

# EL LÉXICO DISCIPLINAR EN LA EDUCACIÓN UNIVERSITARIA

**Mireya Cisneros Estupiñan**

Universidad Tecnológica de Pereira  
Colombia

**Giohanny Olave Arias**

Universidad Industrial de Santander  
Colombia

Mireya Cisneros Estupiñan  
Giohanny Olave Arias  
JEL: A20, A29

## El léxico disciplinar en la educación universitaria

### Resumen

En este artículo, presentamos una investigación en curso centrada en el plano léxico del discurso académico en la educación universitaria; específicamente, nos interesa analizar la dimensión léxica en los textos escritos por egresados de carreras profesionales en el momento en que ingresan a los estudios postgraduales. Teóricamente, articulamos aspectos de los estudios sobre disponibilidad léxica - que pretenden descubrir qué lexías sería capaz de usar un hablante en determinados temas de comunicación y centros de interés - con los estudios socio retóricos sobre géneros discursivos académicos y comunidades disciplinares. El corpus para este artículo está constituido por textos argumentativos. En este primer avance, presentaremos los indicadores de análisis definidos para la delimitación del «léxico académico» y que constituyen la ruta de trabajo a seguir: medición de la riqueza léxica; análisis de similitudes, análisis terminológico con nubes de palabras; análisis Reinert o de mundos lexicales; análisis factorial de correspondencias; análisis de la modelización léxica; análisis de las relaciones léxicas y de la articulación textual. Además, expondremos resultados iniciales alcanzados con las herramientas léxico métricas provistas por el software Iramuteq 0.7 alpha 2.

**Palabras clave:** Léxico académico; lexicometría; escritura en posgrado; comunidad disciplinar.

---

## Les défis de l'enseignement de la discipline d'entreprise

### Résumé

Dans cet article, nous présentons une enquête en cours centrée sur le plan léxico du discours académique dans l'enseignement universitaire; plus précisément, nous nous intéressons à l'analyse de la dimension lexicale dans les textes rédigés par les diplômés de carrières professionnelles au moment de leur entrée en troisième cycle. Théoriquement, nous articulons des aspects des études sur la disponibilité lexicale - qui cherchent à découvrir quels lexiques un locuteur pourrait utiliser dans certains problèmes de communication et centres d'intérêt - avec des études socio-rhétoriques sur les genres académiques discursifs et les communautés disciplinaires. Le corpus de cet article est constitué des textes argumentatifs. Dans cette première avancée, nous présenterons les indicateurs d'analyse définis pour la délimitation du «lexique académique» et qui constituent la voie de travail à suivre: mesure de la richesse lexicale; analyse de similitude, analyse terminologique avec des nuages de mots; Analyse du monde reinert ou lexical; analyse factorielle des correspondances; analyse de modélisation lexicale; analyse des relations lexicales et articulation textuelle. De plus, nous présenterons les premiers résultats obtenus avec les outils de lexique métrique fournis par le logiciel Iramuteq 0.7 alpha 2.

**Mots clés:** Lexique académique; lexicométrie; écriture diplômée; communauté disciplinaire.

---

## The disciplinary lexicon in university education

### Abstract

In this article, we present an ongoing investigation focused on the lexical plane of academic discourse in university education; specifically, we are interested in analyzing the lexical dimension in the texts written by graduates of professional careers at the time they enter postgraduate studies. Theoretically, we articulate aspects of studies on lexical availability - which seek to discover what lexies a speaker would be able to use in certain communication issues and centers of interest - with socio-rhetorical studies on academic discursive genres and disciplinary communities. The corpus for this article is made up of the argumentative texts. In this first advance, we will present the analysis indicators defined for the delimitation of the «academic lexicon» and that constitute the work route to be followed: measurement of lexical wealth; similarity analysis, terminology analysis with word clouds; Reinert or lexical world analysis; factorial analysis of correspondences; lexical modeling analysis; analysis of lexical relationships and textual articulation. In addition, we will present initial results achieved with the metric lexicon tools provided by the software Iramuteq 0.7 alpha 2.

**Keywords:** Academic lexicon; lexicometry; graduate writing; disciplinary community.

Artículo de Investigación

## El léxico disciplinar en la educación universitaria

### INFORMACIÓN DEL ARTÍCULO

Recepción del artículo: 10/05/2018  
Concepto de evaluación: 30/06/2018  
Aceptación del artículo: 22/08/2018

Mireya Cisneros Estupiñán\*  
Universidad Tecnológica de Pereira, Colombia

Giohanny Olave Arias\*\*  
Universidad Industrial de Santander, Colombia

### INTRODUCCIÓN

Los estudios sobre enseñanza de la lectura y la escritura en la universidad constituyen actualmente un campo activo de investigación en Latinoamérica (Ávila, 2017; Navarro et al., 2016; WRAB, 2017). Diversas iniciativas en Colombia han contribuido significativamente en el área (González y Vega, 2013) y siguen explorándola con una diversidad de métodos en sus aproximaciones. Desde el grupo “Estudios del Lenguaje y la Educación”, categoría B de Colciencias, añadimos a los aportes que hemos realizado en ese campo (Cisneros, Olave y Rojas, 2013a, 2013b, 2014, 2016c; Cisneros y Muñoz, 2014) una propuesta de trabajo centrada en el plano lexical del discurso académico universitario.

