada el 30%. Esta diferencia en el uso influye en su peso sobre la calificación, pues el empleo de sustantivos abstractos demostró tener un mayor poder predictivo de la variación en el promedio que el uso de léxico especializado (ver Tablas 11 y 13).

En la teoría presentada como marco de esta investigación se asume que el léxico especializado es más accesible para los aprendices de escritura académica debido a que este tipo de palabras tienen referentes más precisos que los conceptos abstractos (Bello, 2016; Halliday & Martin, 1993; Navarro & Revel, 2013); sobre esa base, se podría hipotetizar que la capacidad de los estudiantes para remontar esta complejidad estaría siendo percibida por los jueces. No obstante, el carácter cuantitativo de esta investigación no permite ir más allá en la interpretación de los valores de estos dos indicadores por separado (abstracción y especialización) como predictores del nivel de dominio de la escritura académica; habría que controlar variables como la extensión de los textos y el entrenamiento de los jueces en el uso de la rúbrica, para poder precisar el origen y el valor de estas diferencias; además de analizar la precisión en el uso real de ambos tipos de léxico, lo cual tendría que considerarse en abordajes futuros de corte cualitativo.

También es necesario precisar que, por supuesto, la adopción de un vocabulario especializado no es la única habilidad académica favorecida por la escolarización, como queda claro gracias a investigaciones recientes de largo alcance en español y en inglés y en varios niveles escolares (Figueroa, 2018; Figueroa et al., 2018; Meneses et al., 2017); pero sí abre una ventana a la exploración de posibilidades didácticas que la promuevan como parte no desdeñable de la literacidad académica.

Esta etapa de la investigación dejó ver que la tarea de escritura de un texto académico motivó, por parte de los estudiantes, la incorporación de un repertorio léxico disciplinar claramente relacionado con categorías del análisis literario y del lenguaje, pero también con otros campos de las ciencias sociales que les ayudaron a desarrollar su discurso. Es notable que aunque la rúbrica empleada por los estudiantes para escribir los textos en su clase no hacía referencia explícita a la obligación de atraer terminología específicamente literaria ni las fuentes de donde debía tomarse ese vocabulario, los jóvenes escritores adoptaron un repertorio terminológico disciplinar básico, aunque con diferencias en su uso. Estas diferencias, y su impacto en la percepción de la calidad de los textos por parte de los jueces, subrayan, una vez más, la necesidad de modelar y acompañar didácticamente el aprendizaje y el uso pertinente del léxico disciplinar en la escritura académica, pues es poco probable que un lenguaje de estas características se adquiriera sin la intermediación del docente en el ámbito escolar (Coxhead, 2017).

La consecución del segundo objetivo de esta tesis implicaba dar respuesta también a la *Pregunta específica 4*. ¿Qué proporción de los sustantivos abstractos en la muestra son a la vez términos especializados, y qué efecto tiene el uso de este léxico abstracto especializado en los textos sobre la variación en la calificación otorgada por jueces expertos?

Dado que en las etapas anteriores de la investigación se había podido constatar que el uso frecuente de los sustantivos abstractos y sus clases, así como de léxico especializado, tiene un impacto positivo sobre la calificación otorgada a los textos por parte de los jueces, se quiso determinar si la concurrencia de ambas características en los sustantivos ayudaría a precisar su efectividad como indicadores de dominio de la escritura académica.

Así pues, para reducir la incertidumbre acerca de los distintos efectos predictivos de los sustantivos abstractos y de los términos especializados, se hizo un análisis de los sustantivos categorizados por el cruce de estas dos características (concretos comunes, concretos especializados, abstractos comunes y abstractos especializados), en la muestra de textos agrupados por su nivel de dominio.

Los resultados permiten inferir que el empleo de sustantivos abstractos especializados es un buen indicador de dominio, pues predice con alta precisión la variación en la calificación otorgada al texto. El carácter abstracto de este tipo de léxico disciplinar es bien valorado en el texto académico, a diferencia de los sustantivos concretos especializados, que fueron empleados de manera semejante en los textos con distinto promedio, y concomitantemente su poder predictivo resultó más débil (ver Tabla 16).

Luego, el que la frecuencia de uso de los sustantivos concretos comunes marque una diferencia en el nivel alto o bajo del promedio solo viene a respaldar la importancia de la densidad nominal en los textos académicos, cuyo peso en la calificación se verificó en la primera parte de este trabajo (ver Apartado 6.1 de la exposición de resultados).

Por lo que toca a la frecuencia de uso de los sustantivos abstractos comunes, su efecto diferenciador del nivel en que se ubican los textos refuerza lo sabido acerca del carácter eminentemente abstracto del discurso académico, mismo que ha sido documentado ampliamente por otros autores (Bello, 2016; García, Hall & Marín, 2005; Schleppegrell,

2001; Snow & Uccelli, 2009; Zwiers, 2008). Definitivamente, esto influyó en la valoración de los textos por parte de los jueces.

Los resultados de esta última etapa de la investigación justifican la necesidad de analizar el contexto discursivo en que aparecen los sustantivos en función de término para determinar si el valor de término se lo confiere el uso y además tiene una acepción no disciplinar, lo cual puede sostener el supuesto de que los estudiantes los están empleando porque discriminan correctamente los contextos de uso y se adecuan a ellos. Sin embargo, lo que el texto pone en evidencia es “el vocabulario conocido y usado” (Torner & Battaner, 2002) por los estudiantes, por lo que toda inferencia al respecto ha de acotarse a la situación que actualiza el texto producido. Asimismo, metodológicamente, se confirma que el carácter disciplinar de una palabra es una función que esta realiza en ciertos contextos pero que puede cambiar en otros, por lo cual es útil emplear al menos dos métodos de identificación (Chung & Nation, 2004), como se hizo en este trabajo de tesis (ver Apartados 5.3.2.7 y 5.3.2.8).

Acercas del impacto de esta clase de léxico en la variación de la calificación se puede inferir que el carácter abstracto o concreto de los términos disciplinares sí es influyente en la forma en que los jueces evalúan este tipo de textos. Queda claro que los estudiantes no pueden utilizar solo léxico abstracto ni especializado en la construcción de sus textos; también manejan palabras comunes que, asociadas local y situacionalmente a un uso disciplinar, cobran un valor de término probablemente transitorio. Sería interesante explorar cualitativamente los sustantivos concretos en función de términos para comprender mejor el alcance de esta incorporación en el discurso de los estudiantes, y saber si pueden diferenciar sus acepciones en los usos vernáculos de la lengua y en el discurso académico.

En síntesis, en esta última etapa de la investigación se pudo determinar que los términos especializados concretos –al menos en este corpus– influyen menos en la forma en

que el texto es evaluado, a diferencia de los sustantivos abstractos especializados –o sea, los términos especializados que cumplen las condiciones de abstracción, sea por la ontología de su referente o porque sustancializan un proceso o evento mediante el recurso de la nominalización–, que, aunque no son muy frecuentes (ver Tabla 15), sí son bien valorados en la escritura académica. Nuevamente, solo se puede especular si la mayor complejidad de la abstracción –que Halliday y Martin (1993) relacionan principalmente con el proceso de la nominalización– cobra un valor particular a criterio de los jueces. Ampliar el análisis cualitativo de este tipo de léxico quizá podría aportar elementos para conocer los modos en que se adquiere y se usa, pues tal vez la comunidad académica valora de manera diferente el empleo de conceptos dependiendo de la distancia o la cercanía con el campo de especialidad, lo cual subyace a la descripción que hace Zwiers (2008) del lenguaje académico como un sistema de sub-registros disciplinares.

Los hallazgos de esta etapa final de la investigación no solo permiten confirmar la importancia que tienen los sustantivos en el léxico académico como portadores de contenido temático, y de los sustantivos abstractos como encapsuladores de información, sino que además agrega interés a la descripción del estatus semántico de esta clase de palabras y la forma en que ayudan a particularizar el discurso académico.

Finalmente, el esbozo de análisis de los ejemplos de uso mostró que el discurso de los estudiantes incorpora términos especializados y conceptos disciplinares que, aunque hayan sido usados de manera incipiente, sea apoyándose en una definición o con un ejemplo, de cualquier forma tuvieron una recepción positiva en el criterio de los jueces. En todo caso, este cruce en la caracterización de los sustantivos permitió ponerlos en perspectiva: como constituyentes del léxico general y como elementos de un vocabulario especializado.

Objetivo específico III. Establecer, con base en la evidencia, una jerarquía de importancia de los indicadores del léxico en la valoración de los textos académicos para hacer recomendaciones acerca de su tratamiento didáctico como componentes de la escritura académica, acotadas al tipo de textos analizados y al nivel educativo en que estos se produjeron. Este último objetivo se contesta con la *Pregunta específica 5.* ¿En qué orden se deberían jerarquizar las características del léxico para su tratamiento didáctico como componentes de la escritura académica en el bachillerato? Para contestarla, se presenta la Tabla 18, que organiza las variables del estudio seleccionadas por su valor predictivo o diferenciador del nivel de dominio de la escritura académica.

Tabla 18. *Propuesta de jerarquía de los indicadores léxicos de dominio de la escritura académica.*

#	Por su efecto predictivo del nivel de dominio (+/-)	F	R ²	p <	#	Por su efecto discriminante del nivel de dominio (alto/bajo)	t	p <
1	Densidad nominal en el léxico	5.47	0.20	0.05	1	Sustantivos abstractos	5.35	0.001
2	Densidad léxica	4.18	0.15	0.05	2	Sustantivos abstractos especializados	4.49	0.001
3	Densidad nominal en la cláusula	3.97	0.14	0.01	3	Sustantivos abstractos clase 'Eventos del mundo físico y estados de cosas'	4.36	0.001
4	Diversidad nominal	3.97	0.14	0.001	4	Términos especializados	3.80	0.001
5	Diversidad léxica	3.77	0.14	0.05	5	Sustantivos abstractos clase 'Constructos teóricos y conceptos'	3.35	0.01

Notas: Modelos de regresión lineal múltiple ordenados por el tamaño del efecto (F), con 7 y 112 grados de libertad, $\alpha=0.05$. Se presenta el coeficiente R cuadrado ajustado. Pruebas t de Student con 58 grados de libertad, $\alpha=0.05$ ordenadas por tamaño del efecto (t).

Las posiciones de los indicadores léxicos en la Tabla 18 se basan en los coeficientes obtenidos en las pruebas estadísticas: modelos de regresión lineal múltiple, que buscan

efectos predictivos de las variables sobre el índice numérico PromEA, para el caso de los indicadores de uso del léxico; pruebas t para grupos independientes, que señalan la diferencia en las medidas de tendencia central de los textos agrupados por su nivel alto o bajo en el índice PromEA.

La mitad izquierda de la Tabla 18 muestra que las mediciones de densidad nominal y léxica ocupan los primeros tres lugares no solo porque tienen el mayor poder predictivo (expresado con el valor de F), sino porque explican hasta un 20% del fenómeno de la variación. Este número refleja la bondad del ajuste del modelo de regresión a la variable que se pretende explicar, incluyendo la penalización por la interacción de variables no significativas.⁴² Con estos datos en orden, se puede proponer que tanto para la didáctica de la escritura académica, como para el diseño de pautas de escritura y rúbricas de evaluación, se ponga atención no únicamente al uso diverso del vocabulario, sino a que los tópicos desarrollados en el texto se expresen empleando un léxico denso, es decir, con alta proporción de palabras de contenido semántico: sustantivos, verbos, adjetivos y adverbios.

Ahora bien, la mitad derecha de la Tabla 18 jerarquiza los resultados de las pruebas dedicadas a establecer diferencias en los textos agrupados por nivel de promedio, alto o bajo, en el índice PromEA, según el uso del léxico abstracto y sus clases, especializado y abstracto especializado. El ordenamiento se hizo a partir del valor de t, que expresa el valor teórico de la diferencia entre los grupos.⁴³ En primer lugar aparece el uso de sustantivos

⁴² Cada uno de los modelos se diseñó para probar la interacción de dos y hasta tres variables en su efecto sobre PromEA. El coeficiente de determinación R^2 ajustado expresa el porcentaje explicado por la varianza de la regresión en relación con la varianza de la variable explicada; en otras palabras, la bondad del ajuste da confiabilidad al poder predictivo de la variable de interés porque considera, y descuenta, la varianza debida a otros factores, expresando de un modo realista el efecto explicativo.

⁴³ El valor crítico de t para 58 grados de libertad es de 1.67. Los coeficientes obtenidos para cada variable puesta a prueba fueron todos mayores al valor crítico, por lo que se valida la diferencia de medias entre los grupos de PromEA alto y bajo, con apoyo del nivel de significancia estadística expresado por el valor de p, siempre muy por debajo del error alfa (0.05).

abstractos, en segundo lugar el uso de sustantivos abstractos especializados, y en tercero una clase en particular de sustantivos abstractos, la cual comprende procesos, actividades, resultados, cualidades y estados de cosas que pueden ser resultado de la nominalización de un verbo, un adjetivo u otro sustantivo, aunque algunos de ellos pueden ser formas base de carácter abstracto por su semántica (*combate, tino, pertinencia, responsabilidad, debate, antigüedad*). Estos tipos de sustantivos serían los más influyentes como indicadores de un alto nivel de dominio de la escritura académica. Por último, el empleo de términos especializados y de la clase de sustantivos abstractos ‘constructos teóricos y conceptos’ ocupan las posiciones cuatro y cinco de la Tabla 18 en razón de que también tienen un alto valor como indicadores del nivel de dominio.

Las implicaciones didácticas de esta jerarquía nuevamente remiten a la necesidad de diseñar acciones didácticas encaminadas a la incorporación sistemática, intencionada y pertinente de este tipo de léxico en el bachillerato. La oportunidad que brinda ese nivel educativo es insuperable para modelar y evaluar el aprendizaje y uso en la escritura de estos recursos lingüísticos que claramente tienen una fuerte influencia en la evaluación hecha por expertos, lo cual se plantea en el siguiente apartado de conclusiones.

8. CONCLUSIONES

Los eventos sociales en que se usa la escritura como medio o artefacto son múltiples y ubicuos. Un grupo de ellos involucra prácticas letradas de gran impacto y trascendencia en la vida de las personas: se trata de aquellas que se realizan al interior de las instituciones educativas, las cuales, de manera general, se suelen cobijar bajo el término *academia*.

