

## UN ESTUDIO LEXICOMÉTRICO A LA ESCRITURA DE LOS ASPIRANTES A UN POSTGRADO EN LINGÜÍSTICA<sup>1</sup>

Mireya Cisneros Estupiñán<sup>2</sup>

Universidad Tecnológica de Pereira (Colombia)

*mireyace@utp.edu.co*

Giohanny Olave Arias

Universidad Industrial de Santander (Colombia)

*giolavar@uis.edu.co*

María Alicia Serna Pinto

Universidad Tecnológica de Pereira (Colombia)

*maalserna@utp.edu.co*

**Recibido:** 03/10/2019 - **Aprobado:** 25/11/2019

**DOI:** doi.org/10.17533/udea.lyl.n77a02

**Resumen:** Se presentará una aproximación a la dimensión léxica de los textos escritos por aspirantes a la Maestría en Lingüística de la Universidad Tecnológica de Pereira. Se articularon aspectos de la disponibilidad léxica con estudios sobre géneros discursivos académicos y comunidades disciplinares. El corpus se constituyó por 81 textos presentados como requisito de admisión durante los años 2014-2015, analizados bajo cinco indicadores definidos para la delimitación del léxico académico. Los resultados preliminares indicaron una baja proximidad entre los significados asociados a los mundos lexicales, por lo que se sugiere articular las representaciones sobre la lengua y el discurso con la educación.

**Palabras claves:** léxico académico; tecnolecto; lexicometría; escritura en postgrado; comunidad disciplinar.

### A LEXICOMETRICAL STUDY OF THE APPLICANTS' WRITINGS FOR A POSTGRADUATE DEGREE IN LINGUISTICS

---

1. Artículo de investigación presentado como ponencia en el *III Congreso Internacional de Investigación Lingüística*, realizado en la Universidad de Antioquia entre el 4 y el 6 de septiembre de 2019.

2. Este artículo de investigación es una versión ampliada y actualizada de otra publicada en la Revista *Panorama Económico*, con el título *El léxico disciplinar en la educación universitaria*, Volumen 27, número 1, enero-marzo de 2019, pp. 249-266.

**Abstract:** An approximation to the lexical dimension of the texts written by candidates for the Master's Degree in Linguistics at the Technological University of Pereira will be presented. Aspects of lexical availability were articulated with studies on academic discursive genres and disciplinary communities. The corpus consisted of 81 texts presented as a requirement for admission during the years 2014-2015, analyzed under five defined indicators for the delimitation of the academic lexicon. The obtained results indicated a low proximity between the meanings associated with the lexical worlds, so it is suggested to link the representations about language and discourse with education.

**Key words:** academic lexicon; technolect; lexicometry; postgraduate writing; disciplinary community.

## 1. Introducción

Los estudios sobre la enseñanza de la lectoescritura en la Universidad constituyen actualmente un campo activo de investigación en Latinoamérica (Ávila, 2017; Natale & Stagnaro, 2018; Navarro, 2017, 2018; Navarro *et al.*, 2016; Trigos, 2019; WRAB, 2017). Diversas iniciativas en Colombia han contribuido significativamente en el área (Calle-Arango, 2018; González & Vega, 2013; Moreno, 2019; Rojas, 2017) y siguen explorándola con una diversidad de métodos en sus aproximaciones.

Desde el grupo Estudios del Lenguaje y la Educación, se añade a los aportes ya realizados en ese campo (Cisneros, Olave & Rojas, 2013a, 2013b, 2014, 2016c; Cisneros & Muñoz, 2014; Olave, 2019) una propuesta de trabajo centrada en el plano lexical del discurso académico universitario. Específicamente, se analizará la dimensión léxica en los textos escritos por aspirantes a la Maestría en Lingüística de la Universidad Tecnológica de Pereira; por tratarse de profesionales que han cursado carreras afines con la lingüística y la educación, los resultados pueden indicar el manejo académico y disciplinar del léxico en la construcción de textos escritos expositivos y argumentativos. El análisis realizado resulta relevante, ya que las fallas en el dominio del léxico se traducen generalmente en fallas conceptuales, lo que evidencia el no aprendizaje de los conceptos básicos en el área de profesionalización, situación que se torna problemática al considerar que los aspirantes ejercen labores de docencia en primaria y secundaria.

En este sentido, la formación universitaria profesional se fundamenta sobre una base lingüística representativa que constituyen los technolectos y las jergas profesionales. Así, pues, se entiende 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) que contribuyen a la conformación de comunidades disciplinares específicas (Hyland, 2004) dentro de las cuales

funcionan esos géneros. En consecuencia, 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, pp. 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 (Hyland & Tse, 2009).