Entendemos el léxico académico como el conjunto de lexías (Pottier, 1977) o de unidades léxicas asociadas estrechamente o usadas recurrentemente en los géneros discursivos y textuales del ámbito académico (Bazerman, 2012; Hyland, 2009; Parodi, 2008) y que contribuyen a la conformación de comunidades disciplinares específicas (Hyland, 2004) dentro de las cuales funcionan esos géneros. En este sentido, la noción resulta indesligable del léxico disciplinar o del uso de lexías con significados precisos y particulares dentro de los límites de un área de estudio, es decir, su relación con el vocabulario especializado, técnico, semitécnico y subtécnico de las profesiones (Coxhead, 2016:177-178), pero también con las variaciones disciplinares que presenta el uso del léxico en el ámbito académico y que forman parte de los recursos que utiliza cada disciplina para conformar sus argumentos y constituirse en cuanto tal (Hyland & Tse, 2009).

La noción de léxico académico presupone la idea de que algunas unidades léxicas son usadas más frecuentemente en textos académicos y disciplinares que en textos pertenecientes a otras actividades sociales. La larga tradición de los estudios en terminología, lexicología y lexicografía (Adelstein, 2004) apoyan este presupuesto, lo mismo que algunas aproximaciones desde la psicología cognitiva del desarrollo, como la Hipótesis de la calidad lexical (Perfetti, 2007), los enfoques desde la Lingüística de corpus

\*Autor de correspondencia: mireyace@gmail.com. ORCID: 0000-0002-5519-7192

\*\*Autor de correspondencia: giolavar@uis.edu.co

(Matsuda, Sadowsky y Sabaj, 2011; Venegas, 2007) y los estudios de enseñanza del inglés con fines específicos, centrados en el léxico (Coxhead, 2014, 2016; Hyland & Tse, 2009; Nation, 2013). No obstante, ninguno de estos enfoques autoriza un paso automático hacia la alfabetización académica, pues es evidente la insuficiencia del aprendizaje de un lexicón para el dominio de la lectura y la escritura en las disciplinas profesionales.

Esta investigación no obvia esa complejidad, sino que propone abordar la dimensión léxica de los textos académicos para 1) atender ese campo de vacancia en los estudios actuales sobre alfabetización académica en español; 2) analizar el desempeño léxico (disponibilidad, diversidad, densidad y relaciones léxicas) en la producción escrita de profesionales aspirantes a cursar estudios posgraduales, y 3) derivar propuestas que se articulen y contribuyan a la formación posgradual.

El uso del léxico especializado de la Lingüística conforma un nivel dentro de la cultura disciplinar (Gunnarson, 2009; Hyland, 2004) de esta profesión, es decir, forma parte de las prácticas que constituyen a los lingüistas como una comunidad de profesionales encargada de investigar sobre la lengua, en los ámbitos teórico y aplicado. Las prácticas de cada cultura disciplinar no se limitan a la regulación de estilos formales de escritura, sino que cubren todos los aspectos que le permiten al profesional actuar y ser reconocido como parte de una comunidad de especialistas.

Enseñar a leer y escribir en la universidad significa ayudar institucionalmente a los estudiantes a ingresar y participar en las prácticas discursivas propias de las profesiones para las cuales se preparan (Carlino, 2013) o en las cuales se especializan (Arnoux, 2009). En el caso de la posgraduación en Lingüística, el desafío resulta mayúsculo, no solo por la activa dinámica del campo disciplinar, en constante demanda de actualización, sino también y sobre todo por la exigencia de estudiar la lengua con la lengua misma, es decir, el solapamiento entre el objeto de estudio y las herramientas para abordarlo. Pensamos que esta especificidad disciplinar hace muy pertinente poner el foco en el ámbito lexical.

En este primer avance, presentaremos los indicadores de análisis definidos para la delimitación del léxico académico y que constituyen la ruta de trabajo a seguir: medición de la riqueza léxica; análisis de similitudes, análisis terminológico con nubes de palabras; análisis Reinert o de mundos lexicales; análisis factorial de correspondencias; análisis de la modalización léxica; análisis de las relaciones léxicas y de la articulación textual. Además, expondremos resultados iniciales alcanzados con las herramientas lexicométricas provistas por el software Iramuteq 0.7 alpha 2.

A partir de las conclusiones de esta investigación, se derivarán pautas de intervención para el mejoramiento de los procesos formativos posgraduales involucrados en la escritura de textos académicos, con impacto directo en los procesos formativos pregraduales y postgraduales.

## **METODOLOGÍA**

Articulamos aspectos de los estudios lexicométricos sobre disponibilidad léxica -que pretenden descubrir qué lexías sería capaz de usar un hablante en determinados temas de comunicación y centros de interés- con los estudios sociorretóricos sobre géneros discursivos académicos y comunidades disciplinares. En principio, utilizamos los índices cuantitativos de riqueza léxica (densidad, variedad y hápax), frecuencias léxicas,

asociaciones semánticas y análisis factorial de coocurrencias. Complementariamente, se analizarán las relaciones léxicas de cohesión y modalización textual, en una segunda fase de la investigación. Utilizamos las herramientas de lexicometría automatizada, provistas por el software Iramuteq 0.7 alpha 2.

El corpus está constituido por los textos argumentativos que los aspirantes a la maestría en Lingüística de la Universidad Tecnológica de Pereira presentaron como requisito de admisión, para las cohortes IV y V, correspondientes a los años 2014 y 2015. La recolección de los textos se realizó a través de una convocatoria abierta institucional, ligada al inicio de los estudios de maestría, en la cual se le solicita a cada aspirante la presentación escrita de dos textos argumentativos: una propuesta de tema de investigación y un ensayo sobre el impacto de la Lingüística en la sociedad. Para las cohortes IV (16 estudiantes) y V (26 estudiantes) se recolectaron 81 textos: 39 propuestas y 42 ensayos. En su mayoría, los colaboradores ejercen como docentes de español en instituciones de educación básica y media, lo que corresponde a su formación profesional en los pregrados, y provienen del ámbito regional del Eje cafetero y de lugares aledaños.