Al interior de esta, la producción de textos escritos es una actividad muy valorada porque puede cumplir, entre otras, funciones eminentemente evaluativas: los pares examinan y validan las contribuciones científicas de sus colegas, los profesores de mayor experiencia examinan la producción académica de los aspirantes a ocupar un lugar en una facultad, los investigadores mantienen o mejoran su estatus en una comunidad mediante publicaciones arbitradas, etcétera. Y, por otro lado, durante todo el periodo de escolarización, los profesores dictaminan si los miembros en formación de estas comunidades alcanzan el perfil deseado para acreditar cursos y avanzar al siguiente nivel educativo, generalmente, por medio de la escritura.

Desde este punto de vista, la competencia en escritura académica es una capacidad de la mayor importancia para los estudiantes de casi cualquier nivel educativo; pero lo es especialmente a partir de la educación media, ya que en el bachillerato se prepara a los jóvenes para optar por una carrera universitaria, es decir, para especializarse.

Por décadas, una de las preocupaciones centrales de los profesores universitarios ha sido la falta de herramientas conceptuales, de recursos lingüísticos y de referentes culturales de los jóvenes que egresan del bachillerato y que logran su ingreso al nivel superior, en muchos casos dificultosamente. Las demandas epistémicas y comunicativas de la universidad, más exigentes en grado y cantidad que las de los niveles previos, desnudan y ponen en evidencia una realidad que tiene muchos flancos, y que, por ello, es difícil de

abordar de una sola vez y desde una sola disciplina: no se ha sabido formar unas competencias lingüísticas y comunicativas sólidas que acompañen y sostengan los aprendizajes de los estudiantes.

La lingüística, en particular, puede contribuir a conocer mejor el problema que tanto preocupa a los sistemas educativos aportando descripciones y explicaciones de los factores léxico-gramaticales, pragmáticos, retóricos y discursivos implicados en el fenómeno.

En este sentido, la principal contribución del presente trabajo de tesis consiste en que, en efecto, se pudo verificar que la tarea de escritura de un texto académico motiva la incorporación, por parte de los estudiantes al final del bachillerato, de un repertorio léxico con características cercanas a las convenciones más valoradas por los expertos: un lenguaje léxicamente denso, diverso, con capacidad para empaquetar montos de información gracias al empleo de sustantivos abstractos, nominalizaciones y terminología especializada, que, en una parte, es también semánticamente abstracta; todo lo cual se observó al relacionar las mediciones de estos datos del léxico con las variaciones en la calificación otorgada por un grupo de jueces.

Resulta notable que, aunque la pauta que emplearon los estudiantes para escribir sus textos en clase no hacía referencia explícita a la obligación de atraer un determinado tipo de léxico, ni las fuentes específicas de donde debían tomarlo, y tampoco consignaba directrices sobre la manera de emplearlo, esos estudiantes reconocieron como propios del discurso solicitado los rasgos referidos y se adecuaron a ellos.

Sin embargo, no todos los estudiantes emplearon en la misma medida tales recursos, lo cual es indicativo de la necesidad de dar acompañamiento planificado, claro –y fundamentado en un marco teórico sólido– a la formación de competencias lingüísticas de los jóvenes que apuntalen a su vez otras competencias de valor académico desde la educación básica, y atender esa formación de manera especial en el bachillerato, haciendo

énfasis en la adquisición del léxico disciplinar y analizando con detalle las características del género que se va a producir.

De este trabajo de tesis se desprende, asimismo, la necesidad de considerar las condiciones de producción de los géneros académicos, su anclaje en el discurso de las disciplinas, su estrecha relación con la evaluación de los aprendizajes en la educación formal, así como con la exigencia de proporcionar a todos los alumnos iguales oportunidades para acceder al conocimiento y para pertenecer y participar en las prácticas letradas que les atañen como miembros de la sociedad.

Como ha quedado dicho a lo largo de este trabajo, alfabetizarse o alfabetizar a alguien académicamente rebasa la habilitación en procesos cognitivos y lingüísticos: significa situar la escritura en un contexto más amplio (cultural, social, disciplinar), cuyas restricciones y exigencias se le imponen a los estudiantes de necesidad (Beke, 2011). El dominio de la escritura académica, junto con su correlato, la oralidad académica, serviría entonces como uno de los parámetros de la capacidad de un estudiante de participar en una comunidad discursiva que le impone metas, mecanismos y medios para comunicarse y obtener retroalimentación, y le coloca, asimismo, en la necesidad de conocer, usar y producir ciertos géneros, emplear un léxico específico y, en general, unas ciertas funciones del lenguaje, guiado por otros miembros expertos de la misma comunidad (Teberosky, 2007).

En el entendido de que el texto académico cristaliza organizaciones retóricas por medio de recursos lingüísticos que difieren funcionalmente de los empleados en la lengua coloquial, se sostiene la tesis de que las principales dificultades de los estudiantes –mismas que son capaces de interferir en su trayectoria académica– radican en la falta de familiaridad con una variedad de discurso con un alto nivel de abstracción, convencionalización y especialización, por lo que es imperativo que los estudiantes dispongan de acompañamiento didáctico en el manejo de este tipo de discurso, no en un

solo taller descontextualizado, y no en una sola ocasión, sino de manera situada, sostenida y a través del currículum.

Como lo señala Ávila (1991), “el aprendizaje fuera del ambiente escolar es circunstancial y aleatorio. Frente a esto, la escuela debe ofrecer –por contraste– una enseñanza planificada y eficiente. Para hacerlo habría que partir de lo que ya sabe el alumno y proceder en consecuencia. La lingüística podría contribuir a ese conocimiento mediante la descripción del lenguaje de los estudiantes, y más específicamente, del léxico” (p. 626).

Subsidiariamente, el conocimiento derivado de esta investigación puede dirigir la atención hacia el tratamiento didáctico del léxico en la escritura académica, que se adecue al nivel educativo enfocado y al tipo de textos que aquí se analizaron. Puede asimismo motivar la reflexión pedagógica acerca de la importancia de promover y fortalecer la competencia en escritura académica en el tránsito hacia el nivel superior; puede motivar el diseño de acciones estratégicas al interior de centros y programas de escritura académica dirigidas a estudiantes de nuevo ingreso en instituciones de ese nivel, y también puede servir como base para otros proyectos lingüísticos que abonen a una profundización en el campo de la escritura académica.

Finalmente, se puede decir que los contrastes cuantitativos en el uso del léxico observados en este trabajo autorizan a inferir la existencia de diferencias cualitativas en el lenguaje empleado por los estudiantes en la construcción de sus textos que merecerían un estudio más detallado. Los aspectos no explorados o no totalmente logrados de esta tesis abren oportunidades para futuras investigaciones que consideren, entre otros aspectos: extender el análisis a poblaciones compuestas por estudiantes en varias etapas de la escolarización formal, incluidos los universitarios; contrastar lo que sucede en la construcción de otros géneros con diferentes propósitos y situados en diversos campos

disciplinares; evaluar cualitativamente los usos del léxico en el discurso producido por los estudiantes, e incluso controlar, experimentalmente, variables que no pueden aislarse en un estudio como el que se presentó aquí.

9. BIBLIOGRAFÍA

- Ağçam, R. (2015). Author stance in academic writing: a corpus-based study on epistemic verbs. *The Journal of Teaching English for Specific and Academic Purposes*, 3(1), 9-20. Recuperado de https://www.academia.edu/12721949/Author_Stance_in_Academic_Writing_A_Corpus-Based_Study_on_Epistemic_Verbs
- Aissen, J. (2006). Taller sobre la Complementación. Centro de Investigaciones y de Estudios Superiores en Antropología Social-Sureste. México, del 31 de julio al 10 de agosto de 2006.
- Aktas, R. N. (2005). *Functions of “shell nouns” as cohesive devices in academic writing: a comparative corpus-based study*. [Retrospective Theses and Dissertations. Paper 16168]. Iowa: Iowa State University Digital Repository. Recuperado de <http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=17167&context=rtd>
- Aktas, R. N. & Cortes, V. (2005). Shell nouns as cohesive devices in published and ESL students writing. *Journal of English for Academic Purposes*, 7(1), 3-14. Recuperado de <https://www.sciencedirect.com/science/article/pii/S1475158508000039>
- Alarcos Llorach, E. (1995). *Gramática de la Lengua Española*. Madrid: RAE-Espasa.
- Alcaraz, E. & Martínez, M. A. (1997). *Diccionario de lingüística moderna*. Barcelona: Ariel.
- Alotaibi, H. (2014). The role of lexical cohesion in writing quality. *International Journal of Applied Linguistics & English Literature*, 4(1), 261-269. Recuperado de <http://www.journals.aiac.org.au/index.php/IJALEL/article/view/1261>
- Álvarez, T. (2010). *Competencias básicas en escritura*. Barcelona: Octaedro.
- Anglin, J. M. (1993). Vocabulary development: a morphological analysis. *Monographs of the Society for Research in Child Development*, 58(10), 238.

- Anstey, M. & Bull, G. (2018). *Foundations of Multiliteracies. Reading, Writing and Talking in the 21st Century*. Nueva York: Routledge.
- ANUIES (2014). *Habilidades lingüísticas de los estudiantes de primer ingreso a las instituciones de educación superior del área metropolitana de la ciudad de México. Desempeño en el EXHALING 2011*. Recuperado de <http://asambleaanui.es.iberomexico.mx/wp-content/uploads/2013/11/Estudio-EXHALING-6-nov.pdf>
- Ávila, R. (1991). Densidad léxica y adquisición del vocabulario: niños y adultos. En C. Hernández, G. De Granda, C. Hoyos, V. Fernández, D. Dietrick & Y. Carballera (coords.). *El español de América. Actas del III Congreso Internacional de El Español de América* (pp. 621-630). Valladolid: Junta de Castilla y León.
- Backhoff, E., Larrazolo, N. & Tirado, F. (2011). Habilidades verbales y conocimientos del español de estudiantes egresados del bachillerato en México. *Revista de la Educación Superior*, 160(4), 9-127. Recuperado de http://publicaciones.anui.es.iberomexico.mx/pdfs/revista/Revista160_S1ES.pdf
- Bailey, A. L. (2007). *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.
- Bajtín, M. M. (2005). *Estética de la creación verbal*. México: Siglo XXI.
- Barton, D. & Hamilton, M. (2010). La littératie: une pratique sociale. *Langage et Société*, 3(133), 45-62. doi: 10.3917/l.s.133.0045
- Bazerman, Ch. (2005). *Reference guide to Writing Across the Curriculum*. Indiana: Parlor Press/The WAC Clearinghouse.
- Bednarek, M. (2006). *Evaluation in Media Discourse*. Londres: Continuum.
- Bello, I. (2016). Cognitive implications of nominalization in the advancement of scientific discourse. *International Journal of English Studies*, 16 (2), 1-23.
- Benitez-Castro, M.-A. (2015). Coming to Grips with Shell-nounhood: A Critical Review of Insights into the Meaning, Function and Form of Shell-noun Phrases. *Australian Journal of Linguistics*, 35(2), 168-194. doi: 10.1080/07268602.2015.1005001

- Beke, R. (2005). El metadiscurso interpersonal en artículos de investigación. *Signos*, 38(57), 7-18. <http://dx.doi.org/10.4067/S0718-09342005000100001> Recuperado de http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342005000100001
- Beke, R. (2011). *Las voces de los otros en el discurso académico*. Caracas: Comisión de Estudios de Postgrado, Facultad de Humanidades y Educación, Universidad Central de Venezuela.
- Bello, I. (2016). Cognitive implications of nominalizations in the advancement of scientific discourse. *International Journal of English Studies*, 16(2), 1-23. <http://dx.doi.org/10.6018/ijes/2016/2/262921>. Recuperado de <http://revistas.um.es/ijes/article/view/262921>
- Bereiter, M. & Scardamalia, C. (1987). *The psychology of written composition*. Hillsdale: Erlbaum.
- Berman R. A. (2004). Between emergence and mastery: The long developmental route of language acquisition. En R. A. Berman (Ed.). *Language development across childhood and adolescence* (pp. 9-34). Amsterdam: John Benjamins.
- Berman, R. & Nir, B. (2009). Cognitive and linguistic factors in evaluating text quality. Global versus local? En V. Evans & S. Pourcel (Eds.). *New direction in cognitive linguistics* (pp. 421-440). Amsterdam: John Benjamins. Recuperado de https://www.researchgate.net/profile/Bracha_Nir/publication/290836613_Cognitive_and_linguistic_factors_in_evaluating_text_quality/links/56a2664808aeef24c585d211/Cognitive-and-linguistic-factors-in-evaluating-text-quality.pdf
- Berman, R. & Nir, B. (2010). The Lexicon in Writing-Speech-Differentiation: Developmental perspectives. *Written Language & Literacy*, 13(2), 183–205. John Benjamins. Recuperado de https://www.researchgate.net/profile/Ruth_Berman/publication/233500108_The_Lexicon_in_Writing-Speech-

Differentiation_Developmental_perspectives/links/55334f040cf20ea0a074cc89/The-Lexicon-in-Writing-Speech-Differentiation-Developmental-perspectives.pdf

- Berman, R. & Slobin, D. (1994). Glossing and Transcript conventions. En R. Berman & D. Slobin (Eds.). *Relating Events in Narrative: A Crosslinguistic Developmental Study* (pp. 655-664). New York: Lawrence Erlbaum.
- Berman, R. & Verhoeven, L. (2002). Cross-Linguistic Perspectives on the Development of Text-Production Abilities: Speech and Writing. *Written Language & Literacy*, 5(1-2), 1-43. doi: 10.1075/wll.5.1.02ber. Recuperado de https://www.researchgate.net/publication/233489499_Cross-Linguistic_Perspectives_on_the_Development_of_Text-Production_Abilities_Speech_and_Writing
- Berman, R., Nayditz, R. & Ravid, D. (2011). Linguistic diagnostics of written texts in two school-age populations. *Written Language & Literacy*, 14(2), 161-187. doi: 10.1075/wll.14.2.01ber. John Benjamins Publishing Company. Recuperado de <https://doritravid.files.wordpress.com/2014/10/berman-nayditz-ravid.pdf>
- Betancourt, M. (2009). *Literatura comparada clásica*. México: McGraw Hill.
- Biber, D. (2006). *University language. A corpus-based study of spoken and written registers*. Filadelfia: John Benjamins (Studies in Corpus Linguistics, 23).
- Biggs, J. & Tang, C. (2011). *Teaching for Quality Learning at University. What the Student Does*. Londres: McGraw-Hill/Open University Press. Recuperado de http://hust.edu.oak.arvixe.com/media/197963/-John_Biggs_and_Catherine_Tang_Teaching_for_Quali-BookFiorg-.pdf
- Bloomfield, L. (1939). *Linguistic aspects of science*. Chicago: The University of Chicago Press.
- Bogard, S. (2004). La estructura de la cláusula. Hacia una metodología del análisis sintáctico. En G. Báez & E. Luna Traill (Eds.). *Disquisiciones sobre Filología Hispánica. In Memoriam Juan M. Lope Blanch* (pp.73-80). México: UNAM.