Los tecnolectos (Montes, 1980) son variedades de la lengua que refieren el uso especializado del lenguaje al interior de una comunidad científico-académico-profesional. Como variante, el tecnolecto únicamente difiere de la variedad estándar y de los usos consuetudinarios de la lengua por la densidad del nivel léxico; esto significa que la lengua no solo sirve como medio de comunicación, sino también como un medio para conformar grupos sociales distintivos a partir del habla (Labov, 1983) o del uso concreto donde se privilegian unas formas lingüísticas sobre otras. Considerando lo anterior, la educación universitaria no puede obviar el fortalecimiento de los repertorios léxicos propios de cada comunidad científica, ya que el aprendizaje y el dominio adecuado de los conceptos y términos disciplinares le permitirán al profesional en formación acceder a estudios cada vez más especializados para impactar de manera positiva en su praxis laboral.

La noción de léxico académico presupone 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 supuesto, lo mismo que algunas aproximaciones desde la psicología cognitiva del desarrollo, como la hipótesis de la calidad lexical (Perfetti, 2007), los estudios dialectológicos sobre los tecnolectos (Suarez, 2003; Solís, 2005; Rumbos & Valles, 2008), los enfoques desde la lingüística de corpus (Matsuda, Sadowsky & 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.

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 facilitarán al profesional actuar y ser

reconocido como parte de una comunidad de especialistas. El uso del léxico especializado de la lingüística conforma un nivel dentro de la cultura disciplinar (Gunnarson, 2009; Hyland, 2004), es decir, forma parte de las prácticas que constituyen, a los profesionales egresados de carreras afines con el estudio del lenguaje, como miembros de una comunidad de profesionales encargada de investigar sobre la lengua, en los ámbitos teórico y aplicado.

En este marco académico, 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). Ese proceso de enculturación (Prior & Bilbro, 2011) dentro de cada disciplina, requiere identificar las múltiples y simultáneas funciones que cumple la lectoescritura en la educación superior (Navarro, 2018). En el caso del ingreso al postgrado en Lingüística, el desafío resulta mayúsculo, no solo por la activa dinámica del campo disciplinar, en constante demanda de actualización, sino 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.

Específicamente, en este artículo se presentarán cinco dimensiones de análisis definidas para la delimitación del léxico académico, que constituirán una ruta de trabajo a seguir: 1) medición de la riqueza léxica, 2) análisis de similitudes, 3) análisis terminológico con nubes de palabras, 4) análisis Reinert o de mundos lexicales y 5) análisis factorial de correspondencias. Además, se exponen resultados iniciales alcanzados con las herramientas lexicométricas provistas por el programa Iramuteq 0.7 alpha 2 (Pélissier, 2016).

## 2. Metodología

Para el desarrollo de este trabajo, se articularon aspectos de los estudios lexicométricos sobre disponibilidad léxica con los estudios sociorretóricos sobre géneros discursivos académicos y comunidades disciplinares. En principio, se utilizaron los índices cuantitativos de riqueza léxica (densidad, variedad y hápax), frecuencias léxicas, asociaciones semánticas y análisis factorial de coocurrencias.

Las pruebas que miden la riqueza léxica de los textos de un corpus son de naturaleza cuantitativa y comparativa (López Morales, 1984). Se tuvieron 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 contó el número total de palabras

(ocurrencias) y el de vocablos (palabras diferentes o formas léxicas) y se calculó la proporción entre ambos valores. 2) intervalo de aparición de las palabras nocionales, que consideró 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 aumentaba según disminuía el intervalo de aparición de lexías nocionales; en contraste, cuanto más amplio era 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 obtuvo aplicando la fórmula  $V$  (número de vocablos) /  $V1$  (suma de palabras de una ocurrencia). Para determinarlo, se dividió la cifra total de lexías por aquellas que solo se empleaban una vez en el texto.

Se utilizaron las herramientas de lexicometría automatizada, provistas por el programa Iramuteq 0.7 alpha 2. El programa apoyó la realización de análisis multidimensionales de textos, a partir de los fundamentos de la lexicometría y la textometría (Pincemin, 2010). Desarrollado por Pierre Ratinaud en el laboratorio de investigación LERASS, de la Universidad de Toulouse (Moreno & Ratinaud, 2015, p. 3), es compatible con todos los tipos de sistemas operativos, totalmente gratuito y descargable directamente desde su página web.

El corpus fue constituido por los textos argumentativos que los aspirantes a la Maestría en Lingüística de la Universidad Tecnológica de Pereira presentaron por correo electrónico, 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 solicitó 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 cuanto a la población, la mayoría de los aspirantes trabajaban como docentes de español o inglés en instituciones de educación básica y media, en correspondencia con su formación profesional de pedagogo.

### 3. Resultados

A continuación, se presentarán los resultados preliminares del análisis, elaborado a partir de los cinco indicadores analíticos del léxico disciplinar. Los indicadores fueron seleccionados a partir de los rasgos genéricos propios de la escritura para el postgrado y, en particular, en las

regularidades y regulaciones de los géneros textuales denominados como propuesta de investigación y ensayo reflexivo, según el nivel de maestría para el área de la lingüística.