## RESULTADOS

### *Indicadores de análisis para estudiar el léxico disciplinar*

Proponemos ocho indicadores analíticos para estudiar el léxico disciplinar, basados en los rasgos genéricos propios de la escritura para el posgrado y, en particular, en las regularidades y regulaciones de los géneros textuales “propuesta de investigación” y “ensayo reflexivo”, en el área de la Lingüística al nivel de maestría. Aclaramos que estos indicadores propuestos se encuentran relacionados entre sí y contribuyen, en su totalidad, al léxico académico tal como lo definimos en la sección introductoria. Presentaremos cada indicador, seguido de algunos avances realizados en el análisis del corpus:

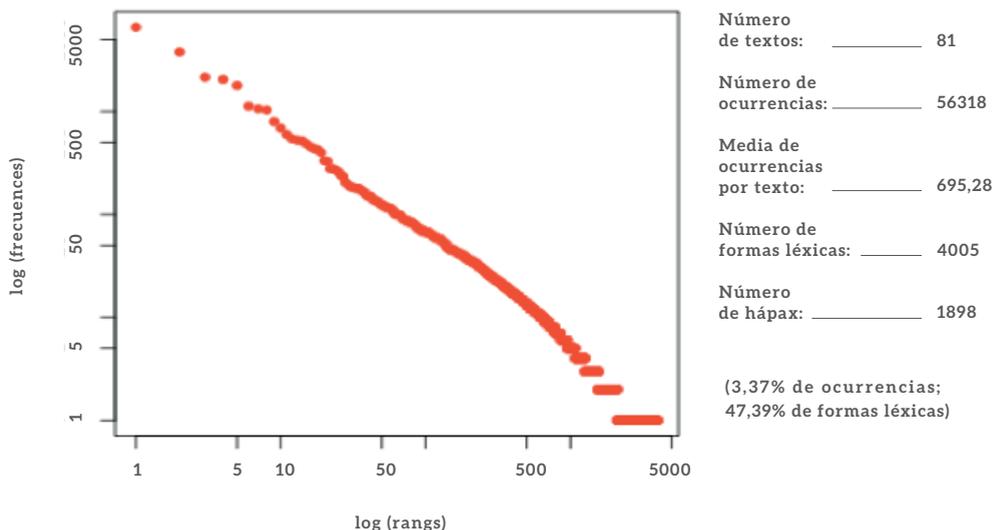
#### **1. Riqueza léxica**

Las pruebas que miden la riqueza léxica de los textos de un corpus son de naturaleza cuantitativa y comparativa (López Morales, 1984). Se tienen en cuenta tres indicadores principales: 1) Densidad léxica, esto es, la proporción de lexías diferentes en relación con el total de lexías de un texto. Para hallar este índice se cuenta el número total de palabras (ocurrencias) y el de vocablos (palabras diferentes o formas léxicas) y se calcula la proporción entre ambos valores. 2) Intervalo de aparición de las palabras nocionales. Este índice considera la cantidad total de palabras del texto dividido por los sustantivos, adjetivos, verbos y adverbios (formas activas), para obtener el intervalo de aparición de estas palabras. El índice aumenta según disminuye el intervalo de aparición de lexías nocionales; en contraste, cuanto más amplio es el intervalo, menor es la riqueza léxica de la muestra. 3) El índice hápax identifica el uso de lexías utilizadas solo una vez en el corpus. El valor de hápax se obtiene aplicando la fórmula  $V(\text{número de vocablos})/V1(\text{suma de palabras de una ocurrencia})$ . Para determinarlo, se divide la cifra total de lexías por aquellas que solo se emplean una vez en el texto.

En síntesis, los textos que contienen el mayor porcentaje de vocablos y el menor intervalo de palabras nocionales y de hápax son los clasificados como de mayor riqueza léxica, si bien está sujeta a las condiciones disciplinares de la terminología especializada. Este tipo de índice permite conocer el dominio del léxico activo de un sujeto o de un grupo y, a partir de allí, programar intervenciones pedagógicas.

En el corpus de escrituras de aspirantes a la Maestría, la Figura 1 muestra los resultados de los índices 1 (densidad léxica) y 3 (hápx):

Figura 1. Logaritmos de frecuencia y de rango en el corpus



La estadística textual indica que en 56318 usos léxicos se utilizan apenas 4005 formas distintas, es decir, cada lexía nocional (nombres, adjetivos, verbos y adverbios) se utiliza en un 7,11% en el total de 81 textos, lo que constituye el porcentaje de vocablos. Este indicador grueso nos ofrece una visión general de la baja diversidad léxica del corpus. En cuanto al intervalo de aparición de lexías con contenido nocional, según es calculada por López Morales (2011), la proporción de usos nocionales en el texto es de 14,06. Esto significa que en los textos es necesario esperar cada 14,06 lexías para que aparezca un uso nuevo de contenido semántico nocional. Teniendo en cuenta que a mayor número de usos nocionales menor es el valor del intervalo y mejor el índice de riqueza léxica, este primer análisis indica que el corpus presenta un vocabulario restringido.