- Bolívar, A. (2004). Análisis crítico del discurso de los académicos. *Signos*, 37 (55), 1-12.
<http://dx.doi.org/10.4067/S0718-09342004005500001>. Recuperado de https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-09342004005500001
- Bolívar, A. (2005). *Discurso e interacción en el texto escrito*. Caracas, Venezuela: Universidad Central de Venezuela/Consejo de Desarrollo Científico y Humanístico.
- Bolívar, A. (2007). El análisis internacional del discurso: del texto a la dinámica social. En A. Bolívar (Comp.) *Análisis del discurso. ¿Por qué y para qué?* (pp. 247-277). Caracas Venezuela: Universidad Central de Venezuela/Los Libros de El Nacional.
- Bolívar, A. & Beke, R. (2014). *Lectura y escritura para la investigación*. Caracas: Universidad Central de Venezuela.
- Bolívar, A. & Parodi, G. (2015). Academic and Professional Discourse. En M. Lacorte (Ed.). *The Routledge Handbook of Hispanic Applied Linguistics* (pp. 459-475). Nueva York: Routledge. Recuperado de https://www.academia.edu/21694608/26_Academic_and_Professional_Discourse
- Bolívar, A., Beke, R. & Shiro, M. (2010). Las marcas lingüísticas de posicionamiento en las disciplinas: Estructuras, voces y perspectivas discursivas. En G. Parodi (Ed.). *Alfabetización académica y profesional en el siglo XXI: leer y escribir desde las disciplinas*. Santiago de Chile: Editorial Planeta.
- Borreguero, M. (2018). Los encapsuladores anafóricos: una propuesta de clasificación. *Caplletra*, 64, 179-203. doi: 10.7203/Caplletra.64.11380
- Bourdieu, P. (1997). *Capital cultural, escuela y espacio social*. México: Siglo XXI Editores.
- Bosque, I. (2011). *Nueva gramática de la lengua española* (eBook pack). Madrid: Espasa/RAE.

- Bravo, N. (2007). *Competencias Proyecto Tuning-Europa, Turing-América Latina*. Informe. Bogotá, Col. Recuperado de http://www.cca.org.mx/profesores/cursos/hmfbcp_ut/pdfs/m1/competencias_proyectotuning.pdf
- Cademártori, Y., Parodi, G. & Venegas, R. (2006). El discurso escrito y especializado: caracterización y funciones de las nominalizaciones en los manuales técnicos. *Literatura y Lingüística*, 17, 243-265.
- Calle, L., Pico, A. L. & Murillo, J. H. (2017). Los Centros de Escritura: entre nivelación académica y construcción del conocimiento. *Cadernos de Pesquisa*, 47(165), 872-895.
- Calsamiglia, H. & Tusón, A. (2004). *Las cosas del decir. Manual de análisis del discurso*. Barcelona: Ariel.
- Camps, A. & Castelló, M. (2013). La escritura académica en la universidad. *Revista de Docencia Universitaria*, 11 (1), 17-36. Recuperado de <http://red-u.net/redu/files/journals/1/articles/579/public/579-2168-1-PB.pdf>
- Carlino, P. (2003). Alfabetización académica. Un cambio necesario, algunas alternativas posibles. *Educere*, 6(20), 409-420.
- Carlino, P. (2012). *Escribir, leer y aprender en la universidad. Una introducción a la alfabetización académica*. México: FCE.
- Cassany, D. (Comp.). (2009). *Para ser letrados. Voces y miradas sobre la lectura*. Barcelona: Paidós.
- Cassany, D. (2012). *En línea. Leer y escribir en la red*. Barcelona: Paidós.
- Cassany, D. & López-Ferrero, C. (2010). De la universidad al mundo laboral: continuidad y contraste entre las prácticas letradas académicas y profesionales. En G. Parodi (Ed.). *Alfabetización académica y profesional en el siglo XXI: leer y escribir desde las disciplinas* (pp. 347-374). Barcelona: Planeta Ariel.
- Castelló, M. (Coord.) (2007). *Escribir y comunicarse en contextos científicos y académicos: conocimientos y estrategias*. Barcelona: Graó.
- Castelló, M. (2014). Los retos actuales de la alfabetización académica: estado de la cuestión y últimas investigaciones. *Enunciación*, 19(2), 346-365. doi: <http://dx.doi.org/10.14483/udistrital.jour.enunc.2014.2.a13>

- CENEVAL (2017). *Resultados del Examen Nacional de Ingreso a la Educación Superior en el año 2017. EXANI-II Diagnóstico: Admisión*. Recuperado de http://www.ceneval.edu.mx/documents/20182/74837/Estadistica_EXANI-II_2017/b26c52bd-2ae1-4788-bd9d-8b3c152163e4
- Chafe, W. L. (1985). Linguistic differences produced by differences between speaking and writing. En D. R. Olson, N. Torrance, N. & A. Hildyard (Eds.). *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing* (pp. 105-123). Cambridge: Cambridge University Press.
- Chafe, W. L. & Danielewicz, J. (1987). Properties of Spoken and Written Language. En R. Horowitz y F. J. Samuels (Eds.). *Comprehending Oral and Written Language*. Nueva York: Academic Press. Recuperado de https://www.nwp.org/cs/public/download/nwp_file/142/TR05.pdf?x-r=pcfile_d
- Charadeau, P. & Maingueneau, D. (2005). *Diccionario de análisis del discurso*. Buenos Aires: Amorrortu.
- Charles, M. (2003). 'This mystery. . .': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, 2, 313-326. Recuperado de www.elsevier.com/locate/jeap
- Charles, M. (2007). Argument or evidence? Disciplinary variation in the use of the Noun that pattern in stance construction. *English for Specific Purposes*, 26, 203-218. Recuperado de www.elsevier.com/locate/esp
- Chung, T. & Nation, I. S. P. (2004). Identifying technical vocabulary. *System*, 32(2), 251-263. doi:10.1016/j.system.2003.11.008. Recuperado de <https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/2004-Chung-Identifying-Tech-Voc.pdf>
- Cisneros, M. (2006). *Lectura y escritura en la universidad*. Pereira: Universidad Tecnológica de Pereira.
- Cisneros, M. & Muñoz, C. (2014). *Tras las huellas de las investigaciones sobre Lectura y Escritura en la Universidad*. Pereira: Universidad Tecnológica de Pereira.
- Clark, E. V. (1995). *The lexicon in acquisition*. Nueva York: Cambridge University Press.

- CONBIOÉTICA (7 de diciembre de 2018). *Capacitación en Ética de Investigación UAQ-CONBIOÉTICA*. Capítulo 3. Diseño de las investigaciones. Disponible en http://uaqvirtual.uaq.mx/campusvirtual/filosofia/pluginfile.php/9973/mod_resource/content/7/Capitulo3.pdf
- CONBIOÉTICA (7 de diciembre de 2018). *Capacitación en Ética de Investigación UAQ-CONBIOÉTICA*. Capítulo 4. Consentimiento informado. Disponible en http://uaqvirtual.uaq.mx/campusvirtual/filosofia/pluginfile.php/9975/mod_resource/content/6/Capitulo4.pdf
- Coulthard, M. (Ed.). (1994). *Advances in Written Text Analysis*. Londres: Routledge.
- Coxhead, A. (2017). Approaches and perspectives on teaching vocabulary for discipline-specific academic writing. En J. Flowerdew & T. Costley (Eds.). *Discipline-Specific Writing. Theory into practice* (pp. 62-76). Nueva York: Routledge.
- Crossley, S. A. & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *Int. J. Continuing Engineering Education and Life-Long Learning* 21 (2-3), 170-191. Recuperado de ftp://129.219.222.66/Publish/pdf/Understand_Expert_Ratings_of_Essay_Quality.pdf
- Crossley, S. A., Kyle, K. & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1-16. <http://dx.doi.org/10.1016/j.jslw.2016.01.003> Recuperado de <https://www.semanticscholar.org/paper/The-development-and-use-of-cohesive-devices-in-L2-Crossley-Kyle/e67c2e2c16f5e285863f0244406d5884e73de009>
- Crossley, S. A., Salisbury, T., & McNamara, D. S. (2011). Predicting the proficiency level of language learners using lexical indices. *Language Testing* 29 (2), 243-263, [doi.org/10.1177/0265532211419331](http://dx.doi.org/10.1177/0265532211419331). Recuperado de <http://journals.sagepub.com/doi/abs/10.1177/0265532211419331#articleCitationDownloadContainer>
- Crossley, S. A., Salisbury, T., McNamara, D. S. & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, 28(4), 561-580, [doi.org/10.1177/0265532210378031](http://dx.doi.org/10.1177/0265532210378031). Recuperado de <http://journals.sagepub.com/doi/abs/10.1177/0265532210378031#articleCitationDownloadContainer>

- Cubo, L. (2014). Escritura de formación en la universidad. En F. Navarro (Coord.) *Manual de escritura para carreras de humanidades* (pp. 17-25). Buenos Aires: Editorial de la Facultad de Filosofía y Letras Universidad de Buenos Aires.
- De la Fuente, S. (2011). *Análisis discriminante*. Universidad Autónoma de Madrid. Recuperado de <http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/SEGMENTACION/DISCRIMINANTE/analisis-discriminante.pdf>
- Denny, P. (1998). El pensamiento racional en la cultura oral y la descontextualización escrita. En D. R. Olson, N. Torrance & A. Hildyard (Comps.). *Cultura escrita y oralidad* (pp. 95-126). Barcelona: Gedisa.
- Díaz, C., Gracida, Y. & Sule, T. (2014). Luces y sombras de la educación lingüística en México. En C. Lomas (Ed.). *La educación lingüística, entre el deseo y la realidad. Competencias comunicativas y enseñanza del lenguaje* (pp. 176-188). México: FLACSO/Octaedro.
- Díez, M. (2018). Abstract nouns as metadiscursive shells in academic discourse. *Caplletra*, 64, 153-178. doi: 10.7203/Caplletra.64.11372
- Du Bois, J. W. (2007). The Stance Triangle. En R. Englebreston (Ed.). *Stancetaking in Discourse. Subjectivity, evaluation, interaction* (pp. 139-182). Filadelfia: John Benjamins Publishing Company.
- Dussel, I. & Southwell, M. (2007). La escuela y las nuevas alfabetizaciones. *Boletín de la Organización de los Estados Iberoamericanos para la Educación, la Ciencia y la Cultura*. Recuperado de <https://www.oei.es/historico/noticias/spip.php?article11326>
- Enríquez, A. & Sánchez, E. (Coords.). (2017). *Dominio del español de estudiantes universitarios. Retos y perspectivas*. Morelia, México: Universidad Michoacana de San Nicolás de Hidalgo.
- Ezcurra, A. M. (2011). Abandono estudiantil en educación superior. Hipótesis y conceptos. En N. Gluz (Ed.). *Admisión a la universidad y selectividad social. Cuando la democratización es más que un problema de "ingresos"* (pp. 23-62). Los Polvorines, Buenos Aires: Universidad Nacional de General Sarmiento.

- Fang, Z., Schleppegrell, M. J. & Cox, B. E. (2006). Understanding the Language Demands of Schooling: Nouns in Academic Registers. *Journal of Literacy Research*, 38(3), 247–273. Recuperado de http://journals.sagepub.com/doi/pdf/10.1207/s15548430jlr3803_1
- Figueroa, J. (2018). *Escribir explicaciones y argumentaciones en la escuela: calidad de la escritura y lenguaje académico (habilidades y recursos) en estudiantes de 8o. básico*. [Tesis doctoral]. Santiago de Chile: Pontificia Universidad Católica de Chile-Facultad de Educación. Recuperado de https://repositorio.uc.cl/bitstream/handle/11534/22051/Compendio%20de%20art%C3%ADculos_Javiera_Figueroa_Miralles.pdf?sequence=1&isAllowed=y
- Figueroa, J., Meneses, A. & Chandía, E. (2018). Academic language and the quality of written arguments and explanations of Chilean 8th graders. *Reading and Writing*, 31(3), 703–723. doi: 10.1007/s11145-017-9806-5
- Foley, W. (2007). A typology of information packaging in the clause. En T. Shopen (Ed.). *Language Typology and Syntactic Description* (pp. 362-446). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511619427.007. Recuperado de https://www.academia.edu/28896589/A_typology_of_information_packaging_in_the_clause
- Flowerdew, J. (Ed.) (2002). *Academic discourse*. Cambridge: Cambridge University Press.
- Flowerdew, J. (2003). Signalling nouns in discourse. *English for Specific Purposes*, 22, 329-346. Recuperado de https://www.academia.edu/5662093/Signalling_nouns_in_discourse
- Flowerdew, J. (2006). Use of signalling nouns in a learner corpus. *International Journal of Corpus Linguistics*, 11(3), 345-362. Recuperado de https://www.academia.edu/3654989/Use_of_signalling_nouns_in_a_learner_corpus
- Flowerdew, J. (2015). Revisiting metadiscourse: Conceptual and methodological issues concerning signalling nouns. *Ibérica*, 29, 15-34. Recuperado de <https://www.academia.edu/21274904/>

Revisiting_metadiscourse_Conceptual_and_methodological_issues_concerning_signalling_nouns

- Flowerdew, J. & Costley, T. (Eds.). (2017). *Discipline-Specific Writing. Theory into practice*. Nueva York: Routledge.
- Flowerdew, J. & Forest, R. W. (2015). *Signalling nouns in English. A corpus-based discourse approach*. Cambridge: Cambridge University Press.
- Francis, G. (1994). Labelling discourse. En Coulthard, M. (Ed.). *Advances in Written Text Analysis* (pp. 83-101). Londres: Routledge.
- Galán, R. M. & Ormsby, L. (2010). El Centro de Escritura como puente en la educación superior: el Centro de Aprendizaje, Redacción y Lenguas del ITAM. En G. Ruiz, M. G. López & A. Carrasco (Coords.). *Memoria de las ponencias del II Seminario Internacional de Lectura en Instituciones de Educación Media Superior y Superior* (173-181). Puebla: Consejo Puebla de Lectura/Universidad Autónoma de Aguascalientes.
- García, R. I. (2005). Mecanismos de cohesión léxica: análisis comparativo de algunas propuestas clasificatorias. *Interlingüística*, 16(1), 503-515. Recuperado de <https://dialnet.unirioja.es/descarga/articulo/2517066.pdf>
- García, M. M, Hall, B. & Marín, M. (2005). Ambigüedad, abstracción y polifonía del discurso académico: Interpretación de las nominalizaciones. *Revista signos*, 38(57), 49-60. <https://dx.doi.org/10.4067/S0718-09342005000100004>
- Garrido, M. A. (2017). *Diccionario Español de Términos Literarios Internacionales*, DETLI. Madrid: Instituto de Lengua, Literatura y Antropología/Consejo Superior de Investigaciones Científicas/Union Académique Internationale. Disponible en <http://www.proyectos.cchs.csic.es/detli/node/1>
- Garton, A. & Pratt, C. (2000). *Learning to be literate. The development of spoken and written language*. Oxford: Blackwell Publishers.
- Gili Gaya, S. (2002). *Curso superior de sintaxis española*. Barcelona: Vox.
- Givón, T. (Ed.) (1983). *Topic continuity in discourse*. Filadelfia: John Benjamins.
- Givón, T. (1990). *Syntax. A functional-typological introduction*, 2 vol. Filadelfia: John Benjamins.