### 3.1. Riqueza léxica

Los textos que contienen el mayor porcentaje de vocablos y el menor intervalo de palabras nocionales y de hápax fueron los clasificados como de mayor riqueza léxica, si bien dicha clasificación estuvo 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 (hápax):

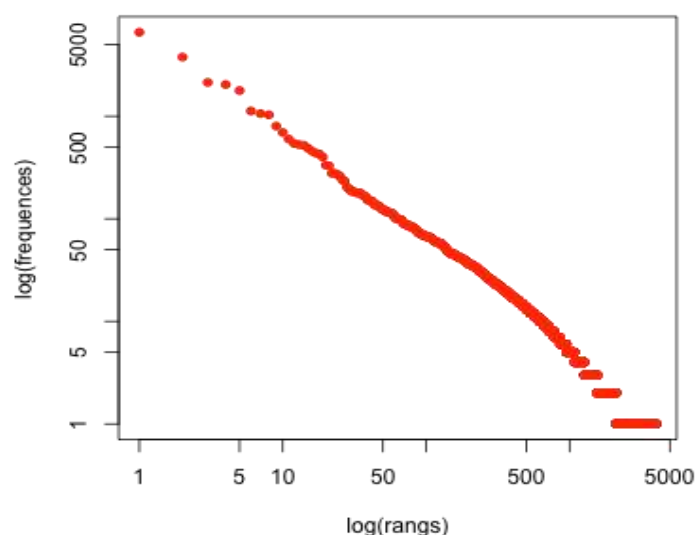


Figura 1. Logaritmos de frecuencia y de rango en el corpus de léxico disciplinar

- Número de textos: 81
- Número de ocurrencias: 56 318
- Media de ocurrencias por texto: 695,28
- Número de formas léxicas: 4 005
- Número de hápax: 1 898

— (3,37 % de ocurrencias y 47,39 % de formas léxicas)

La estadística textual indica que en 56 318 usos léxicos se utilizaron apenas 4 005 formas distintas, es decir, cada lexía nocional (nombres, adjetivos, verbos y adverbios) se utilizaron en un 7,11 % en el total de 81 textos, lo que constituye el porcentaje de vocablos. Este indicador grueso 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 indicó 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 lo señalaron Matsuda, Sadowsky y Sabaj (2011), el mayor porcentaje de hápax 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.

También, se pudieron 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 su nivel de riqueza léxica.

### ***3.2. Análisis terminológico con nubes de palabras***

A partir de tablas de frecuencias y nubes de palabras, se analizó la aparición de lexías repetidas en el corpus y se les representó con tamaños de letra distintos. Las tablas de frecuencia deben ser revisadas para identificar 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 «comodín» y de usos coloquiales, indican niveles más restringidos de léxico académico. La Figura 2 presentará la nube correspondiente al corpus de lexías activas, o no lematizadas.



Figura 2. Nube de palabras de lexías repetidas

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 ya se sabe, los primeros dos términos presentan diferencias semánticas sustanciales (Areiza, Cisneros & Tabares, 2019, pp. 11-13); por tanto, habrá que evaluar estos usos en cada contexto, para determinar si evidencian sus especificidades disciplinares. Este análisis cualitativo será tarea de nuevos trabajos donde se realicen análisis lexicométricos.

En la Tabla 1 se mostrará 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 56 318 ocurrencias):

	<b>Cat.</b>	<b>Frec.</b>	<b>Frec.</b>
<b>Forma léxica</b>	<b>gram.</b>	<b>absoluta</b>	<b>relativa</b>
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
			<hr/>
			<b>5,92</b>

Tabla 1. Frecuencias léxicas del 6 % del corpus

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 contexto 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*).

### 3.3. Análisis de similitudes léxicas

Se expondrán las relaciones entre las lexías activas del corpus, a través de un gráfico de red (Garnier & Guérin, 2010; Péliissier, 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 representarán las relaciones de similitud entre las frecuencias léxicas más altas del 6 % del corpus.