En cuanto a las lexías usadas solo una vez en el corpus, representan una mínima parte (3,37%) del total de ocurrencias y casi la mitad (47,39%) del total de formas léxicas distintas. El resultado indica una muy baja presencia de esas lexías únicas; como señalan Matsuda et al. (2011), el mayor porcentaje de hápx se vincula con una alta presencia de terminología y supone un esfuerzo máximo en la comunicación, para alcanzar un registro altamente elaborado destinado a una audiencia especializada y reducida. Queda pendiente aplicar otros índices complementarios, extrapolados de la enseñanza de segundas lenguas, como el índice de complejidad léxica y el de fluidez, que describen la expresividad de una muestra de textos escritos (Larsen-Freeman, 2006; Bulté & Housen, 2012). para determinar con mayor precisión el nivel de riqueza léxica de los textos escritos por los colaboradores.

## 2. Análisis terminológico con nubes de palabras

A partir de tablas de frecuencias y nubes de palabras, se analiza la aparición de lexías repetidas en el corpus y se les representa con tamaños de letra distintos. Las tablas de frecuencia deben ser revisadas para identificar cuál es la frecuencia relativa del uso del léxico especializado y de la terminología lingüística (Adelstein, 2004). Un nivel bajo de uso de este léxico y, en contraste, un nivel alto de lexías “comodin” y de usos coloquiales, indica niveles más restringidos de léxico académico. La Figura 2 presenta la nube correspondiente al corpus de lexías activas, no lematizadas.

Figura 2. Nube de palabras



Las lexías más visibles, lenguaje, lengua y lingüística, presentan la frecuencia de aparición más alta y frecuencias similares entre ellas (0,68%; 0,67% y 0,53%, respectivamente). En esta disciplina, como es sabido, los primeros dos términos presentan diferencias semánticas sustanciales (Areiza, Cisneros y Tabares, 2012:8-9); por tanto, habrá que evaluar estos usos en cada contexto, para determinar si evidencian sus especificidades disciplinares. Este análisis cualitativo es una tarea pendiente.

La Tabla 1 muestra el resultado descendente de todas las lexías activas que conforman el 5,92% de usos en el total del corpus (formas con ocurrencias mayores o iguales a 100, en un total de 56318 ocurrencias):

**Tabla 1.**  
**Frecuencias léxicas del 6% del corpus**

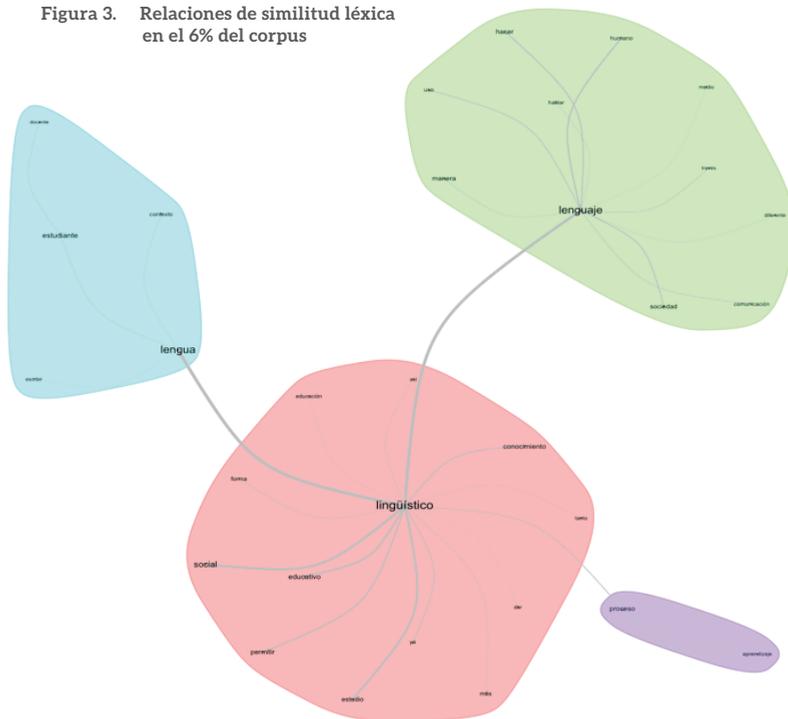
Forma léxica	Cat. gram.	Frec. absoluta	Frec. relativa
Lenguaje	nom	381	0,68
Lengua	nom	377	0,67
Lingüística	adj	299	0,53
Social	adj	171	0,30
Manera	nom	165	0,29
Más	adv	153	0,27
Estudiantes	nom	153	0,27
Sociedad	nom	149	0,26
Conocimiento	nom	141	0,25
Educación	nom	123	0,22
Ya	adv	121	0,21
Uso	nom	121	0,21
Estudio	nom	121	0,21
Comunicación	nom	120	0,21
Aprendizaje	nom	115	0,20
Diferentes	adj	110	0,20
Tanto	adv	105	0,19
Forma	nom	104	0,18
Sociales	adj	103	0,18
Proceso	nom	102	0,18
Así	adv	100	0,18
			<b>5,92</b>

Este análisis preliminar debe ser complementado con una estadística de lexías complejas (sintagmas, locuciones y fórmulas fijas) y, en segundo término, con una categorización cualitativa de usos coloquiales y comodines léxicos, a partir del contraste con la bibliografía especializada en lingüística. Por ejemplo, es necesario desagregar la lexía “manera”, que aparece con una frecuencia alta (0,29%), para diferenciar sus funciones sustantivas de las locucionales en cotexto preposicional (de cualquier manera, de todas manera(s), de una manera u otra, etc.), pues estos usos afectan el registro formal en los textos del corpus. El caso puede ser similar para los adverbios de la Tabla 1 (más, ya, tanto, así).