- Gómez, G., Sotomayor, C., Bedwell, P., Domínguez, A. M. & Jéldrez, E. (2016). Analysis of lexical quality and its relación to writing quality for 4th grade, primar y school students in Chile. *Read Writ*, 29, 1317-1336. doi: 10.1007/s11145-016-9637-9. Recuperado de <http://repositorio.uchile.cl/handle/2250/142408>
- Hallebeek, J. (1986) Las palabras funcionales del español. *Boletín de la Asociación Europea de Profesores de Español*, XVIII(34-35), primavera-otoño de 1986. Recuperado de Centro Virtual Cervantes http://cvc.cervantes.es/ensenanza/biblioteca_ele/aepe/pdf/boletin_34-35_18_86/boletin_34-35_18_86_24.pdf
- Halliday, M. A. K. (1990). *Spoken and written language*. Oxford: Oxford University Press.
- Halliday, M. A. K. (1994). The construction of knowledge and value in the grammar of scientific discourse, whit reference to Charles Darwin's *The Origin of Species*. En M. Coulthard (Ed.). *Advances in Written Text Analysis* (pp. 136-156). Londres: Routledge.
- Halliday, M. A. K. (2013). *El lenguaje como semiótica social. La interpretación social del lenguaje y del significado*. México: FCE.
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. Londres: Longman.
- Halliday, M. A. K. & Martin, J. (1993). *Writing Science*. Londres: The Falmer Press.
- Hess, K. (2013). Desarrollo léxico en la adolescencia: Un análisis de sustantivos en narraciones orales y escritas. *Actualidades en Psicología*, 27(115), 113-127. Repositorio Institucional de la Universidad de Costa Rica. Recuperado de <http://repositorio.ucr.ac.cr/handle/10669/12686>
- Hunston, S. (2011). *Corpus Approaches to Evaluation. Phraseology and Evaluative Language*. Nueva York: Routledge.
- Hyland, K. (2002). Genre, language, context, and literacy. *Annual Review of Applied Linguistics*, 22, 113-135. Cambridge University Press.
- Hyland, K. (2004). *Disciplinary discourse: social interactions in academic writing*. Londres: Longman.
- Hyland, K. (2005a). *Metadiscourse: Exploring interaction in writing*. Londres: Continuum.

- Hyland, K. (2005b). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7, 173-192. Sage Publications. Recuperado de <http://dis.sagepub.com/cgi/content/refs/7/2/173>
- Hyland, K. (2015) What is metadiscourse? En K. Tracy (Ed.). *The international encyclopedia of language and social interaction* (pp. 1-11). Nueva Jersey: John Wiley & Sons. Recuperado de https://www.researchgate.net/publication/285591598_Metadiscourse
- Hyland, K. & Tse, P. (2004). Metadiscourse in Academic Writing: A Reappraisal. *Applied Linguistics* 25 (2), 156-177. Recuperado de https://www.researchgate.net/publication/31418041_Metadiscourse_in_Academic_Writing_A_Reappraisal
- INEE (2019). *Informe de resultados PLANEA EMS 2017. El aprendizaje de los alumnos de Educación Media Superior en México. Lenguaje y Comunicación y Matemáticas*. México: INEE. Recuperado de <https://www.inee.edu.mx/publicaciones/resultados-planea-ems-2017/>
- Jarpa, M. (2015). Escritura en las disciplinas: Géneros académicos evacuativos en un programa de posgrado de Biotecnología. En G. Parodi & G. Burdiles (Eds.). *Leer y escribir en contextos académicos y profesionales: géneros, corpus y métodos* (pp. 221-255). Santiago de Chile: Ariel.
- Jiang, F. & Hyland, K. (2015). ‘The fact that’: Stance nouns in disciplinary writing. *Discourse Studies*, 17(5), 1-22. Sage Publications. Recuperado de https://www.academia.edu/14734868/_The_fact_that_Stance_nouns_in_disciplinary_writing
- Jiang, F. & Hyland, K. (2016). Nouns and Academic Interactions: A Neglected Feature of Metadiscourse. *Applied Linguistics*, 23, 1-25. Oxford University Press. Recuperado de https://www.academia.edu/27356259/Nouns_and_Academic_Interactions_A_Neglected_Feature_of_Metadiscourse

- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: a developmental perspective. *Working Papers in Linguistics*, 53, 61-79. Recuperado de <http://journals.lub.lu.se/index.php/LWPL/article/view/2273>
- Johansson, V. (2009). *Developmental Aspects of Text Production in Writing and Speech*. [Tesis doctoral]. Lund: Lund University.
- Kaplan, N. (2004). Nuevos desarrollos en el estudio de la evaluación en el lenguaje: La Teoría de la Valoración. *Boletín de Lingüística*, 22, julio-diciembre, 52-78. Recuperado de <http://www.redalyc.org/articulo.oa?id=34702203>
- Kerlinger, F. N. & Lee, H. B. (2002). *Investigación del comportamiento. Métodos de investigación en ciencias sociales*. México: McGrawHill.
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155-163. doi.org/10.1016/j.jcm.2016.02.012. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4913118/>
- Kroeger, P. R. (2005). *Analyzing grammar: An introduction*. Nueva York: Cambridge University Press.
- Langacker, R. W. (1987). *Foundations of cognitive grammar*, 2 vol. Stanford: Stanford University Press.
- Lea, M. R., & Street, B. V. (2006). The “Academic Literacies” Model: Theory and Applications. *Theory Into Practice*, 45(4), 368-377. Recuperado de https://www.researchgate.net/publication/47343136_The_Academic_Literacies_Model_Theory_and_Applications
- Lemke, J. (1998). Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text. En J.R. Martin & R. Veel (Eds.), *Reading science* (pp. 87-113). Londres: Routledge. Recuperado de https://www.researchgate.net/publication/246905867_Multiplying_meaning_Visual_and_verbal_semiotics_in_scientific_text
- Levshina, N. (2015). *How to do linguistics with R*. Filadelfia: John Benjamins.

- López, A. (2011). *La categorización de entidades del discurso en la escritura profesional. Las etiquetas discursivas como mecanismo de cohesión léxica*. [Tesis doctoral]. Barcelona: Universidad de Barcelona-Facultad de Filología. Recuperado de http://diposit.ub.edu/dspace/bitstream/2445/35053/7/02.ALS_2de4.pdf
- López, A. (2013). Las etiquetas discursivas: del mantenimiento a la construcción del referente. *ELUA, Estudios de Lingüística de la Universidad de Alicante*, 27, 167-197. Recuperado de <http://rua.ua.es/dspace/handle/10045/36627>
- López, A. (2015). Etiquetas discursivas, hiperónimos y encapsuladores: una propuesta de clasificación de las relaciones de cohesión referencial. *RILCE, Revista de Filología Hispánica* 31 (2), 435-462. Recuperado de <https://www.unav.edu/publicaciones/revistas/index.php/rilce/article/view/1551>
- López, A. (2018). La encapsulación nominal en el discurso académico-científico oral y escrito: patrones de aparición. *Caplletra*, 64, 129-152. doi: 10.7203/Caplletra.64.11369
- López, H. (2011). Los índices de ‘riqueza léxica’ y la enseñanza de lenguas. En De Santiago Guervós, J., Bongaerts, H., Sánchez Iglesias, J. J. y Seseña, M. (coords.). *Del texto a la lengua: La aplicación de los textos a la enseñanza-aprendizaje del español L2-LE*, vol. I, 15-28.
- Lyons, J. (1977). *Semantics*, vol. 2. Nueva York: Cambridge University Press.
- MacCarthy, Ph. M. & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392. Recuperado de <http://link.springer.com/article/10.3758/BRM.42.2.381>
- MacDonald, S. P. (1994). *Professional Academic Writing in the Humanities and the Social Sciences*. Carbondale: Southern Illinois University Press.
- Malvern, D. D., Richards, B. J., Chipere, N. & Durán, P. (2004). *Lexical diversity and language development: quantification and assessment*. Londres: Palgrave Macmillan.

- Marinkovich, J. (2007). La nominalización como un recurso de cohesión léxica en los manuales de la formación técnico-profesional. En G. Parodi (Ed.). *Lingüística de corpus y discursos especializados: puntos de mira* (pp. 97-113). Valparaíso, Chile: Ediciones Universitarias de Valparaíso. Recuperado de http://www.giovanniparodi.cl/giovanniparodi/site/artic/20150930/asocfile/20150930091236/2007linguistica_de_corpus_y_discurso_especializado.pdf
- Martin, J. R. & White, P. R. R. (2005). *The Language of Evaluation. Appraisal in English*. Nueva York: Palgrave Macmillan.
- Martín, J. C. (2004). *El vocabulario del discurso tecnocientífico*. Madrid: Arco Libros.
- Martínez, R. (2010). *Evaluar la escritura en Chile. La dimensión de escritura del Sistema de Medición de la Calidad de la Educación Chilena (SIMCE) desde los modelos cognitivos de la producción escrita*. [Tesis de magíster]. Santiago de Chile: Universidad de Chile, Facultad de Filosofía y Humanidades.
- Matsuda, P. K., & Jeffery, J. V. (2012). Voice in student essays. En K. Hyland y C. Sancho-Guinda (Eds.). *Voice and Stance in Academic Writing* (pp. 151-165). Nueva York: Palgrave Macmillan.
- McEnery, T. & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Nueva York: Cambridge University Press.
- McNamara, D. S., Floyd, R. G., Best, R. & Louwse, M. (2004). World knowledge driving young readers' comprehension difficulties. En *Proceedings of the 6th international conference on Learning sciences (ICLS '04)*. International Society of the Learning Sciences (pp. 326-333). Recuperado de https://www.researchgate.net/publication/228875498_World_knowledge_driving_young_readers%27_comprehension_difficulties
- McNamara, D. S., Crossley, S. A. & McCarthy, P. M. (2009). Linguistic Features of Writing Quality. *Written Communication*, 27(1), 57-86. Recuperado de <http://journals.sagepub.com/doi/abs/10.1177/0741088309351547>
- Meneses, A., Uccelli, P., Santelices, M. V., Ruiz, M., Acevedo, D. & Figueroa, J. (2017). Academic language as a predictor of reading comprehension in monolingual spanish-

- speaker readers: evidence from chilean early adolescents. *Reading Research Quarterly*, 0(0), 1-25. International Literacy Association. doi: 10.1002/rrq.192
- Molina, V. (2014). Centros de Escritura: una mirada retrospectiva para entender el presente y futuro de estos programas en el contexto latinoamericano. *Legenda*, 18(18), 9-32.
- Monippally, M. M. & Shankar Pawar, B. (2010). *Academic writing. A guide for management students and researchers*. Nueva Delhi: Response.
- Montes, M. & López, G.(2017). Literacidad y alfabetización disciplinar: enfoques teóricos y propuestas pedagógicas. *Perfiles Educativos*, 34(155), 162-178.
- Montoya, S. (2018). *Defining literacy*. Hamburgo: GALM Fifth Meeting, UNESCO.
- Morales, F. (2004). Evaluar la escritura, sí... Pero ¿Qué y cómo evaluar? *Acción Pedagógica*, 13(1), 38-48. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=2971903>
- Munguía, I. (2017). *Gramática de la lengua española: clases de palabras*. México: Gedisa/ UAM Iztapalapa.
- Narasihman, R. (1998). La cultura escrita: caracterización e implicaciones. En D. R. Olson, N. Torrance & A. Hildyard (Comps.). *Cultura escrita y oralidad* (pp. 237-261). Barcelona: Gedisa.
- Navarro, F. (2012). ¿Qué son los géneros profesionales? Apuntes teórico-metodológicos para el estudio del discurso profesional. En A. Cristófalo y J. Ledesma (Eds.). *Actas del IV Congreso Internacional de Letras “Transformaciones Culturales. Debates de la teoría, la crítica y la lingüística en el Bicentenario”* (Buenos Aires, 22-27 noviembre de 2010). Universidad de Buenos Aires. Recuperado de https://www.researchgate.net/publication/305637478_Que_son_los_generos_profesionales_Apuntes_teorico-metodologicos_para_el_estudio_del_discurso_profesional/related
- Navarro, F. (Coord.) (2014). *Manual de escritura para carreras de humanidades*. Buenos Aires: Editorial de la Facultad de Filosofía y Letras Universidad de Buenos Aires.
- Navarro, F. & Brown, A. (2014). Lectura y escritura de géneros académicos. Conceptos básicos. En F. Navarro (Coord.). *Manual de escritura para carreras de humanidades* (pp. 55-100). Buenos Aires: Editorial de la Facultad de Filosofía y Letras Universidad de Buenos Aires.