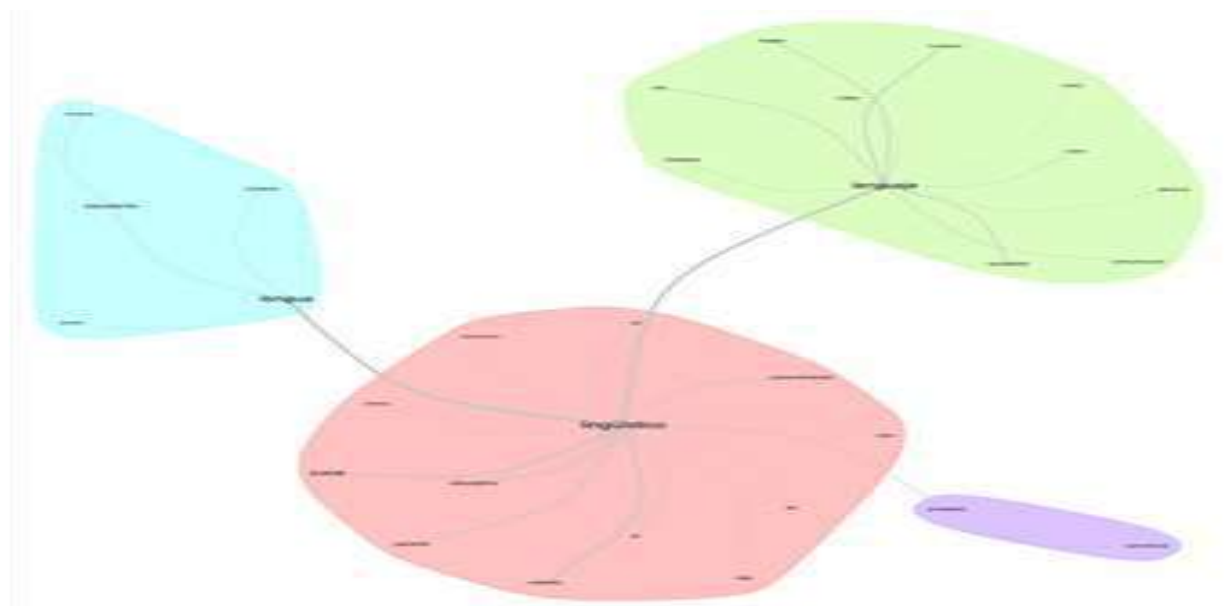


Figura 3. Relaciones de similitud léxica del 6 % del corpus

A través de los halos de color, el gráfico muestra los grupos léxicos asociados 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 (cuatro 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 —o 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 toma una muestra representativa del 6 % del corpus, la distribución asociativa se mantiene en los gráficos del corpus completo.

### **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 la prueba chi cuadrado que calcula, mediante un algoritmo, la correlación entre variables comparando la distribución observada de los datos con la distribución esperada de ellos. Las categorías aparecerán en dendogramas y representan agrupaciones de puntos de vista de los enunciadores con respecto a los temas sobre los cuales escriben. La Figura 4 ofrecerá, visualmente, la organización jerarquizada de los agrupamientos lexicales o clases de lexías, de acuerdo con sus mundos de referencia.

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

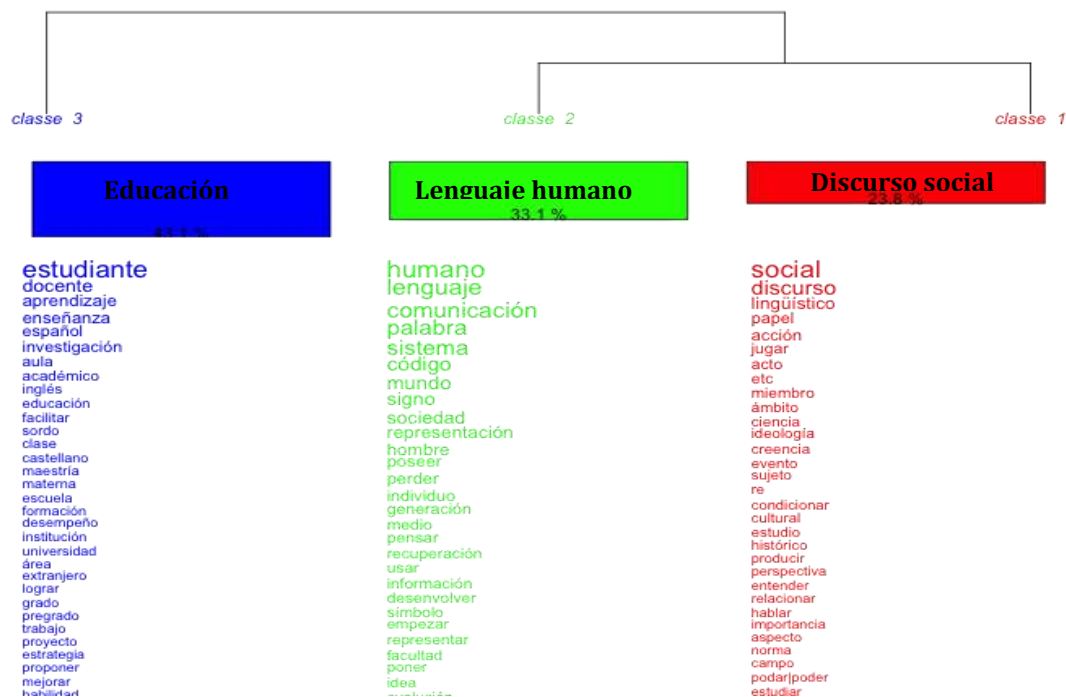


Figura 4. Agrupamientos léxicos jerarquizados

El análisis de los mundos léxicos muestra que, en el corpus, los tres campos de significado 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 de 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 & 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 léximas 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 léximas claves en cada núcleo de significado, como *discurso* (para la clase 1), *gramática* (para la clase 2) e *investigación* (para la clase 3).