### 3. Análisis de similitudes léxicas

Se muestran las relaciones entre las lexías activas del corpus, a través de un gráfico de red (Garnier et Guérin, 2010; Pélissier, 2016). El tamaño de las lexías representa su frecuencia; los enlaces pueden aparecer más gruesos o más delgados y más cercanos o más lejanos, de acuerdo con la mayor o menor fuerza asociativa entre los significados. En la Figura 3 se representan las relaciones de similitud entre las frecuencias léxicas más altas del 6% del corpus.

Figura 3. Relaciones de similitud léxica en el 6% del corpus



A través de los halos de color, el gráfico muestra los grupos léxicos que se asocian alrededor de las tres lexías más recurrentes. Se concibe más cercanía de la lingüística con la lengua que con el lenguaje, aunque los significados asociados a la primera resultan ser más restringidos, tanto en cantidad (4 lexías: estudiante-docente, contexto y escribir) como en mundo de referencia (educativo).

La red semántica establecida dentro del halo “lenguaje” parece ser más compleja y reenvía tanto a acciones (hablar, hacer, uso) como a sociedad (comunicación, humano, diferente, medio, manera), pero llama la atención que las ramificaciones indican una cierta desconexión (direccionalidad opuesta) entre esas acciones y la comunicación en sociedad.

En cuanto al aprendizaje como proceso, aparece desprendido de la lingüística, pero exterior a su campo de dominio (halo violeta), según la distancia asociativa que queda representada en el gráfico. El halo central ha sido constituido con la lexía lingüístico y sus lexías asociadas, por lematización; en este campo, el significado más cercano es educativo y se presenta un vínculo fuerte con social y estudio. Si bien la Figura solo representa el 6% del corpus, la distribución asociativa se mantiene en los gráficos del corpus completo, no incluidos aquí por cuestiones de espacio y para privilegiar la claridad.

#### **4. Análisis Reinert o de mundos lexicales**

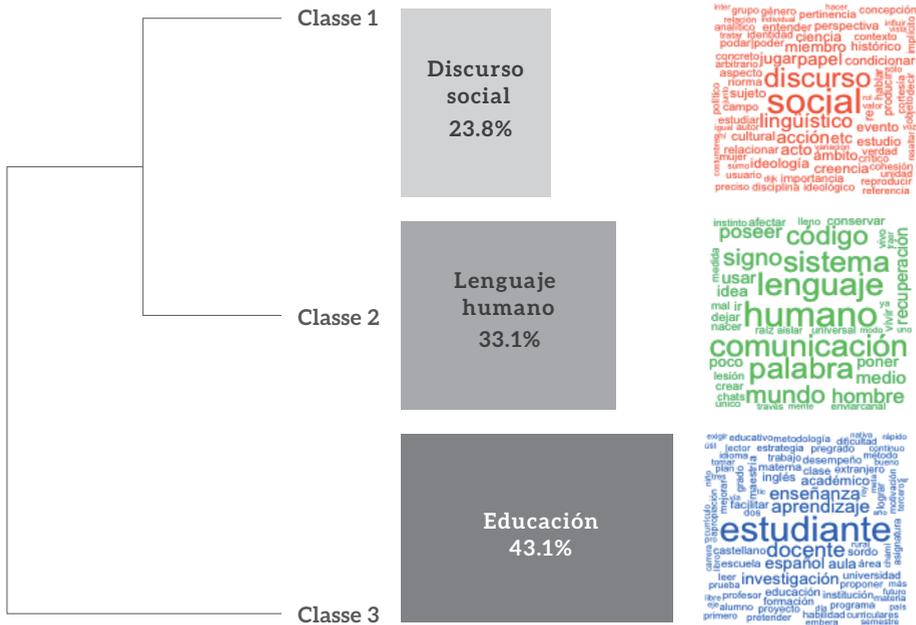
Con base en un análisis factorial de coocurrencias, se identifican los “mundos lexicales” (Reinert, 1993) o grupos de lexías organizados en torno a núcleos de significado; las similitudes y diferencias entre usos léxicos permiten constituir categorías conceptuales. La clasificación de las lexías se realiza en orden jerárquico descendente a través de un test que calcula, mediante un algoritmo, la correlación entre variables comparando la distribución observada de los datos con la distribución esperada de ellos. Las categorías se muestran en dendogramas y representan agrupaciones de puntos de vista de los enunciadores con respecto a los temas sobre los cuales escriben. Las Figuras 4 y 5 ofrecen visualmente la organización jerarquizada de los agrupamientos lexicales o clases de lexías, de acuerdo con sus mundos de referencia.

Las clases 1 y 2, que hemos categorizado como “discurso social” y “lenguaje humano”, atendiendo a las lexías de mayor frecuencia en cada agrupamiento, constituyen un poco más de la mitad del corpus (56,9%), mientras que la clase 3, “educación”, el 43,1%. La asociación cuantitativa de las clases “discurso social” y “lenguaje humano” responde a una vinculación cualitativa que presenta relaciones de proximidad lexical más fuertes entre ellas que la clase “educación”, claramente desligada en un mundo lexical independiente.

Figura 4. Agrupamientos lexicales jerarquizados



Figura 5. Agrupamientos lexicales jerarquizados



El análisis de mundos lexicales muestra que en el corpus estos tres campos de significado, se conforman dos bloques diferenciados y equilibrados: el de la Lingüística (Discurso social + Lenguaje humano) y el de la educación. Este último tiene una conexión más débil que la presentada entre el discurso y el lenguaje, es decir, que emerge un punto de vista sobre la educación como un asunto independiente del estudio científico del lenguaje. Una perspectiva así orientada parece confirmar lexicométricamente lo que se ha hallado cualitativamente en otras investigaciones precedentes (Cisneros, Olave y Rojas, 2016a, 2016b), al respecto de la desconexión empírica entre las ciencias del lenguaje y las ciencias de la educación.