- Navarro, F. & Revel, A. (2013). *Escribir para aprender. Disciplinas y escritura en la escuela secundaria*. Buenos Aires: Paidós.
- Navarro, F., Ávila, N. & Gómez, G. (2019). Validez y Justicia: hacia una evaluación significativa en pruebas estandarizadas de escritura. *Avaliação, 11*(31), 1-35. Recuperado de <http://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/2045>
- Nippold, M. A. (2000). Language development during the adolescent years: aspects of pragmatics, syntax, and semantics. *Topics on Language Disorders, 20*(2), 15-28. Aspen Publishers. Recuperado de https://www.researchgate.net/publication/240595827_Language_Development_during_the_Adolescent_Years_Aspects_of_Pragmatics_Syntax_and_Semantics
- Nippold, M. A. (2007a). *Later language development. School-age children, adolescents, and young adults* (3a. ed.). Austin: Pro-Ed.
- Nippold, M. A. (2007b). School-Age Children Talk About Chess: Does Knowledge Drive Syntactic Complexity? *Journal of Speech Language and Hearing Research, 52*(4), 856-71. doi: 10.1044/1092-4388(2009/08-0094). Recuperado de https://www.researchgate.net/publication/24378788_School-Age_Children_Talk_About_Chess_Does_Knowledge_Drive_Syntactic_Complexity
- Nippold, M. A. (2010). Explaining complex matters: how knowledge of a domain drives language. En M. A. Nippold & C. M. Scott (Eds.) (2010). *Expository discourse in children, adolescents, and adults. Development and disorders* (41-61). Nueva York: Taylor and Francis.
- Nippold, M. A. & Scott, C. M. (Eds.) (2010). *Expository discourse in children, adolescents, and adults. Development and disorders*. Nueva York: Taylor and Francis.
- Nir, B. & Berman, R. (2010). Complex syntax as a window on contrastive rhetoric. *Journal of Pragmatics, 42*(3), 744-765. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0378216609001829>
- OCDE (2000). *Literacy in the in the Information Age. Final Report of the International Adult Literacy Survey*. París: Organisation for Economic Co-operation and Development.

- OCDE (10 de enero de 2019). *Presentación de los estudios de la OCDE “El Futuro de la Educación Superior en México: Promoviendo Calidad y Equidad” y “La Educación Superior en México: Resultados y Relevancia para el Mercado Laboral”*. México: OCDE. Recuperado de <http://www.oecd.org/about/secretary-general/estudios-de-la-ocde-sobre-educacion-superior-en-mexico-january-2019-sp.htm>
- Olive, T., & Levy, C. M. (Eds.). (2001). *Contemporary tools and techniques for studying writing*. Dordrecht: Kluwer Academic Publishers (Studies in Writing, vol. 10).
- O’Loughlin, K. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. *Language Testing*, 12(2), 217-237. doi.org/10.1177/026553229501200205. Recuperado de <http://journals.sagepub.com/doi/abs/10.1177/026553229501200205>
- Olson, D. R. (1985a). Introduction. En D. R. Olson, N. Torrance & A. Hildyard (Eds.). *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing* (pp. 1-18). Cambridge: Cambridge University Press.
- Olson, D. R. (1985b). *El mundo sobre papel. El impacto de la escritura y la lectura en la estructura del conocimiento*. Barcelona: Gedisa.
- Olson, D. R., Torrance, N. & Hildyard, A. (Comps.). (1998). *Cultura escrita y oralidad*. Barcelona: Gedisa.
- Ong, W. J. (1986). *The written world. Literacy in transition*. Oxford: Clarendon Press.
- Ong, W. J. (2016). *Oralidad y escritura. Tecnologías de la palabra*. México: FCE.
- Palancar, E. & Alarcón, L. J. (2007). Predicación secundaria depictiva en español. *Revista Española de Lingüística (RSEL)*, 37, 337-370.
- Parodi, G. (2007a). El discurso especializado escrito en el ámbito universitario y profesional: constitución de un corpus de estudio. *Revista signos*, 40(63), 147-178. Recuperado de <http://dx.doi.org/10.4067/S0718-09342007000100008>
- Parodi, G. (2007b). *Lingüística de corpus y discursos especializados: puntos de mira*. Valparaíso, Chile: Ediciones Universitarias de Valparaíso-Pontificia Universidad Católica de Chile.
- Parodi, G. (2008). Géneros del discurso escrito: Hacia una concepción integral desde una perspectiva sociocognitiva. En G. Parodi (Ed.). *Géneros académicos y géneros*

- profesionales: Accesos discursivos para saber y hacer* (pp. 17-37). Valparaíso, Chile: Ediciones Universitarias de Valparaíso-Pontificia Universidad Católica de Chile. Recuperado de http://www.euv.cl/archivos_pdf/generos.pdf
- Parodi, G. (2010). Multisemiosis y lingüística de corpus: Artefactos (multi)semióticos en los textos de seis disciplinas en el Corpus PUCV-2010. *Revista de Lingüística Teórica y Aplicada*, 48(2), 33-70. Recuperado de https://scielo.conicyt.cl/scielo.php?pid=S0718-09342014000200002&script=sci_arttext
- Parodi, G. & Burdiles, G. (Eds.). (2015). *Leer y escribir en contextos académicos y profesionales: géneros, corpus y métodos*. Santiago de Chile: Ariel.
- Pecorari, F. (2014). Anaphoric encapsulation and presupposition: Persuasive and stereotypical uses of a cohesive strategy. *Cadernos de Letras da UFF - Dossiê: Anáfora e correferência: temas, teorias e métodos*, 49, 175-195
- Perelman, L. C., Paradis, J. & Barret, E. (1998). *The Mayfield handbook of technical and scientific writing*. Mountain View, CA: Mayfield Publishing Company.
- Perry, K. (2012). What is Literacy? –A critical overview of sociocultural perspectives. *Journal of Language and Literacy Education [Online]*, 8(1), 50-71 Recuperado de http://jolle.coe.uga.edu/wp-content/uploads/2012/06/What-is-Literacy_KPerry.pdf
- Peterson, C. & McCabe, A. (1983). Developmental psycholinguistics: Three ways of looking at a child's narrative. *Journal of Child Language*, 13(2), 450-454. doi:10.1017/S0305000900008217
- Proyecto Tuning (2007). *Informe Final del Proyecto Tuning América Latina: Reflexiones y perspectivas de la Educación Superior en América Latina*. Bilbao: Universidad de Deusto. Recuperado de <http://tuning.unideusto.org/tuningal/index.php?>
- Ravid, D. (2006). Semantic development in textual contexts during the school years: Noun Scale analyses. *Journal of Child Language*, 33, 791-821. Recuperado de <http://www.cambridge.org/core/terms>. <http://dx.doi.org/10.1017/S0305000906007586>
- Real Academia Española (2014). *Diccionario de la lengua española*, 23.^a ed. [Online] Madrid: Espasa.
- Ribera, J. & Marín, M. J. (Coords.). (2018). Monogràfic “L'encapsulació lèxica: cohesió, coherència i metadiscurs”. *Caplletra*, 64, 105-252.

- Richardson, M. (2008, nov. 7). Writing is not just a basic skill. *The Chronicle of Higher Education*. Recuperado de <http://www.chronicle.com/article/Writing-Is-Not-Just-a-Basic/22146>
- Rijlaarsdam, G., van den Bergh, H. & M. Couzijn (Eds.) (2005). *Effective learning and teaching of writing. A handbook of writing in education*. Dordrecht: Kluwer Academic Publishers (Studies in Writing, vol. 14).
- Rivero, A. (2010). El aprendizaje del español en el nivel medio-superior: Diagnóstico y propuestas. *Revista de la Educación Superior*, 39(156), 35-51. Recuperado en 10 de octubre de 2018, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-27602010000400003&lng=es&tlng=es.
- Russell, D. (1991). *Writing in the academic disciplines, 1870-1990. A curricular history*. Carbondale, IL: Southern Illinois University Press. Recuperado de https://lib.dr.iastate.edu/engl_books/4
- Saita, L. & Zucker, J. D. (2013). Abstraction in different disciplines. En *Abstraction in artificial intelligence and complex systems* (pp. 11-47). Nueva York: Springer Science and Business Media. doi: 10.1007/978-1-4614-7052-6_2.
- Sanders, T. & Pander, H. (2006). Cohesion and coherence: Linguistic approaches. En *Encyclopedia of Language & Linguistics*. Elsevier. doi: 10.1016/B0-08-044854-2/00497-1 Recuperado de http://www.hum.uu.nl/medewerkers/h.l.w.pandermaat/coherence_doc/cohesionencoherence2006.pdf
- Schleppegrell, M. J. (2001). Linguistic Features of the Language of Schooling. *Linguistics and Education*, 12(4): 431-459. Recuperado de http://blogs.ubc.ca/lled3602015/files/2015/08/Schleppegrell_Linguistic_Features.pdf
- Schleppegrell, M. J. (2004). *The language of schooling. A functional linguistics perspective*. Nueva Jersey: Lawrence Erlbaum.
- Schleppegrell, M. J. (2009). Language in academic subject areas and classroom instruction: what is academic language and how can we teach it. En *Workshop on The role of language in school learning sponsored by The National Academy of Sciences, Menlo Park, CA*. Recuperado de <https://www.mydigitalchalkboard.org/cognoti/content/file/resources/documents/98/98c3e7f4/98c3e7f49b44eaa5ee60b45939df619b4593afc7/Schleppegrell.pdf>

- Schmid, H.-J. (1997). Constant and ephemeral hypostatization: *Thing, Problem* and other “shell nouns”. En B. Caron (Ed.). *Proceedings of the 16th. International Congress of Linguists*. París: Elsevier.
- Schmid, H.-J. (1999a). Cognitive effects of shell nouns. En K. van Hoek, A. Kibrik & L. Noordman (Eds.). *Discourse Studies in Cognitive Linguistics* (pp. 111-132). Berlín: Mouton de Gruyter.
- Schmid, H.-J. (1999b). Towards a functional-cognitive lexicology of nouns. En W. Falkner & H.-J. Schmid (Eds.). *Words, lexemes, concepts, approaches to the lexicon: Studies in honour of Leonhard Lipka* (pp. 213-226). Tubinga: G. Narr.
- Schmid, H.-J. (2000). *English abstract nouns as conceptual shells: From corpus to cognition*. Berlín: Walter de Gruyter.
- Schmid, H.-J. (2018). Shell nouns in english. A personal roundup. *Caplletra*, 64, 109-128. doi: 10.7203/Caplletra.64.11368
- Scott, C. M. (2010). Assessing expository texts produced by school-age children and adolescents. En M. A. Nippold, & C. M. Scott (Eds.) (2010). *Expository discourse in children, adolescents, and adults. Development and disorders* (pp. 191-214). Nueva York: Taylor and Francis.
- SEP (2016). *Documento base del bachillerato general*. México: Secretaría de Educación Pública. Recuperado de http://www.dgb.sep.gob.mx/informacion-academica/programas-de-estudio/documentobase/DOC_BASE_16_05_2016.pdf
- SEP (2017a). *Planes de estudio de referencia del componente básico del marco curricular común de la Educación Media Superior*. México: Secretaría de Educación Pública. Recuperado de <https://www.gob.mx/cms/uploads/attachment/file/241519/planes-estudio-sems.pdf>
- SEP (2017b). *Publicación de Resultados 2016. Plan Nacional para la Evaluación de los Aprendizajes (PLANEA EMS)*. México: SEP-Subsecretaría de Planeación, Evaluación y Coordinación. Recuperado de <http://planea.sep.gob.mx/ms/>
- SEP (2018). Subsecretaría de Educación Media Superior [sitio web]. México: SEP. Disponible en <http://www.sems.gob.mx>

- Shrout, P. E. y Fleiss, J. L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, 86(2), 420-428. Recuperado de http://www.aliquote.org/cours/2012_biomed/biblio/Shrout1979.pdf
- Sinclair, J. (2004). *Trust the text. Language, corpus and discourse*. Nueva York: Routledge.
- Snow, C. E. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review*, 53, 165-189.
- 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. Recuperado de <http://nrs.harvard.edu/urn-3:HUL.InstRepos:11654980>
- Snyder, L. & Caccamise, D. (2010). Comprehension processes for expository text: building meaning and making sense. En M. A. Nippold & C. M. Scott (Eds.) (2010). *Expository discourse in children, adolescents, and adults. Development and disorders* (13-39). Nueva York: Taylor and Francis.
- Street, B. V. (1984). *Literacy in theory and practice*. Cambridge: Cambridge University Press.
- Street, B. V. (1995). *Social literacies. Critical approaches to Literacy in Development, Ethnography and Education*. Londres: Longman.
- Street, B. V. (2005). *Understanding and defining literacy*. Paper commissioned for the EFA Global Monitoring Report 2006, Literacy for Life. UNESCO. Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000146186>
- Street, B. V. (2015). Academic Writing: theory and practice. En *International Conference: "Literacy in Contemporary Society: Compiling social, political, philosophical, pedagogical and technological aspects"*. Nicosia: Pedagogical Institute of Cyprus. Recuperado de http://www.pi.ac.cy/pi/files/epimorfosi/synedria/literacy/Brian_Street_Academic_Cyprus_Conf_presentation.pdf

- Strömquist, S., Johansson, V., Kriz, S., Ragnarsdóttir, H., Aisenman, R. & Ravid, D. (2002). Toward a cross-linguistic comparison of lexical quantain speech and writing. *Written Language & Literacy*, 5(1), 45-67, doi 10.1075/wll.5.1.03str. Recuperado de https://www.researchgate.net/publication/233624481_Toward_a_cross-linguistic_comparison_of_lexical_quantain_speech_and_writing
- Swales, J. (1990). *Genre Analysis: English in academic and research settings*. Nueva York: Cambridge University Press.
- Swales, J. (2013). *Research genres. Explorations and applications*. Nueva York: Cambridge University Press.
- Taki, S. & Jafarpour, F. (2012). Engagement and Stance in Academic Writing: A Study of English and Persian Research Articles. *Mediterranean Journal of Social Sciences*, 3(1), 157-168. Recuperado de https://www.academia.edu/3105269/Engagement_and_Stance_in_Academic_Writing_A_Study_of_English_and_Persian_Research_Articles
- Tallerman, M. (2011). *Understanding syntax*. Londres: Hodder Education.
- Tárnyiková, J. (2018). Constructions with shell nouns in English: their dual role in information packaging. *Caplletra*, 64, 205-225. doi: 10.7203/Caplletra.64.11381
- Teberosky, A. (2007). El texto académico. En M. Castelló (Coord.). *Escribir y comunicarse en contextos científicos y académicos: conocimientos y estrategias* (pp. 17-46). Barcelona: Graó.
- Torner, S. & Battaner, P. (2005). *El corpus PAAU 1992: estudios descriptivos, textos y vocabulario*. [e-book] Barcelona: IULA.
- Uccelli, P., Dobbs, C. L. & Scott, J. (2012). Mastering Academic Language: Organization and Stance in the Persuasive Writing of High School Students. *Written Communication*, 30(1), 36-62. doi:10.1177/0741088312469013
- Uccelli, P., Barr, C. D., Dobbs, C. L., Phillips Galloway, E., Meneses, A., & Sánchez, E. (2014). Core Academic Language Skills (CALs): An expanded operational construct