### 3.5. Análisis factorial de correspondencias léxicas

Permitió analizar la aparición simultánea de lexías. La coincidencia entre ellas indicó cercanías entre significados, con las cuales se constituyeron agrupamientos y se confirmó así el análisis Reinert. Estos grupos se acercan o se alejan de otros grupos de lexías coocurrentes, todo lo cual ayudó a identificar las asociaciones semánticas que realizaron los informantes. En este nivel analítico del léxico se pudieron generar hipótesis interpretativas sobre el direccionamiento de las asociaciones semánticas y el tratamiento de los temas sobre los cuales se les solicitó escribir a los informantes.

En la Figura 5, 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 1 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. Así, una aproximación analítica a la Figura 5 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; se considera que estos significados, asumidos como representaciones disciplinares, lideran el tránsito hacia el acercamiento de la lingüística (de la lengua y del discurso) hacia el ámbito de lo educativo.

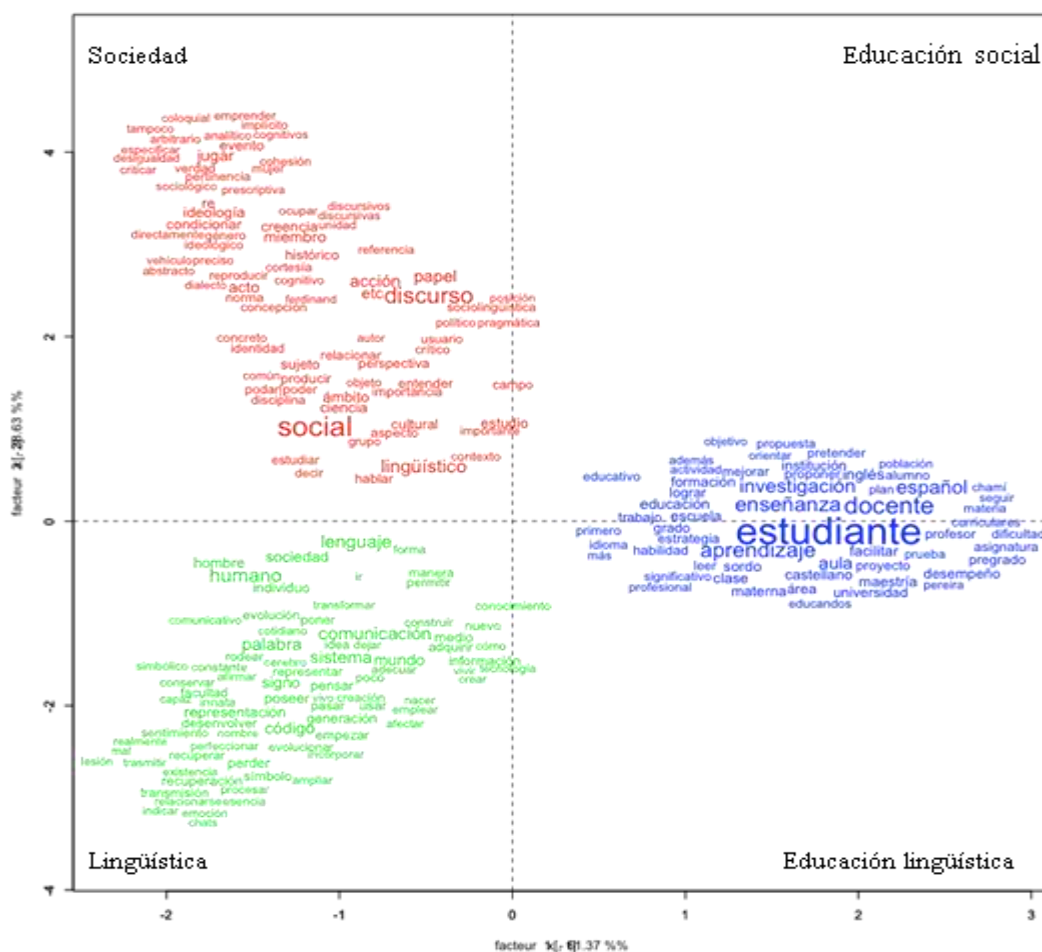


Figura 5. Análisis factorial de correspondencias

#### 4. Conclusiones

En los apartados anteriores, se presentaron cinco dimensiones de análisis o indicadores definidos para estudiar el léxico académico de un corpus de argumentaciones escritas por los aspirantes en el momento crucial de ingreso a cursar una Maestría en Lingüística, los cuales permitirán definir el modo de acceso de los estudiantes de carreras afines a la lingüística en materia de uso de la lengua con fines académicos. El trabajo representa 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 postgrado; especialmente de la formación postgradual en lingüística, por sus desafíos epistemológicos y metadiscursivos.