Para profundizar en esta hipótesis y complejizar el análisis, será necesario indagar en los “perfiles” o listados completos de lexías que constituyen cada clase según el valor de la prueba Chi cuadrado; los “antiperfiles” o formas disociadas de cada clase, así como la generación de gráficos de proximidad léxica para lexías claves en cada núcleo de significado, como “discurso” (para la clase 1), “gramática” (para la clase 2) e “investigación” (para la clase 3).

#### 4. Análisis Reinert o de mundos lexicales

Con base en un análisis factorial de coocurrencias, se identifican los “mundos lexicales” (Reinert, 1993) o grupos de lexías organizados en torno a núcleos de significado; las similitudes y diferencias entre usos léxicos permiten constituir categorías conceptuales. La clasificación de las lexías se realiza en orden jerárquico descendente a través de un test que

calcula, mediante un algoritmo, la correlación entre variables comparando la distribución observada de los datos con la distribución esperada de ellos. Las categorías se muestran en dendogramas y representan agrupaciones de puntos de vista de los enunciadores con respecto a los temas sobre los cuales escriben. Las Figuras 4 y 5 ofrecen visualmente la organización jerarquizada de los agrupamientos lexicales o clases de lexías, de acuerdo con sus mundos de referencia.

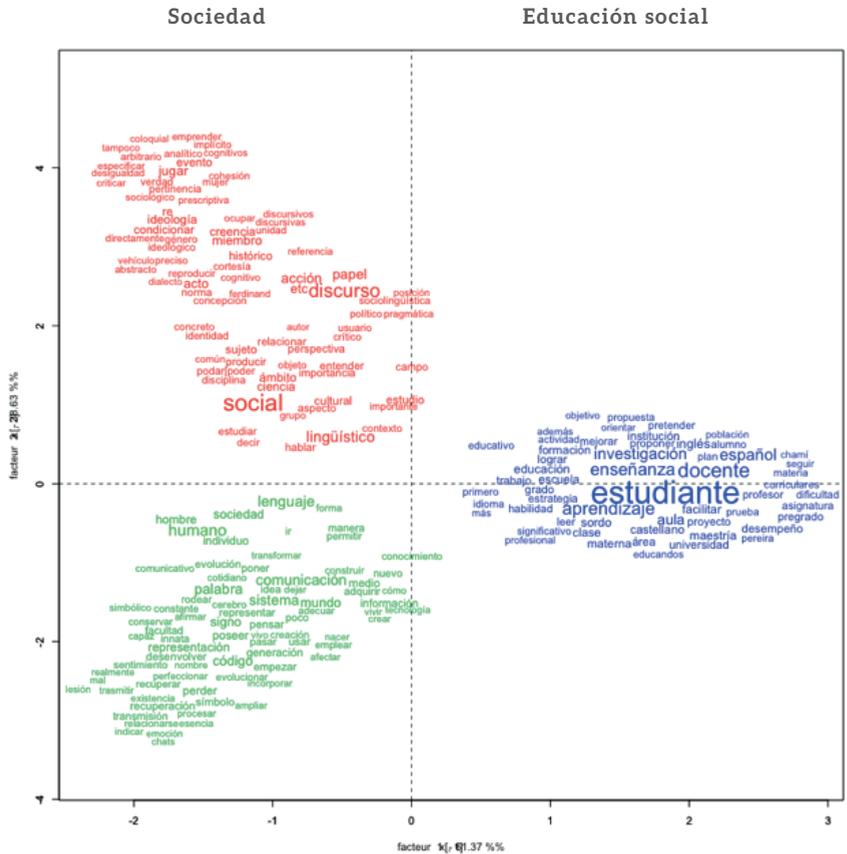
Las clases 1 y 2, que hemos categorizado como “discurso social” y “lenguaje humano”, atendiendo a las lexías de mayor frecuencia en cada agrupamiento, constituyen un poco más de la mitad del corpus (56,9%), mientras que la clase 3, “educación”, el 43,1%. La asociación cuantitativa de las clases “discurso social” y “lenguaje humano” responde a una vinculación cualitativa que presenta relaciones de proximidad lexical más fuertes entre ellas que la clase “educación”, claramente desligada en un mundo lexical independiente.

### **5. Análisis factorial de correspondencias léxicas**

Permite analizar la aparición simultánea de lexías. La coincidencia entre ellas indica cercanías entre significados, con las cuales se constituyen agrupamientos y se confirma el análisis Reinert. Estos grupos se acercan o se alejan de otros grupos de lexías coocurrentes, todo lo cual ayuda a identificar las asociaciones semánticas que realizan los informantes. En este nivel analítico del léxico se pueden generar hipótesis interpretativas sobre el direccionamiento de las asociaciones semánticas y el tratamiento de los temas sobre los cuales se les ha pedido escribir a los informantes.

En la Figura 6, los mundos lexicales hallados antes se distribuyen en plano factorial, conservando los colores asociados con las clases ya categorizadas. La sumatoria de los factores resume el 100% de la información del corpus (Factor 1: 38,63%; Factor 2: 61,37%). En el eje horizontal, es clara la oposición o aislamiento de la clase “educación” con respecto a las demás, si bien resulta ser un agrupamiento más cohesionado (de proximidades léxicas más fuertes) que los otros dos, más dispersos. La posición de esta clase es muy interesante, pues tiene potencial mediador entre las otras dos clases, aunque se ubique oposicionalmente.