- and a novel instrument to chart school-relevant language proficiency in per-adolescent and adolescent learners. *Applied Psycholinguistics*, 5, 1075-1107. doi: 10.1017/S0142716400006X.
- Uccelli, P., Phillips Galloway, E., Barr, C., Meneses, A., & Dobbs, C. (2015). Beyond vocabulary: Exploring cross-disciplinary academic language proficiency and its association with reading comprehension. *Reading Research Quarterly*, 50(3). doi:10.1002/rrq.104
- UNESCO (2005). *Aspects of Literacy Assessment: Topics and issues from the UNESCO Expert Meeting, 10-12 June, 2003*. París: UNESCO.
- UNESCO (2008). *The Global literacy challenge: a profile of youth and adult literacy at the mid-point of the United Nations Decade 2003-2012*. (M. Richmond, C. Robinson & M. Sachs-Israel, Eds.). París: UNESCO-Education Sector. Division for the Coordination of United Nations Priorities in Education.
- UNESCO (2014). *Alfabetización: una ruta de aprendizaje multimodal para toda la vida. Consideraciones sobre las prácticas de lectura y escritura para el ejercicio ciudadano en un contexto global e intercomunicado*. Bogotá: CERLAC-UNESCO/ Universidad de Antioquia.
- Ure, J. (1971). Lexical density and register differentiation. En G. Perren & J. Trim (Eds.). *Applications of linguistics* (pp. 443-452). Cambridge: Cambridge University Press.
- Van Hoek, K. (1997). *Anaphora and conceptual structure*. Chicago: University of Chicago Press.
- Witte, S. P. & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication*, 32(2), 189-204. Language Studies and Composing. Recuperado de <http://www.jstor.org/stable/356693>
- Zamudio, C. (2016). *Evaluación del Corpus EXCALE de escritura*. México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/C/153/P1C153.pdf>

- Zamudio, V. (2016). La expresión de opiniones y puntos de vista en textos académicos estudiantiles sobre literatura. *Lenguaje*, 44(1), 35-59. Recuperado de <http://www.scielo.org.co/pdf/leng/v44n1/v44n1a03.pdf>
- Zavala, V. (2009). La literacidad o lo que la gente *hace* con la lectura y la escritura. En D. Cassany, (Comp.). *Para ser letrados. Voces y miradas sobre la lectura* (pp. 23-35). Barcelona: Paidós.
- Zwiers, J. (2008). *Building academic language. Essential practices for content classrooms, grades 5-12*. San Francisco, CA: Jossey-Bass/IRA.

10. ANEXOS

A. Resumen de los Antecedentes de investigación

Tabla 19. *Aportes conceptuales y metodológicos de los estudios previos*

Estudio	Antecedente teórico-conceptual	Antecedente metodológico	Observaciones
Torner y Battaner (2002)	<i>Competencia en lengua escrita</i>	Análisis de corpus escrito por aspirantes a ingresar la universidad. Selección de una muestra de producciones de estudiantes de bachillerato privado con altos estándares. Uso de herramientas de análisis léxico y descripción manual.	Utilidad de usar muestras provenientes de bachilleratos de alta exigencia para que se evidencien las características lingüísticas que se desea describir. Acotación de los hallazgos del análisis a la situación de producción del texto y al género.
Berman y Nir (2009)	<i>Calidad global del texto</i>	Análisis de corpus de textos evaluados por jueces. Mediciones de la extensión de las palabras, la densidad léxica, el registro y la abstracción nominal, en correlación con evaluación de jueces.	Influencia de la producción (en cantidad de palabras) sobre la valoración de los jueces. Necesidad de que los jueces sean distintos a los investigadores.
Morales (2004)	<i>Calidad del texto</i>	<i>No aplica.</i>	Necesidad de diseñar pautas holísticas para evaluar la escritura.
Atorresi (2005)	<i>Calidad del texto</i>	<i>No aplica.</i>	Necesidad de orientar a los docentes sin formación o entrenamiento lingüístico para redactar consignas dirigidas a evaluar la escritura.
Martínez (2010)	<i>Calidad de la escritura</i>	<i>No aplica.</i>	Necesidad de incorporar los avances de las ciencias cognitivas al diseño de pautas de evaluación.

McNamara, Crossley y McCarthy (2009)	<i>Calidad del texto</i>	Análisis de corpus de textos calificados por expertos con base en una rúbrica. Mediciones de cohesión, complejidad sintáctica, diversidad de palabras, frecuencia, concreción e imaginabilidad. Uso de modelos de regresión para predecir indicadores eficientes del constructo.	Mayor poder predictivo de los indicadores léxicos respecto de los indicadores de cohesión textual, en particular, palabras de baja frecuencia, menos concretas e imaginables.
Crossley y McNamara (2011)	<i>Competencia de escritura</i>	Análisis de corpus de textos evaluados por estudiantes universitarios entrenados. Mediciones de indicadores de cohesión y de diversidad léxica y frecuencia de las palabras.	Mayor poder predictivo de los indicadores léxicos respecto de los indicadores de cohesión textual, en particular, léxico de baja frecuencia y menor grado de concreción.
Uccelli, Dobbs y Scott (2012)	<i>Calidad de la escritura</i>	Análisis de corpus de textos calificados por profesores. Mediciones de indicadores léxicos y gramaticales. Uso de modelos de regresión para predecir indicadores eficientes del constructo.	Relevancia de la diversidad léxica como indicador de calidad, junto con aspectos sintácticos y discursivos.
Uccelli, Barr, Dobbs, Phillips Galloway, Meneses y Sánchez (2014)	<i>Competencia en lenguaje académico</i>	Construcción de un aparato conceptual e instrumental llamado CALS para evaluar la competencia.	Implicaciones cognitivas y pedagógicas de la adquisición de habilidades del constructo CALS.
Uccelli, Phillips Galloway, Barr, Meneses y Dobbs (2015)	<i>Competencia en lenguaje académico</i>	Ampliación del constructo CALS.	Implicaciones cognitivas y pedagógicas de la adquisición de habilidades del constructo CALS.

Meneses, Uccelli, Santelices, Ruiz, Acevedo y Figueroa (2017); Figueroa (2018), y Figueroa, Meneses y Chandía (2018) Navarro, Ávila y Gómez (2019)	<i>Competencia en lenguaje académico</i> <i>Competencia de escritura</i>	Adaptación al español del constructo CALS: Habilidades Clave de Lenguaje Académico (HCLA) y diseño de instrumentos para evaluar el constructo. <i>No aplica.</i>	Capacidad predictiva del dominio de recursos léxico-gramaticales como el manejo de vocabulario académico. Necesidad de diseñar pautas holísticas para evaluar la escritura.
Torner y Battaner (2002)	<i>Vocabulario conocido y usado en producciones de estudiantes</i>	Análisis del vocabulario disciplinar en un corpus de escritura de estudiantes al final del bachillerato; uso de términos científicos y sustantivos abstractos.	Acotación de los hallazgos en el uso de vocabulario a la situación de producción del texto y al género. No generalización a otros contextos.
Strömqvist, Johansson, Kriz, Ragnarsdóttir, Aisenman y Ravid (2002)	<i>Recursos léxicos de la escritura</i>	Análisis de producciones orales y escritas. Mediciones de diversidad, densidad léxica y extensión de las palabras, con el programa CLAN.	Uso diferenciado de recursos léxicos en las tareas escritas respecto de las tareas orales.
Chung y Nation (2004)	<i>Vocabulario técnico</i>	Análisis contrastivo de cuatro métodos de identificación de términos técnicos, sobre dos corpus especializados: uso de escalas, diccionarios técnicos, pistas en el texto y frecuencia en corpus.	Mayor precisión de las escalas de valoración y los diccionarios como método de identificación de vocabulario técnico. Presencia de 20% a 30% de este léxico en corpus disciplinares.
Ravid (2006)	<i>Escala nominal</i>	Análisis de textos orales y escritos de dos géneros. Puntuación de sustantivos con una escala nominal de diez niveles basada en Lyons (1977).	Mayor densidad y abstracción nominal en textos escritos respecto de los orales.

Berman y Nir (2010)	<i>Escala nominal</i>	Análisis de textos orales y escritos de dos géneros. Mediciones de tamaño de la palabra, registro, densidad léxica, diversidad léxica y abstracción de los sustantivos (con base en Ravid, 2006).	Mayor abstracción de los sustantivos y complejidad de la escritura de adultos con mayor escolaridad.
Berman, Nayditz y Ravid (2011)	<i>Calidad léxica en la escritura</i>	Análisis de textos escritos de dos niveles educativos con diferentes historiales académicos. Mediciones de léxico y sintaxis: extensión en cláusulas, cantidad de palabras, densidad sintáctica, diversidad y densidad léxica, en relación con un diagnóstico del texto.	Relevancia de indicadores como la densidad léxica, los sustantivos abstractos y nominalizaciones en la valoración de los textos.
Hess (2013)	<i>Escala nominal y lenguaje académico</i>	Análisis de textos orales y escritos agrupados por el acceso previo a la cultura escrita. Análisis de contenido semántico-pragmático de los sustantivos con base en la escala nominal de Ravid (2006).	Mayor uso de sustantivos abstractos y de lenguaje académico literario en textos escritos de adolescentes con mayor inmersión en la cultura escrita.
Gómez, Sotomayor, Bedwel, Domínguez y Jéldrez (2016)	<i>Calidad léxica en la escritura</i>	Análisis de textos escritos de tres géneros, evaluados con base en una rúbrica. Mediciones de diversidad léxica, frecuencia y densidad léxica. Modelos de regresión respecto de la calificación de los textos.	Influencia diferenciada de los indicadores léxicos en la calificación, en relación con los distintos géneros.

Crossley, Salsbury, McNamara y Jarvis (2010)	<i>Dominio léxico</i>	Análisis de textos escritos evaluados por jueces entrenados con base en una rúbrica. Mediciones de tamaño de las palabras, diversidad léxica, frecuencia, concreción, etc., con Coh-Metrix. Modelos de regresión respecto de la calificación.	Mayor poder productivo de la diversidad léxica, la hiperonimia y la frecuencia de las palabras.
Crossley, Salsbury y McNamara (2011)	<i>Competencia léxica</i>	Análisis de textos escritos por estudiantes de segunda lengua, puntuados con la prueba TOEFL. Mediciones de indicadores léxicos con medios automatizados.	Mayor poder predictivo del constructo de la imaginabilidad, la frecuencia, la diversidad léxica y la familiaridad de las palabras en textos de estudiantes de segunda lengua.

B. Indicadores de literacidad según organismos internacionales y nacionales

Para contar con referentes adecuados acerca de lo que se entiende por *literacidad*, enseguida se presentan una serie de definiciones avaladas por instituciones dedicadas al desarrollo social, económico y educativo, además de los estándares esperados en materia de lenguaje en el sistema educativo mexicano para los egresados del nivel medio superior.

UNESCO: “Literacy is the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve her or his goals, to develop her or his knowledge and potential, and to participate fully in her or his and wider society” (Montoya, 2018; UNESCO, 2005).

European Literacy Policy Network: en la Declaración Europea del Derecho a la Literacidad (*European Declaration of the Right to Literacy*), esta red indica que la literacidad refiere a la habilidad para leer y escribir en un nivel en el cual los individuos

puedan usar y entender efectivamente la comunicación escrita en todos los medios (impresos y electrónicos), incluyendo la literacidad digital (Montoya, 2018).

OCDE: la Organización para la Cooperación y el Desarrollo Económico, en la operación de un programa de seguimiento a las habilidades de adultos, define este mismo concepto como entender, evaluar, usar e involucrarse con textos para participar en la sociedad, para lograr objetivos propios y desarrollar el propio conocimiento y potencial (Montoya, 2018).

Banco Mundial: en su programa Skills Towards Employability and Productivity (STEP), este organismo financiero agrega la definición de habilidades cognitivas, que comprenden la capacidad de entender ideas complejas, de adaptarse efectivamente al medio, de aprender de la experiencia e involucrarse en varias formas de razonamiento para superar obstáculos actuando cognitivamente sobre ellos (Montoya, 2018).

Esta diversidad de definiciones puede ser problemática porque, según el mismo documento de trabajo de la UNESCO (Montoya, 2018), aunque este organismo mundial propone que la evaluación del logro de la literacidad cumpla algunos criterios, como considerar un rango amplio de habilidades que se puedan ubicar en un continuo, existen diferencias en la manera de operacionalizar las capacidades que quedan incluidas al interior del concepto, según las distintas organizaciones internacionales.

Así, en algunos programas de evaluación de distintos países la literacidad implica “poder leer una parte de una oración”, “poder leer y escribir”, “no tener dificultades para leer y escribir”, en términos muy amplios; o en términos más restringidos como: “leer, entender y escribir una historia corta sobre una parte rutinaria de la propia vida”, “escribir una carta personal o leer el periódico”. También difieren los enfoques en cuanto a la lengua en que se debe ser competente, si debe ser solo en la propia o en otras lenguas, en variedades dialectales e incluso entre sistemas de escritura o alfabetos (Montoya, 2018).

Para el caso de México,⁴⁴ resulta pertinente hacer una revisión de los aprendizajes que se esperan al término del bachillerato en materia de lenguaje. En el Acuerdo Secretarial 444 se establecen las competencias que constituyen el Marco Curricular Común (MCC) del Sistema Nacional del Bachillerato (SNB); estas se dividen en competencias genéricas, competencias disciplinares básicas y, según la modalidad y el subsistema, también en disciplinares extendidas y profesionales (SEP, 2018). A continuación se cita el desglose de la competencia genérica “Se expresa y comunica”, del MCC de bachillerato.

Competencia: Escucha, interpreta y emite mensajes pertinentes en distintos contextos mediante la utilización de medios, códigos y herramientas apropiados.

Atributos:

Expresa ideas y conceptos mediante representaciones lingüísticas, matemáticas o gráficas.

Aplica distintas estrategias comunicativas según quienes sean sus interlocutores, el contexto en el que se encuentra y los objetivos que persigue.

Identifica las ideas clave en un texto o discurso oral e infiere conclusiones a partir de ellas.

Se comunica en una segunda lengua en situaciones cotidianas.

Maneja las tecnologías de la información y la comunicación para obtener información y expresar ideas (SEP, 2016, p. 56).

En cuanto a las competencias disciplinares básicas de comunicación, de acuerdo al *Documento base del bachillerato general* (SEP, 2016), estas están referidas a la capacidad de los estudiantes de comunicarse efectivamente en español y en lo esencial en una segunda lengua en diversos contextos, mediante el uso de distintos medios e instrumentos.

