

**ANÁLISIS DEL LÉXICO ACADÉMICO Y DISCIPLINAR EN LAS
TESIS DE LICENCIATURA DE CIENCIAS SOCIALES Y
HUMANIDADES DE 2007 AL 2014 DE LA UASLP, MÉXICO:
PRIMER ANÁLISIS**

ANALYSIS OF THE ACADEMIC AND DISCIPLINARY LEXICON IN
THE BACHELOR'S THESES OF SOCIAL SCIENCES AND
HUMANITIES FROM 2007 TO 2014 OF THE UASLP, MEXICO:
FIRST ANALYSIS

Marco Antonio Pérez Durán

Universidad Autónoma de San Luis Potosí, México

marco.duran@uaslp.mx

<https://orcid.org/0000-0003-0854-3109>

RESUMEN: Este trabajo se enmarca en los estudios léxico-estadísticos de carácter comparativo. Su objetivo es identificar las diez palabras más frecuentes del léxico académico y disciplinar presentes en las secciones de introducción y conclusión de 135 tesis de grado, correspondientes a cinco programas educativos de la Facultad de Ciencias Sociales y Humanidades de la Universidad Autónoma de San Luis Potosí (UASLP), México, elaboradas entre 2007 y 2014. Se emplea una metodología léxico-descriptiva y contrastiva, con un enfoque mixto, exploratorio, asincrónico y de corte transversal. Los resultados muestran que las diez unidades léxicas más frecuentes están marcadas por la presencia

de verbos, característicos del vocabulario académico, y sustantivos, propios del léxico especializado.

PALABRAS CLAVE: léxico académico, léxico disciplinar, licenciatura, tesis de grado, vocablo.

ABSTRACT: This study is framed within the field of comparative lexical-statistical research. Its objective is to identify the ten most frequent words from academic and disciplinary vocabulary founded in the introduction and conclusion sections of 135 undergraduate theses from five academic programs within the Faculty of Social Sciences and Humanities at the Universidad Autónoma de San Luis Potosí (UASLP), Mexico, produced between 2007 and 2014. A lexical-descriptive and contrastive methodology is used, with a mixed, exploratory, asynchronous, and cross-sectional approach. The results show that the ten most frequent lexical units are characterized by the presence of verbs, which are typical of academic vocabulary, and nouns, which are representative of specialized vocabulary.

KEYWORDS: academic lexicon, disciplinary lexicon, degree, degree thesis, word.

Recibido: 2 de agosto de 2024

Aceptado: 6 de junio de 2025

INTRODUCCIÓN

El dominio del léxico académico es una herramienta clave para la construcción de conocimiento disciplinar en textos universitarios. Sin embargo, su adquisición y uso en el contexto adecuado sigue siendo un desafío para los estudiantes. Analizar la forma cómo se manifiesta este léxico en los textos de titulación permite comprender el proceso de

formación lingüística en la educación superior. Este trabajo pertenece a una serie de publicaciones que forman parte del proyecto de investigación *Estudio del léxico académico de estudiantes universitarios de la Universidad Autónoma de San Luis Potosí, México* (PREDO, 2024¹). Su objetivo general consiste en identificar las diez palabras más frecuentes del léxico académico y disciplinar presentes en las secciones de introducción y conclusión de 135 tesis de grado, correspondientes a cinco programas educativos de la Facultad de Ciencias Sociales y Humanidades de la Universidad Autónoma de San Luis Potosí (UASLP), México, elaboradas entre 2007 y 2014.

El trabajo utiliza una metodología mixta-descriptiva aplicada a un corpus de 270 textos (135 por sección) que permite describir detalladamente los resultados obtenidos. La premisa de la que parte este estudio es que una parte del léxico académico y disciplinar encontrado en ambas secciones se aprende desde la educación básica o inicial, mientras que la otra se adquiere durante la educación superior. Este análisis busca aportar evidencia empírica relacionada con el uso real del léxico en las tesis de cada disciplina, lo que puede contribuir al diseño de estrategias orientadas a fortalecer la enseñanza del vocabulario en la educación superior. La pregunta de investigación que se estudia es: ¿cuál es el léxico académico más frecuente por licenciatura?

Es necesario recordar que la tesis de grado es un texto sofisticado que refleja un alto nivel de elaboración conceptual, en el que se sintetiza el dominio de conocimiento alcanzado por el estudiante y, también, se fortalece y renueva el conocimiento académico de un área de estudio. En ella, las teorías y los métodos involucrados tienen como base un léxico académico y disciplinar que se observa en dos secciones de la tesis: introducción y conclusión. En estas secciones se aprecian las marcas lingüísticas enunciativas de quienes han escrito el texto. Estas marcas se denominan *voz del emisor*. Según Tutin et al. (2017), en el apartado de introducción el emisor debe justificar el interés por el tema a estudiar y, en la conclusión, debe subrayar generalmente los aportes específicos del trabajo y resumir los resultados del estudio comprendido.