Una aproximación analítica a la Figura 6 presenta la tendencia de los cuadrantes hacia una construcción de significados más o menos delimitados en el corpus de escrituras de los aspirantes a la Maestría en Lingüística.



En los cuadrantes superiores queda evidenciada la distancia entre los núcleos semánticos “Sociedad” y “Educación social”, así como en los cuadrantes inferiores sucede de la misma forma con los núcleos “Lingüística” y “Educación lingüística”. La frontera central es atravesada tímidamente por la sociolingüística y la pragmática, en el eje superior, y por conocimiento y tecnología, en el eje inferior: consideramos que estos significados, que asumimos como representaciones disciplinares, lideran el tránsito hacia el acercamiento de la Lingüística (de la lengua y del discurso) hacia el ámbito de lo educativo.

### 6. Análisis de la modalización léxica

A partir de la tabla de lexías activas del corpus extraeremos las categorías gramaticales “adjetivo” y “adverbio”, así como los términos con derivaciones sufijales apreciativas (RAE, 2009), para analizar la cantidad y modos en que los informantes califican o se posicionan frente a los significados que construyen. Las modalizaciones ubican los lugares textuales donde se introducen opiniones y puntos de vista de manera explícita, con lo cual se puede evaluar en esos tramos la adecuación y efectividad persuasiva de la subjetividad dentro de las fronteras discursivas del género académico.

## 7. Análisis de las relaciones léxicas

Se analizan dos tipos de relaciones léxicas: las reiteraciones y las colocaciones (Halliday & Hasan, 1976; Martínez, 2015[1994]). Las primeras se refieren a la cantidad y modalidades de persistencia de los significados: repeticiones literales, sinonimias o cuasisinonimias, superordenaciones y generalizaciones. Las segundas se refieren a los modos de construcción de cadenas semánticas a partir de la relación entre lexías asociadas y colocadas a lo largo de los textos. Ambos mecanismos, reiteraciones y colocaciones, son indicadores de los modos de construcción de la cohesión textual.

Figura 6. Análisis factorial de correspondencias

Lingüística

Educación lingüística

## 8. Análisis de la articulación textual

Se analiza el uso de marcadores discursivos y conectores textuales utilizados por los informantes para contribuir a la construcción de la cohesión de sus textos (Montolío, 2001). Los mecanismos de conexión en el discurso académico son específicos y exigen el uso de recursos altamente variados y de precisión en la relación entre segmentos, para evitar problemas de coherencia o de ambigüedad.

## CONCLUSIONES

El trabajo que hemos presentado es una ruta a seguir, sobre la base de que analizar el nivel léxico a las puertas del proceso formativo para ingresar a una comunidad de especialistas, puede contribuir positivamente en el direccionamiento curricular y formativo de los programas de posgrado en cualquier área de conocimiento o especialidad académica, por sus desafíos epistemológicos y metadiscursivos.

Como un primer acercamiento a este reto, presentamos los indicadores de análisis definidos para estudiar el «léxico académico» en un corpus constituido por argumentaciones escritas por profesionales que egresaron de carreras afines al postgrado que pretenden realizar. Esta particularidad nos permite ver cómo egresan los profesionales en materia de manejo de la lengua escrita con fines académicos y/o en contextos académicos, y también con qué fundamentos lexicales que deben ser parte del dominio de la disciplina, asumen una nueva etapa de su formación, que en este caso es de mayor atención porque se trata de estudios postgraduales.

Insistimos, sin embargo, en que abordar este nivel en el marco de los estudios sobre alfabetización académica es necesario, pero no suficiente. Nuestra contribución busca realizar una entrada al problema, focalizando en los desafíos léxico-semánticos de las disciplinas y en el papel de la escritura en su construcción, en busca de alternativas que se sumen a los esfuerzos actuales por cualificar la comprensión y producción textual en los procesos educativos que forman para el desempeño de las profesiones.

Los resultados iniciales, alcanzados con algunas herramientas lexicométricas, confirman que es necesario articular las representaciones sobre la lengua –como sistema- y el discurso –como acción social-, con la educación –como praxis-. La aparición de esos dominios en las escrituras señala que la formación académica previa permitió una baja proximidad entre los significados asociados a estos mundos léxicos, lo cual se refleja en las dificultades para proponer investigaciones interdisciplinarias que enriquezcan tanto la ciencia lingüística como la educativa. A quienes creemos en esa interdisciplinariedad nos resta, entonces, mucho trabajo por hacer.