⁴⁴ El sistema educativo mexicano comprende tres niveles: la educación básica (preescolar, primaria y secundaria), la educación media superior o bachillerato (EMS) y la educación superior (ingenierías, licenciaturas y posgrados). El bachillerato tiene varias modalidades: escolarizado, no escolarizado, mixto, certificación por examen, capacitación para el trabajo y una dirigida a personas con discapacidad. La modalidad escolarizada tiene varios subsistemas adscritos al bachillerato tecnológico (con varios centros y colegios bajo su coordinación) o al bachillerato general, coordinado por la Dirección General de Bachilleratos (DGB), incluyendo el Colegio de Bachilleres (COLBACH). Las escuelas de este nivel pueden ser de sostenimiento público o privado (SEP, 2018).

Los estudiantes que hayan desarrollado estas competencias podrán leer críticamente y comunicar y argumentar ideas de manera efectiva y con claridad oralmente y por escrito. Además, usarán las tecnologías de la información y la comunicación de manera crítica para diversos propósitos comunicativos. (...) Las competencias de comunicación están orientadas además a la reflexión sobre la naturaleza del lenguaje y a su uso como herramienta del pensamiento lógico (SEP, 2016, pp. 25-26).

Finalmente, las competencias disciplinares básicas de comunicación consideradas en el MCC son las siguientes:

- 1) Identifica, ordena e interpreta las ideas, datos y conceptos explícitos e implícitos en un texto, considerando el contexto en el que se generó y en el que se recibe.
- 2) Evalúa un texto mediante la comparación de su contenido con el de otros, en función de sus conocimientos previos y nuevos.
- 3) Plantea supuestos sobre los fenómenos naturales y culturales de su entorno con base en la consulta de diversas fuentes.
- 4) Produce textos con base en el uso normativo de la lengua, considerando la intención y situación comunicativa.
- 5) Expresa ideas y conceptos en composiciones coherentes y creativas, con introducciones, desarrollo y conclusiones claras.
- 6) Argumenta un punto de vista en público de manera precisa, coherente y creativa.
- 7) Valora y describe el papel del arte, la literatura y los medios de comunicación en la recreación o la transformación de una cultura, teniendo en cuenta los propósitos comunicativos de distintos géneros.
- 8) Valora el pensamiento lógico en el proceso comunicativo en su vida cotidiana y académica.
- 9) Analiza y compara el origen, desarrollo y diversidad de los sistemas y medios de comunicación.
- 10) Identifica e interpreta la idea general y posible desarrollo de un mensaje oral o escrito en una segunda lengua, recurriendo a conocimientos previos, elementos no verbales y contexto cultural.

11) Se comunica en una lengua extranjera mediante un discurso lógico, oral o escrito, congruente con la situación comunicativa.

12) Utiliza las tecnologías de la información y comunicación para investigar, resolver problemas, producir materiales y transmitir información. (SEP, 2016, p. 60).

Como se vio en este desglose de definiciones y parámetros, son muy diversas las maneras de conceptualizar el alcance del término *literacidad*; por ello, en esta investigación se toma como referente el concepto extendido propuesto por la UNESCO, y se incluyen los fines de participación ciudadana considerados por la OCDE. En ambos organismos participa México, y por ello sus estándares en las competencias de lenguaje y comunicación abarcan las diversas prácticas letradas que son indispensables para acceder al conocimiento en el ámbito académico, lo mismo que para participar en la comunidad, aunque la definición operante en esta tesis debe restringirse a las prácticas de literacidad escolar.

C. Pauta de escritura de los textos

Elementos Ponderación general	Criterios a evaluar Ponderación específica	✓	Evaluación
TIPO TEXTUAL	Es un ensayo con estructura IDC con tesis explícita y argumentación lógicamente organizada. No es un relato, un resumen o una reseña. (En esos casos el trabajo se anula con calificación de 10/100).		
ORTOGRAFÍA	Respetar las reglas de acentuación y de uso de las grafías a lo largo de todo su texto. Acostumbra usar diccionarios para corregir tus escritos. (Se restará una décima por cada 3 errores, aunque la palabra se repita. 10 errores anulan el ensayo con calificación de 30/100).		
ESTRUCTURA Y CONTENIDO 40%	<u>Título</u> : es atractivo y referente al tema.		
	<u>Introducción</u> : presenta el tema del ensayo, el contexto de las obras y el propósito del ensayo comparativo.		
	<u>Desarrollo</u> : Se organiza como una secuencia comparativa a partir de tres variables que deberán estar definidas; la comparación se sustenta con tres citas directas de cada obra comparada.		
	<u>Conclusión</u> : Recupera brevemente la argumentación y los hallazgos del análisis.		
REDACCIÓN 40%	<u>Estructura de párrafos</u> : Presenta una organización en párrafos con oraciones alrededor de una idea central.		
	<u>Precisión y variedad de vocabulario</u> : Usa sinónimos y antónimos para evitar repeticiones y cacofonías; emplea el sentido denotativo preciso de los términos; usa un lenguaje académico adecuado al propósito del texto. No usa ambigüedades como "cosa", "algo", etc.		
	<u>Cohesión y coherencia</u> : emplea conectores para relacionar lógicamente las ideas; usa anáforas y pronombres para reiterar el referente; hace uso de la puntuación para delimitar segmentos de significado.		
	<u>Personalización</u> : usa lenguaje referencial, objetivo e impersonal. No usa la primera persona.		
FORMATO 10%	<u>Bibliografía, citas y referencias</u> : Usa el formato APA en citas, referencias y bibliografía.		
PRESENTACIÓN 10%	Presenta un documento limpio, legible, ordenado y adecuadamente distribuido en la hoja, con márgenes y sangrías.		

Figura 30. Pauta de escritura de los textos.

Fuente: Material de la clase de Español y literatura (2013).

D. Rúbrica original y versiones adaptadas

SCORE OF 6: Demonstrates clear and consistent mastery, although it may have a few minor errors. Effectively and insightfully develops a point of view on the issue and demonstrates outstanding critical thinking, using clearly appropriate examples, reasons, and other evidence to support its position is well organized and clearly focused, demonstrating clear coherence and smooth progression of ideas exhibits skillful use of language, using a varied, accurate, and apt vocabulary, meaningful variety in sentence structure, free of most errors in grammar, usage, and mechanics.

SCORE OF 5: Demonstrates reasonably consistent mastery, although it will have occasional errors or lapses in quality. A typical essay effectively develops a point of view on the issue and demonstrates strong critical thinking, generally using appropriate examples, reasons, and other evidence to support its position is well organized and focused, demonstrating coherence and progression of ideas exhibits facility in the use of language, using appropriate vocabulary, variety in sentence structure, generally free of most errors in grammar, usage, and mechanics.

SCORE OF 4: Demonstrates adequate mastery, although it will have lapses in quality. A typical essay develops a point of view on the issue and demonstrates competent critical thinking, using adequate examples, reasons, and other evidence to support its position is generally organized and focused, demonstrating some coherence and progression of ideas exhibits adequate but inconsistent facility in the use of language, using generally appropriate vocabulary demonstrates some variety in sentence structure has some errors in grammar, usage, and mechanics.

SCORE OF 3: Demonstrates developing mastery and is marked by ONE OR MORE of the following weaknesses: develops a point of view on the issue, demonstrating some critical thinking, but may do so inconsistently or use inadequate examples, reasons, or other evidence to support its position is limited in its organization or focus, or may demonstrate some lapses in coherence or progression of ideas displays developing facility in the use of language but sometimes uses weak vocabulary or inappropriate word choice lacks variety or demonstrates problems in sentence structure contains an accumulation of errors in grammar, usage, and mechanics.

SCORE OF 2: Demonstrates little mastery, and is flawed by ONE OR MORE of the following weaknesses: develops a point of view on the issue that is vague or seriously limited, and weak critical thinking, providing inappropriate or insufficient examples, reasons, or other evidence to support its position is poorly organized and/or focused or demonstrates serious problems with coherence or progression of ideas displays very little facility in the use of language, using very limited vocabulary or incorrect word choice demonstrates frequent problems in sentence structure contains errors in grammar, usage, and mechanics so serious that meaning is somewhat obscured.

SCORE OF 1: Demonstrates very little or no mastery and is severely flawed by ONE OR MORE of the following weaknesses: develops no viable point of view on the issue, or provides little or no evidence to support its position is disorganized or unfocused, resulting in a disjointed or incoherent essay displays fundamental errors in vocabulary demonstrates severe flaws in sentence structure contains pervasive errors in grammar, usage, or mechanics that persistently interfere with meaning. (McNamara, Crossley y McCarthy, 2009, pp. 77-78.)

Figura 31. Rúbrica original que se adaptó para esta investigación.

Fuente: McNamara et al. (2009).

Puntuación de 6. Demuestra un **dominio claro y consistente de la escritura académica**, aunque puede tener algunos errores menores. **Desarrolla efectivamente un punto de vista sobre el tema** y demuestra pensamiento crítico, utilizando ejemplos apropiados, razones y otras evidencias para apoyar su posición. El texto está bien organizado y claramente enfocado al tema, demostrando coherencia y una clara progresión de las ideas. Exhibe un **uso hábil del lenguaje**, utilizando un vocabulario variado, preciso y apropiado, una variedad significativa en la estructura de las oraciones, libre de la mayoría de errores gramaticales.

Puntuación de 5. Demuestra un **dominio razonablemente consistente de la escritura académica**, aunque puede tener errores ocasionales o fallas menores de calidad. **Desarrolla un punto de vista sobre el tema** y demuestra pensamiento crítico, usando ejemplos apropiados, razones y otras evidencias para apoyar su posición. El texto está bien organizado y enfocado al tema, demostrando coherencia y progresión de las ideas. Hace **buen uso del lenguaje**, utilizando un vocabulario apropiado; muestra variedad en la estructura de oraciones; está libre de la mayoría de errores gramaticales.

Puntuación de 4. Demuestra un **dominio adecuado en la escritura académica**, aunque puede tener fallas de calidad. **Desarrolla un punto de vista sobre el tema** y demuestra pensamiento crítico, aunque no suficientemente riguroso; utiliza algunos ejemplos, razones y otras evidencias para apoyar su posición. El texto está organizado y enfocado al tema, demostrando cierta coherencia y progresión de las ideas. Exhibe **facilidad adecuada en el uso del lenguaje pero puede tener pasajes incoherentes**. Usa un vocabulario generalmente apropiado; demuestra cierta variedad en la estructura de oraciones, y tiene algunos errores gramaticales.

Puntuación de 3. Demuestra un **desarrollo en el dominio de la escritura académica**, con una o más de las siguientes debilidades: **desarrolla un punto de vista sobre el tema**, demostrando algún pensamiento crítico, pero puede hacerlo de manera inconsistente o usar ejemplos inadecuados, razones u otra evidencia para apoyar su posición. **La organización del texto es limitada o pierde el enfoque en el tema**, o puede demostrar algunas fallas en la coherencia o la progresión de las ideas. Muestra facilidad en el uso del lenguaje, pero a veces presenta una **elección de palabras inapropiada**. Carece de variedad en la estructura de las oraciones o demuestra problemas en su estructuración, y acumula errores gramaticales.

Puntuación de 2. Demuestra **poco dominio en la escritura académica** y es deficiente por una o más de las siguientes debilidades: **desarrolla un punto de vista vago o seriamente limitado sobre el tema**, muestra un pensamiento crítico débil, proporcionando ejemplos inapropiados o insuficientes razones u otras evidencias para apoyar su posición. **El texto está mal organizado y/o enfocado en el tema**. Muestra serios problemas con la coherencia o la progresión de las ideas que afectan el significado. Exhibe **muy poca facilidad en el uso del lenguaje**, empleando un vocabulario muy limitado o la elección de palabras incorrectas. Demuestra problemas frecuentes en la estructura de las oraciones y contiene errores en la gramática.

Puntuación de 1. Demuestra **poco o ningún dominio en la escritura académica**, cuya calidad se ve severamente afectada por una o más de las siguientes debilidades: **no desarrolla ningún punto de vista viable sobre el tema** o proporciona poca o ninguna evidencia para apoyar su posición. **El texto está desorganizado o desenfocado**, resultando en un escrito inconexo o incoherente. Presenta **errores importantes en la elección y el uso del vocabulario**. Demuestra severas fallas en la estructura de las oraciones y contiene errores gramaticales que persistentemente interfieren con el significado.

Figura 32. Primera adaptación de la rúbrica.

Fuente: Elaboración propia con base en McNamara et al. (2009).

Puntuación de 6. Demuestra un **dominio notablemente consistente de la escritura académica**, aunque puede tener errores menores. El texto se **estructura** a partir de una introducción en que presenta el propósito del escrito y plantea de manera explícita un punto de vista, que es desarrollado y sustentado utilizando ejemplos apropiados, razones y otras evidencias para apoyar su posición, misma que recupera en la conclusión. Muestra una **organización claramente enfocada al desarrollo del tema**, demostrando coherencia y una eficiente progresión de las ideas. Exhibe un **uso hábil del lenguaje**, utilizando un vocabulario variado, preciso y adecuado al propósito del texto. Demuestra una habilidad gramatical notable que se muestra en la estructuración y conexión de las oraciones, misma que favorece el significado.

Puntuación de 5. Demuestra un **dominio razonablemente consistente de la escritura académica**, aunque puede tener errores o fallas menores. El texto se **estructura** a partir de una introducción en que presenta el propósito del escrito y plantea de manera explícita un punto de vista, que es desarrollado y sustentado utilizando ejemplos apropiados, razones y otras evidencias para apoyar su posición, misma que recupera en la conclusión. Muestra una **organización claramente enfocada al desarrollo del tema**, demostrando coherencia y una clara progresión de las ideas, aunque puede tener alguna falla en un pasaje delimitado. Exhibe un **uso apropiado del lenguaje**, utilizando un vocabulario variado y adecuado al propósito del texto. Demuestra un buen nivel de habilidad gramatical visible en la estructuración y conexión de las oraciones, aunque puede tener algunos errores que no afectan mayormente el significado.

Puntuación de 4. Demuestra un **dominio adecuado de la escritura académica**, aunque presenta algunas fallas que afectan la calidad. El texto se **estructura** a partir de una introducción en que presenta el propósito del escrito y plantea de manera explícita un punto de vista, que es desarrollado y sustentado utilizando ejemplos apropiados, razones y otras evidencias para apoyar su posición, misma que recupera en la conclusión. Muestra una **organización enfocada al desarrollo del tema**, demostrando cierta coherencia y una progresión suficientemente fluida de las ideas. Exhibe **facilidad en el uso del lenguaje** pero puede tener fallas en la precisión del vocabulario. Demuestra un nivel aceptable de habilidad gramatical, aunque comete errores en la estructuración y conexión de las oraciones que llegan a dificultar el significado.