En general, en esta fase inicial de la investigación, los indicadores lexicométricos permitieron establecer las relaciones semánticas en torno a núcleos de significado. Los valores contrastivos entre los usos léxicos muestran redes conceptuales que actúan como marcos interpretativos y referenciales de la realidad, los cuales regulan el modo en que un sujeto percibe, interpreta y participa en los hechos del mundo.

De esta manera, los resultados parciales permiten establecer que el análisis del léxico es significativo para los procesos didácticos de la lengua en el nivel superior. Específicamente, la didáctica del léxico disciplinar toma el léxico como nivel para orientar la enseñanza, ya que se parte de la idea de que el léxico mental almacena las palabras por su similitud formal y semántica; a partir de ahí, se considera que, generalmente, se aprendan conceptos especializados con sus pares asociativos. En otros términos, las palabras se almacenan en forma de red y no como una lista de palabras aisladas; por ende, proveer los nodos de esa red y dotarla de un carácter sociocultural y disciplinar debe formar parte de la planeación de los currículos.

Según Bybee (1988), las palabras forman esquemas o modelos de identidad cuya fuerza depende del número de palabras que la integren y de la cantidad de rasgos compartidos entre ellas; de este modo, las palabras con conexiones más fuertes guardan una relación de proximidad en la memoria del hablante. Al respecto, Aitchison (1987) planteó la metáfora de la de la biblioteca para ilustrar que la recuperación de las palabras depende de su frecuencia de uso y de sus relaciones léxicas: así como en una biblioteca se ubican en las estanterías cercanas los libros que se usan con mayor frecuencia, si bien están organizados según algún criterio (materia, el autor, el año, etc.) sucede lo mismo con las formas léxicas.

En términos pedagógicos, el léxico no solo opera como fuente de consulta de vocabulario. Los mundos lexicales actúan como complejas asociaciones mentales donde residen los conceptos. En este sentido, dominar el léxico disciplinar equivale a dominar los conceptos disciplinares para ser utilizados adecuadamente en procesos investigativos y docentes.

Los vacíos léxicos en determinada red léxica, así como las asociaciones débiles entre las palabras, inciden de manera negativa en la educación especializada del postgrado, pues se afecta la posibilidad de reelaborar los esquemas mentales en el proceso de andamiaje de los saberes declarativos más complejos que, en última instancia, son los saberes que regulan la capacidad de ejercer los saberes procedimentales (Díaz & Hernández, 2010). Dado que la organización del léxico depende de las experiencias del sujeto, es imperativo realizar estudios de corte lexicométrico, con el fin de detectar las posibles fallas asociativas que afectan el

aprendizaje y la reconceptualización en determinado momento de la vida, esto es, la formación de ideas equivocadas o *misconceptions* (Lucariello & Naff, 2019, p. 31).

En el caso de los estudiantes de postgrado, las concepciones alternativas frente a los saberes disciplinares generan fallas conceptuales que se replican en sus escenarios laborales y debilitan los procesos de formación y producción científica en el nivel doctoral y posdoctoral. Por tanto, es imperativo incluir el análisis lexicométrico con el fin de establecer aquellos patrones que necesitan una reconfiguración durante la formación postgradual.

Los análisis lexicométricos expuestos anteriormente indican que existe una desconexión léxica entre la educación y la lingüística, la cual dificulta pensar en relaciones interdisciplinarias y repercute en el desarrollo profesional de los aspirantes. De acuerdo con los datos obtenidos, los aspirantes relacionan los conceptos de lenguaje y lengua mediante conexiones débiles y no por las razones semánticas que se esperarían en el ámbito especializado de sus carreras. Por tanto, la formación postgradual debe incluir el desarrollo léxico de los tecnolectos y el metalenguaje lingüístico.

En este artículo no se pretendió plantear los modos de realizar esas intervenciones pedagógicas o didácticas, pero los resultados preliminares del análisis permiten derivar recomendaciones para incrementar el desempeño escritural disciplinar en su dimensión léxica dentro de la formación profesional.

Se considera que, tal como se ha trabajado en la enseñanza de las lenguas extranjeras con fines específicos (Molina & Gentry, 2019), los ejercicios didácticos para el enriquecimiento léxico deben orientarse hacia la integración del contexto sociocultural dentro de las profesiones en las cuales se especializan los estudiantes de postgrado. Esta orientación involucra los intereses personales del estudiante y las demandas léxicas de los campos profesionales, de manera que se haga consciente el hecho de que profundizar en el conocimiento de una disciplina profesional implica adquirir modos específicos del decir y del escribir en dicho ámbito.

Sin embargo, es pertinente insistir en que abordar este nivel en el marco de los estudios sobre alfabetización académica es necesario, pero no suficiente. La presente contribución busca realizar una entrada al problema, focalizando la atención 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 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 dichos dominios en las escrituras al inicio del postgrado en Lingüística, más que leerse como una tendencia de los aspirantes a la Maestría en Lingüística, refleja dificultades para proponer investigaciones interdisciplinarias que enriquezcan tanto la ciencia lingüística como la educativa.