## REFERENCIAS BIBLIOGRÁFICAS

- Adelstein, A. (2004). Unidad léxica y valor especializado: Estado de la cuestión y observaciones sobre su representación. Tesis de Doctorado, Universidad Pompeu Fabra, Barcelona.
- Areiza, R., Cisneros, M. y Tabares, L. (2012). Sociolingüística. Enfoques pragmático y variacionista (2ª ed.). Bogotá: Ecoe.
- Arnoux, E. N. de (Dir.) (2009). Escritura y producción de conocimiento en las carreras de posgrado. Buenos Aires: Santiago Arcos.
- Ávila, N. (2017). Postsecondary writing studies in Hispanic Latin America: Intertextual dynamics and intellectual influence. *London Review of Education*, 15(1), 21-37. DOI: <https://doi.org/10.18546/LRE.15.1.03>
- Bazerman, Ch. (2012). Géneros textuales, tipificación y actividad. Puebla, México: Benemérita Universidad Autónoma de Puebla.
- Bulté, B. & Housen, A. (2012) Defining and operationalising L2 complexity. En: A. Housen, F. Kuiken y I. Vedder (eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (pp. 21-46). Amsterdam: John Benjamins.
- Carlino, P. (2013). Alfabetización académica diez años después. *Revista Mexicana de Investigación Educativa*, 18(57), 355-381.
- Cisneros, M. y Muñoz, C. (2014). Tras las huellas de las investigaciones sobre lectura y escritura en la universidad. Pereira: Universidad Tecnológica de Pereira.
- Cisneros, M., Olave, G. y Rojas, I. (2013a). Alfabetización académica y lectura inferencial. Bogotá: Ecoe.
- Cisneros, M., Olave, G. y Rojas, I. (2013b). Deserción universitaria y alfabetización académica. *Educación y Educadores*, 16(3), 455-471. Disponible en <http://educacionyeducadores.unisabana.edu.co/index.php/eye/article/view/2726>
- Cisneros, M., Olave, G. y Rojas, I. (2014). Cómo escribir la investigación. Desde el proyecto hasta la defensa. Bogotá: Ediciones de la U.
- Cisneros, M., Olave, G. y Rojas, I. (2016a). Didáctica de la lengua materna en Colombia. Currículos y visiones docentes. Pereira: Universidad Tecnológica de Pereira.
- Cisneros, M., Olave, G. y Rojas, I. (2016b). Hacia una desfragmentación de la lengua materna. Pereira: Universidad Tecnológica de Pereira.

- Cisneros, M., Olave, G. y Rojas, I. (2016c). Alfabetización académica y pedagogía de género discursivo en la Lingüística Sistemática Funcional. Una experiencia de trabajo. *Signos. Estudios de Lingüística*, 49(S1), 224-246. DOI: 10.4067/S0718-09342016000400011
- Coxhead, A. (2014). Identifying specialised vocabulary. En: A. Coxhead (ed.), *New ways in teaching vocabulary, revised* (pp. 251-252). Alexandria, VA: TESOL.
- Coxhead, A. (2016). Acquiring academic and disciplinary vocabulary. En: K. Hyland & P. Shaw (eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 177-190). London/ New York: Routledge.
- Garnier, B. y Guérin-Pace, F. (2010). *Appliquer les méthodes de la statistique textuelle*. París: UMR-CEPED.
- González, B. y Vega, V. (2013). Lectura y escritura en la educación superior colombiana: Herencia y deconstrucción. *Revista Interacción*, 12, 195-201.
- Gunnarsson, B. (2009). *Professional Discourse*. New York: Continuum.
- Halliday, M. & Hasan, R. (1976). *Cohesion in English*. London/New York: Longman.
- Hyland, K. (2004). *Disciplinary Discourses. Social Interactions in Academic Writing*. Michigan: The University of Michigan Press.
- Hyland, K. (2009). *Academic Discourse. English in a global context*. New York: Continuum.
- Hyland, K. & Tse, P. (2009). Academic Lexis and Disciplinary Practice: Corpus Evidence for Specificity. *International Journal of English Studies*, 9(2), 111-129.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27, 590-619.
- López Morales, H. (1984). *La enseñanza de la lengua materna. Lingüística para maestros de español*. Madrid: Editorial Playor.
- López Morales, H. (2011). Los índices de riqueza léxica y la enseñanza de lenguas. En: J. de Santiago Guervós, H. Bongaerts, J. Sánchez y M. Seseña (coords.), *Del texto a la lengua: La aplicación de los textos a la enseñanza-aprendizaje del español L2-LE* (pp. 15-28). Salamanca : Asociación para la enseñanza del español como lengua extranjera.
- Martínez, M. C. (2015[1994]). *Cohesión en español* (2 ed.). Cali: Univalle.
- Matsuda, K., Sadowsky, S. y Sabaj, O. (2011). Índice de Palabras de Contenido (IPC) y Distribución Porcentual de Legomena (DPL) en artículos de investigación en español. *Signos*, 45(78), 70-82. DOI: 10.4067/S0718-09342012000100005
- Montolío, E. (2001). *Conectores de la lengua escrita*. Barcelona : Ariel.
- Nation, I.S.P. (2013). *Learning vocabulary in another language* (3rd ed.). Cambridge: Cambridge University Press.
- Navarro, F., Ávila, N., Tapia, M., Cristovão, V., Moritz, M., Narváez, E. y Bazerman, Ch. (2016). Panorama histórico y contrastivo de los estudios sobre lectura y escritura en educación superior publicados en América Latina. *Signos. Estudios de Lingüística*, 49(S1), 100-126. DOI: 10.4067/S0718-09342016000400006
- Parodi, G. (ed.) (2008). *Géneros discursivos y géneros profesionales. Accesos discursivos para saber y hacer*. Valparaíso, Chile: Pontificia Universidad Católica de Valparaíso.

- Pélissier, D. (2016). Initiation à la lexicométrie. Approche pédagogique à partir de Iramuteq. Disponible en <https://presnumorg.hypotheses.org/>
- Perfetti, Ch. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11, 357-383. Disponible en <http://dx.doi.org/10.1080/10888430701530730>
- Pottier, B. (1977). *Lingüística general*. Madrid: Gredos.
- RAE (Real Academia Española) (2009). *Nueva gramática de la lengua española*. Tomo I: Morfología. Madrid: Autor.
- Reinert, M. (1993). Les "mondes lexicaux" et leur "logique" à travers l'analyse statistique d'un corpus de récits de cauchemars. *Langage et société*, 66, 5-39.
- Venegas R. (2007). Clasificación de textos académicos en función de su contenido léxico-semántico. *Signos*, 40(63), 239-271.
- WRAB (2017). Sitio web de la cuarta versión del Writing Research Across Borders. Febrero 15-18, Bogotá. Disponible en <http://wrab2017.com/javeriana/>