Puntuación de 3. Demuestra un **desarrollo en el dominio de la escritura académica**, con una o más de las siguientes debilidades: El texto se **estructura** a partir de una introducción en que se presenta el propósito del escrito y se expresa una opinión sobre el tema general, aunque éste es desarrollado de manera inconsistente o usando ejemplos, razones o evidencias inadecuados; hay una conclusión pero no recupera explícitamente el punto de vista antes presentado. El texto muestra una **organización limitada** o pierde el enfoque en el tema, mostrando algunas fallas en la coherencia o un interrupciones en la progresión de las ideas. Exhibe **cierta facilidad en el uso del lenguaje**, pero puede presentar una elección inapropiada de palabras. Demuestra un nivel bajo de habilidad gramatical, con errores en la estructuración y conexión de las oraciones que afectan notoriamente el significado.

Puntuación de 2. Demuestra **poco dominio en la escritura académica**, que es deficiente por una o más de las siguientes debilidades: El texto se **estructura** a partir de una introducción pero o bien no se presenta el propósito del escrito o el punto de vista sobre el tema es vago o impreciso, el cual es desarrollado proporcionando ejemplos inapropiados o insuficientes razones u otras evidencias para apoyar su posición, que puede no ser recuperada en una conclusión claramente definida. El texto está **desorganizado** o no enfocado en el tema. Muestra serios problemas con la coherencia o la progresión de las ideas, habiendo incluso pasajes incoherentes. Exhibe **muy poca facilidad en el uso del lenguaje**, empleando un vocabulario muy limitado o la elección de palabras incorrectas. Demuestra problemas gramaticales frecuentes que afectan notoriamente el significado.

Puntuación de 1. Demuestra **poco o ningún dominio en la escritura académica**, cuya calidad se ve severamente afectada por varias de las siguientes debilidades: aunque el texto muestre cierta **estructura** no desarrolla ningún punto de vista viable sobre el tema; proporciona poca o ninguna evidencia para apoyar una posición, que puede no ser explícita, o no hay una conclusión claramente definida. El texto está **desorganizado** o desenfocado, resultando en un escrito inconexo o incoherente. Presenta una visible **dificultad en el uso del lenguaje**, con errores importantes en la variedad, la elección y el uso del vocabulario. Demuestra severas fallas en la estructura de las oraciones y contiene errores gramaticales que persistentemente interfieren con el significado.

Figura 33. Versión final de la rúbrica.

Fuente: Elaboración propia con base en McNamara et al. (2009).

E. Perfiles de los jueces

En congruencia con los criterios que ayudaron a definir la estrategia de aplicación descritos en el Apartado 5.1, se reunió un equipo de cinco jueces, anonimizados aquí con letras capitulares, que no tuvieron contacto entre sí durante el proceso.

La Juez A es licenciada en Derecho titulada en 2001. Cuenta con 10 años de experiencia en la docencia en nivel medio superior en el sector público. Tiene el título de maestra en Educación Práctica Educativa (2015), además del diploma del Programa para el Fortalecimiento de la Educación Media Superior (PROFORDEMS, 2013). Se certificó en diciembre del mismo año en el Proceso de Certificación de Competencias Docentes para la Educación Media Superior (CERTIDEMS), y en 2015 obtuvo la acreditación como evaluadora del mismo proceso por el Instituto Nacional de Evaluación Educativa (INEE). Ha impartido materias humanísticas como Derecho, Ciencias Sociales, Filosofía, Ética, Literatura e Inglés.

La Juez B es licenciada en Educación Media con especialidad en Inglés. Se tituló en 2004. Cuenta con 19 años de experiencia en la docencia en nivel medio superior en el sector público. Ha impartido las materias de inglés, filosofía y logística. Tiene el título de maestra en Educación Práctica Educativa (2015), además del diploma del Programa para el Fortalecimiento de las Competencias Docentes en Educación Media Superior (PROFORDEMS, 2013). Se certificó ese el mismo año en el Proceso de Certificación de Competencias Docentes para la Educación Media Superior (CERTIDEMS), y en 2015 fue acreditada como evaluadora del mismo proceso por el Instituto Nacional de Evaluación Educativa (INEE).

La Juez C es licenciada en Ciencias de la Comunicación (1992), maestra en Lingüística aplicada con especialidad en enseñanza y adquisición de L1 y L2 (2010); cuenta con 6 años de experiencia continua en preparatoria impartiendo, mayormente, inglés

como segunda lengua, lectura y redacción y literatura. Actualmente realiza labores de gestión en centros escolares del mismo nivel educativo y es escritora de microrrelatos.

La Juez D es licenciada en Geografía (2016), tiene un año de experiencia en la docencia en el Nivel Medio Superior y ha enseñado Geografía e Historia.

El Juez E es licenciado en Filosofía con estudios en Letras Españolas. Tiene la maestría en Educación (2009) y fue instructor certificado del Programa para el Fortalecimiento de las Competencias Docentes en Educación Media Superior (PROFORDEMS, 2011 a 2014). Ha impartido materias de las Humanidades en los niveles Medio Superior y Superior desde hace 19 años. Fundó un proyecto de Escritura Académica en una escuela de formación de profesores, donde también imparte la asignatura de Producción de Textos Académicos.

F. Postulados de Bioética de la Universidad Autónoma de Querétaro

En el Curso de Capacitación en Ética de Investigación UAQ-CONBIOÉTICA, Capítulo 3. Diseño de las Investigaciones, se indica la obligación de informar si mediaron pagos a los participantes en una investigación: “En ensayos clínicos y estudios sociales, el interés en la protección de sujetos humanos radica en que afecta o pone en riesgo a las personas. Las Pautas del CIOMS, en su versión en español, contienen una lista extensiva de los elementos esenciales del protocolo para consideración ética. Además de los datos de los investigadores, los patrocinadores, instituciones participantes, etc., los requisitos más importantes son: (...) 7. *Un informe de incentivos económicos a los sujetos*” (CONBIOÉTICA, 2018; resaltado nuestro).

El Capítulo 4, Consentimiento informado, expone el criterio basado en el monto y en la condición del participante: “Otro asunto relacionado con la coacción es la cuestión de incentivos o pagos a los participantes, tema ya mencionado con respecto al diseño del

protocolo. Se dice a veces que pueden ser pagos coaccionados o incluso un ofrecimiento de tratamiento que tiene un efecto de coacción. El punto es que estos incentivos pueden invalidar el consentimiento, ya sea *si la recompensa es excesivamente alta o si dirigen las promociones a poblaciones vulnerables*. Los incentivos pueden inducir a la gente a hacer lo que de otro modo iría en contra de su mejor juicio. Es un debate y hasta ahora no ha tenido solución” (CONBIOÉTICA, 2018; resaltado nuestro).

En el mismo capítulo se define lo que constituye una retribución aceptable, en contraste con una inaceptable: “*Retribución aceptable*. Puede reembolsarse a los sujetos de investigación gastos de transporte y otros, incluyendo ganancias no percibidas, asociadas con su participación en la investigación. *Aquellos que no reciben beneficios directos de la investigación podrían también recibir una pequeña cantidad de dinero por la inconveniencia de participar en la investigación*. Todos los sujetos pueden recibir servicios médicos no relacionados con la investigación y tener acceso a procedimientos y exámenes en forma gratuita” (CONBIOÉTICA, 2018; resaltado nuestro).

“*Retribución inaceptable*. Los pagos en dinero o en especie a los sujetos de investigación no debieran ser tan elevados como para persuadirlos de correr riesgos innecesarios o de ofrecerse como voluntarios en forma poco racional. *Los pagos o retribuciones que pongan en peligro la capacidad de la persona de ejercitar una elección libre invalidan el consentimiento*. Puede ser difícil distinguir entre retribución adecuada e influencia indebida para que alguien participe en una investigación. *Una persona sin empleo o un estudiante puede considerar una retribución de forma diferente que una persona con trabajo (...)*” (CONBIOÉTICA, 2018; resaltado nuestro).

G. Resumen de la propuesta de categorización de los sustantivos

Tabla 20. *Propuesta de la lexicología cognitivo-funcional de sustantivos de Schmid.*

Conjuntos de funciones cognitivas	Desempeñadas predominantemente por	Ejemplos
Etiquetar	Nombres propios	Leonhard Lipka, Rover, la Torre Eiffel, Londres, Titanic
Categorizar, etiquetar	Sustantivos comunes de nivel básico	agua, leche, oro, madera, mujer, hombre, escarabajo, caracol, cuerda, auto
Categorizar, destacar, poner en perspectiva, etiquetar	Sustantivos superordinados Sustantivos subordinados Sustantivos funcionales Sustantivos relacionales Sustantivos actitudinales	cubierto, arma, vehículo silla de cocina, lámpara de noche policía, tutor, supervisor jefe, grupo, pieza, córner rufián, novio, basura
Categorizar, partir, reificar, etiquetar	Sustantivos eventivos Sustantivos locativos Sustantivos temporales	conversación, almuerzo, nacimiento área, región, milla momento, día, segundo
Categorizar, partir, reificar, concretizar, etiquetar	Sustantivos abstractos	inflación, comunismo libertad, amor, pasión conversión, transformación dinámica, ionización, inducción ritardando, contrapunto, fuga
Categorizar, partir temporalmente, reificar temporalmente, concretizar temporalmente, etiquetar	Sustantivos encapsuladores	situación, evento, acción, paso hecho, problema, idea, noticia

Fuente: Elaboración propia con base en Schmid (1999b).

H. Resumen de criterios de inclusión, exclusión y conteo de sustantivos

(a) Se contabilizaron las instancias de cada sustantivo en su forma lematizada; tanto en el conteo de frecuencias como en el cálculo de su dispersión (la cantidad de textos diferentes

en los que aparece un mismo sustantivo en el total de textos, dato de interés para futuras indagaciones sobre el corpus). Por ejemplo, en: “Hemingway en su obra El viejo y el mar habla de cómo a pesar de la *adversidad*, el hombre debe seguir luchando” (15M30); “Albert Camus en El extranjero y Ernest Hemingway en El viejo y el mar presentan dos maneras opuestas de seguir adelante a pesar de las *adversidades*” (15M30). En este caso hay una frecuencia de 2 y dispersión de 1/120. Las proporciones de cualquier sustantivo se calcularon con la frecuencia del lema en relación con el total de sustantivos del corpus completo (n=120), o bien de la muestra (n=60).

(b) La clase sintáctica tuvo que desambiguarse para evitar incluir adjetivos o verbos en la lista de sustantivos etiquetada automatizadamente. Las bases fueron el Diccionario de la Lengua Española de la RAE (2104) y el contexto discursivo de la palabra, con ayuda de las líneas de concordancia obtenidas con AntConc. Son los casos de *consecuente*: ejemplo: “como se observa en la cita *consecuente*” (adjetivo) (15M10); *característica*: “se hace notar esta *característica* en la siguiente cita” (sustantivo) (14H24); “Meursault (...) tiene una visión muy *característica* de la vida” (adjetivo) (14M18), o *causa*: “esto le *causa* una cantidad enorme de problemas” (verbo) (14H04), “Un ciclo infinito de *causa*-efecto en donde la *causa* es la consecuencia del acto de decidir” (sustantivos) (14M26).

(c) Un mismo sustantivo abstracto puede tener distintas acepciones según su referente, por ello fue necesario analizar el contexto discursivo de cada caso para poderlo categorizar en ‘concreto’/‘abstracto’ y ‘especializado’/‘no especializado’, más sus clases. Si un sustantivo abstracto funciona en el texto como término especializado, se enlistó y contabilizó en la base de datos como sustantivo abstracto, y se le adjuntó también la etiqueta de especializado. Por ejemplo, *alusión*: “la variable [intertextualidad] se puede conocer como las semejanzas que hay entre dos o más textos y que estas pueden ser por medio de *alusiones*, metáforas, imitación, influencia y demás” (15H09), en este texto nombra una figura retórica, pero es asimismo un sustantivo abstracto que designa el

resultado o efecto de un acto verbal: ‘referirse a algo del contexto’, por lo que lleva las dos etiquetas. Asimismo, *juicio*: en “todos los que estaban presentes en su *juicio* tenían esperanzas de verlo libre”(14M06), se refiere al acto jurídico en un tribunal; mientras que “el narrador no puede emitir un *juicio* de valor. No sería el más adecuado.” (14H02), remite a un proceso intelectual; en tanto que en “los tipos de personaje son protagonista, antagonista, personaje pivote, *elemento de juicio*, etc.” (14H17) el sustantivo compone un término especializado de la teoría narrativa. Otro caso es *creación*: “una de sus aportaciones más importantes es la *creación* de la técnica del iceberg” (14M21), en que el sustantivo denota el efecto de crear. En cambio, en “la *creación* de Samuel Beckett pertenece al género dramático” (14M28), el sustantivo es cuasi sinónimo de la obra literaria. De modo que un mismo sustantivo puede aparecer en en la lista de los abstractos y en la lista de los términos especializados.

(d) No entraron en la contabilización de sustantivos abstractos latinismos, anglicismos, germanismos, etc., pero sí se consideraron en la lista de los términos disciplinares aunque sean adopciones de otras lenguas, como *übermensch*, *alter ego* o *flash-back*, debido a que este tipo de adopciones es común en el lenguaje técnico o disciplinar, y eliminarlas restaría precisión al conteo.

(e) No se incluyeron sustantivos que formaran marcadores discursivos o conectores como *de acuerdo con* o *por consiguiente*: “*Por consiguiente* este ensayo propone...” (15H19); *a causa de*: “siendo ejecutado *a causa de* un crimen” (15H10), *en cambio*: “*en cambio* el hombre no [se] ve dañado ante la sociedad” (14H23), *a continuación*: “*A continuación* se muestra un fragmento de ello” (14H16); o frases cristalizadas por el uso (locuciones) como [dar/esperar/lograr algo] *a cambio* [de]: “lograr aquello que se proponen, *a cambio de* luchar por ello” (15H26).

(f) No entraron en la categoría de sustantivos abstractos las unidades de tiempo y medida espacial (como *año*, *lustro*, *momento*, *sinfin*, *tramo*), por las razones expresadas en la descripción de esta variable (ver Apartado 5.3.2.6).