### Referencias bibliográficas

1. Adelstein, A. (2004). *Unidad léxica y valor especializado: Estado de la cuestión y observaciones sobre su representación*. (Tesis de doctorado), Universidad Pompeu Fabra, Barcelona.
2. Aitchinson, J. (1987). *Words in the Mind. An Introduction to the Mental Lexicon*. Cambridge: Blackwell.
3. Areiza, R., Cisneros, M. & Tabares, L. (2019). *Sociolingüística. Enfoques pragmático y variacionista*. 3ª edición. Bogotá: Ecoe.
4. Arnoux, E. N. de (Dir.) (2009). *Escritura y producción de conocimiento en las carreras de postgrado*. Buenos Aires/Santiago: Arcos.
5. 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>
6. Bazerman, C. (2012). *Géneros textuales, tipificación y actividad*. Puebla, México: Benemérita Universidad Autónoma de Puebla.
7. Bulté, B. & Housen, A. (2012) Defining and Operationalising L2 Complexity. En Housen, A., Kuiken, F., & Vedder, I. (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (pp. 21-46). Amsterdam: John Benjamins.
8. Bybee, J. (1988). Morphology as Lexical Organization. En Hammond, M. & Noonan, M. (Eds.), *Theoretical Morphology. Approaches in Modern Linguistics* (pp. 119-141). San Diego: Academic Press.
9. Calle-Arango, L. (2018). Prácticas de lectura en los centros y programas de escritura académica. *Teoría de la Educación. Revista Interuniversitaria*, 30(2), 155-175.

10. Carlino, P. (2013). Alfabetización académica diez años después. *Revista Mexicana de Investigación Educativa*, 18(57), 355-381.
11. Cisneros, M. & Muñoz, C. (2014). *Tras las huellas de las investigaciones sobre lectura y escritura en la universidad*. Pereira: Universidad Tecnológica de Pereira.
12. Cisneros, M., Olave, G. & Rojas, I. (2013a). *Alfabetización académica y lectura inferencial*. Bogotá: Ecoe.
13. Cisneros, M., Olave, G. & Rojas, I. (2013b). Deserción universitaria y alfabetización académica. *Educación y Educadores*, 16(3), 455-471. Recuperado de <http://educacionyeducadores.unisabana.edu.co/index.php/eye/article/view/2726>
14. Cisneros, M., Olave, G. & Rojas, I. (2014). *Cómo escribir la investigación. Desde el proyecto hasta la defensa*. Bogotá: Ediciones de la U.
15. Cisneros, M., Olave, G. & Rojas, I. (2016a). *Didáctica de la lengua materna en Colombia. Currículos y visiones docentes*. Pereira: Universidad Tecnológica de Pereira.
16. Cisneros, M., Olave, G. & Rojas, I. (2016b). *Hacia una desfragmentación de la lengua materna*. Pereira: Universidad Tecnológica de Pereira.
17. Cisneros, M., Olave, G. & Rojas, I. (2016c). Alfabetización académica y pedagogía de género discursivo en la lingüística sistémico-funcional. Una experiencia de trabajo. *Signos. Estudios de Lingüística*, 49(S1), 224-246. DOI: 10.4067/S0718-09342016000400011
18. Coxhead, A. (2014). Identifying Specialised Vocabulary. En Coxhead, A. (Ed.), *New Ways in Teaching Vocabulary, Revised* (pp. 251-252). Alexandria, VA: TESOL.
19. Coxhead, A. (2016). Acquiring Academic and Disciplinary Vocabulary. En Hyland, K. & Shaw, P. (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 177-190). London/New York: Routledge.
20. Díaz, F. & Hernández, G. (2010). *Estrategias docentes para el aprendizaje significativo: una interpretación constructivista*. México: Mc Graw Hill.
21. Garnier, B. & Guérin-Pace, F. (2010). *Appliquer les méthodes de la statistique textuelle*. París: UMR-CEPED.
22. González, B. & Vega, V. (2013). Lectura y escritura en la educación superior colombiana: Herencia y deconstrucción. *Revista Interacción*, 12, 195-201.
23. Gunnarsson, B. (2009). *Professional Discourse*. New York: Continuum.
24. Halliday, M. & Hasan, R. (1976). *Cohesion in English*. London/New York: Longman.

25. Hyland, K. (2004). *Disciplinary Discourses. Social Interactions in Academic Writing*. Michigan: The University of Michigan Press.
26. Hyland, K. (2009). *Academic Discourse. English in a Global Context*. New York: Continuum.
27. Hyland, K. & Tse, P. (2009). Academic Lexis and Disciplinary Practice: Corpus Evidence for Specificity. *International Journal of English Studies*, 9(2), 111-129.
28. Labov, W. (1983). *Modelos sociolingüísticos*. Madrid: Cátedra.
29. 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.
30. López Morales, H. (1984). *La enseñanza de la lengua materna. Lingüística para maestros de español*. Madrid: Editorial Playor.
31. López Morales, H. (2011). Los índices de riqueza léxica y la enseñanza de lenguas. En Santiago Guervós, J., Bongaerts, H., Sánchez, J. & Seseña, M. (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 -ASELE.
32. Lucariello, J. & Naff, D. (2019). *How do I Get My Students Over Their Alternative Conceptions (Misconceptions) for Learning? Removing Barriers to Aid in the Development of the Student*. Recuperado de <https://www.apa.org/education/k12/misconceptions.aspx>
33. Martínez, M. C. (2015[1994]). *Cohesión en español*. 2a. edición. Cali: Universidad del Valle.
34. Matsuda, K., Sadowsky, S. & 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
35. Molina, P. & Gentry, J. (2019). El enriquecimiento lexical en inglés con fines específicos en grupos universitarios de perfiles profesionales múltiples. *Transformación*, 15(2), 24-48.
36. Montes, J. (1980). Lengua, dialecto y norma. *Thesaurus*, xxxv, 237-257.
37. Montolío, E. (2001). *Conectores de la lengua escrita*. Barcelona: Ariel.
38. Moreno, E. (2019). Lectura académica en la formación universitaria: tendencias en investigación. *Lenguaje*, 47(1), 91-119.
39. Moreno, M. & Ratinaud, P. (2015). *Manual de uso de Iramuteq*. Recuperad de <http://www.iramuteq.org/documentation/fichiers/guia-iramuteq>

40. Natalie, L. & Stagnaro, D. (2019). *La lectura y la escritura en las disciplinas: lineamientos para su enseñanza*. Los Polvorines, (Argentina): Universidad Nacional de General Sarmiento.
41. Nation, I.S.P. (2013). *Learning Vocabulary in Another Language*. 3rd Edition. Cambridge: Cambridge University Press.
42. Navarro, F. (2017). Estudios latinoamericanos de la escritura en educación superior y contextos profesionales: hacia la configuración de un campo de un campo disciplinar propio. *Lenguas Modernas*, 50, 9-14.
43. Navarro, F. (2018). Más allá de la alfabetización académica: las funciones de la escritura en la educación superior. En Alves, M. A. & Iensen Bortoluz, V. (Eds.), *Formação de professores: Ensino, linguagens e tecnologias* (pp. 13-49). Porto Alegre: Editora Fi.
44. Navarro, F., Ávila, N., Tapia, M., Cristovão, V., et al. (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
45. Olave, G. (2019). Dimensión crítica de la reescritura académica. *Folios*, 50, 41-50.
46. Parodi, G. (Ed.) (2008). *Géneros disursivos y géneros profesionales. Accesos discursivos para saber y hacer*. Valparaíso, Chile: Pontificia Universidad Católica de Valparaíso.
47. Péliissier, D. (2016). *Initiation à la lexicométrie. Approche pédagogique à partir de Iramuteq*. Recuperado de <https://presnumorg.hypotheses.org/>
48. Perfetti, Ch. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11, 357-383. Recuperado de <http://dx.doi.org/10.1080/10888430701530730>
49. Pincemin, B. (2010). Semántica interpretativa y textometría. *Tópicos del Seminario*, 23, 15-55.
50. Pottier, B. (1977). *Lingüística general*. Madrid: Gredos.
51. Prior, P. & Bilbro, R. (2011). Academic Enculturation: Developing Literate Practices and Disciplinary Identities. En Castelló, M. (Ed.), *University Writing: Selves and Texts in Academic Societies* (pp. 19-32). London: Emerald.
52. RAE (Real Academia Española) (2009). *Nueva gramática de la lengua española. Tomo I: Morfología*. Madrid: Autor.
53. 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.

54. Rojas, I. (2017). Enseñar a leer y escribir en las disciplinas. Estado de la cuestión en las universidades colombianas. *Folios*, 45, 29-49.
55. Rumbos, H. & Valles, B. (2008). El tecnolecto de la terapia del lenguaje: aproximación a una comprensión epistemológica. *Revista Textura Centro de Estudios Textuales*, 7(10), 150-160.
56. Solís, N. (2005). *Análisis lingüístico de los anglicismos en el tecnolecto de la informática*. (Trabajo de grado inédito) Universidad Nacional Mayor de San Marcos.
57. Suárez, A. (2003). De léxicos, tecnolectos y otras jergas: Definiendo el vocabulario de la comunicación y el diseño. En XI *Jornadas de Reflexión Académica*. Universidad de Palermo.
58. Trigos, L. (2019). Una perspectiva sociocultural crítica a las literacidades académicas en América Latina. *Íkala, Revista de Lenguaje y Cultura*, 24(1),13-26.
59. Venegas R. (2007). Clasificación de textos académicos en función de su contenido léxico-semántico. *Signos*, 40(63), 239-271.
60. WRAB (2017). *Sitio web de la cuarta versión del Writing Research Across Borders*. 15-18 de febrero, Bogotá. Recuperado de <http://wrab2017.com/javeriana/>