¹ Programa de Estímulos al Desempeño del personal docente de la Universidad Autónoma de San Luis Potosí.

El léxico con que se construye una tesis está directamente relacionado con factores pedagógicos (lectura, terminología y aplicación del léxico) que el estudiante ha constituido a lo largo de su vida escolar. La enseñanza del léxico especializado en la universidad se ha vuelto un reto para las instituciones y, principalmente, para los propios profesores, quienes no solo deben encargarse de transmitir el conocimiento de su materia, sino también acompañar al alumno en el dominio, uso y apropiación del compendio léxico y disciplinar necesario para expresar con precisión sus ideas a través de la escritura académica. No saber utilizarlo se centra en el bajo dominio sobre la escritura, la competencia léxica y el uso correcto de los géneros textuales, habilidades que el estudiantado no ha desarrollado durante su vida académica y que se observa con mayor claridad a medida que este alcanza el nivel superior. En este nivel, el reto en la universidad no es determinar las causas por las que el estudiantado sabe o no utilizar el discurso académico, sino en mejorar sus deficiencias, principalmente, aquellas relacionadas con la redacción (Cisneros y Olave, 2019; Parodi, 2005; Torres Perdígón, 2018; Venegas et al., 2013).

Este trabajo robustece la línea de investigación de corpus académico y disciplinar relacionado con el estudio del léxico extraído de las propias disciplinas (Parodi, 2005; Fernández-Silva y Rojas, 2015; Meza et al., 2022), en que el nivel léxico es relevante pues es la base del discurso especializado que permea la construcción de los textos académicos universitarios. En este ámbito se ubican los trabajos de López Ferrero (2002), Nazar (2011), Da Cunha y Montané (2019), Porrás-Garzón y Vidal-Sabanés (2019), Fregoso Peralta y Aguilar González (2019). Lo novedoso de este estudio es que, en la revisión hecha, no hay trabajos parecidos sobre léxico académico y disciplinar de dos secciones de las tesis de Ciencias Sociales y Humanidades de una universidad pública estatal.

1. MARCO TEÓRICO

Entendamos el léxico como el compendio de unidades que pertenecen a la lengua y el vocabulario al compendio de unidades que pertenecen al habla (Gómez Molina, 2004; Pérez, 2025). El léxico se conforma por el conjunto de lexías o unidades léxicas usadas

de forma recurrente en los géneros discursivos dentro del ámbito académico (Bazerman, 2012; Cisneros y Olave, 2019; Parodi, 2005). Este se diferencia del vocabulario coloquial, pues el primero se adquiere en un contexto escolarizado (Venegas et al., 2013; Tutin et al., 2017; Torres Perdigón, 2018), mientras que el segundo pertenece a la comunicación cotidiana. Así, conforme los estudiantes cambian de grado, se acrecienta su conocimiento del léxico académico. Cuando los estudiantes llegan al nivel superior, adquieren términos propios de su disciplina y crean un vocabulario disciplinar que les permitirá aprender de su comunidad discursiva y a manifestar sus conocimientos. En este sentido, el vocabulario que pertenece al léxico académico puede afinar y precisar su significado acorde con el área de estudio para expresar conceptos técnicos, semitécnicos y subtécnicos (Cisneros y Olave, 2019; Coxhead, 2016). Así, el vocabulario disciplinar se puede ver como una subespecialización del vocabulario académico, donde el sentido de un término puede cambiar de acuerdo con la disciplina científica que hace uso de este (Pérez y Rojas, 2023).

Desde una perspectiva lingüística, la organización del vocabulario en léxico académico y léxico disciplinar se constituye por tres aspectos: lo sintáctico, lo semántico y lo pragmático, es decir, por la estructura oracional que es el punto de transmisión y contextualización de este vocabulario; por lo semántico, que apunta a la especificidad del significado de los términos; y por el sentido pragmático, en el que el término se especializa según los requerimientos de la disciplina que hace uso de este.

La escritura es una habilidad que se constituye de otras habilidades básicas socio-discursivas, cognitivas y lingüísticas (Rojas Corona et al., 2024) que se fortalecen de forma transversal a lo largo de la formación académica. Dentro de las habilidades lingüísticas destaca el dominio de léxico académico y disciplinar, mientras que, dentro de lo socio-cultural, se encuentra el reconocimiento del género académico que el escritor debe producir. La tesis de grado es parte de este tipo de géneros académicos y, en sus secciones de introducción y conclusión, se identifica un aspecto relevante para este estudio: la voz del emisor.

El concepto de voz del emisor se refiere a la manera como el escritor manifiesta su postura frente al tema que desarrolla (Castello et al., 2011). Cuando el emisor presenta sus ideas hace uso de verbos propios no tomados de otras fuentes, aunque mantienen un diálogo con otras fuentes. Al hacer esto, el texto que presenta evidencia del léxico-académico y disciplinar que realmente conoce y del que dispone. Este léxico real puede evidenciar la apropiación conceptual que el estudiante ha hecho de su disciplina y de su capacidad para producir textos.

Dentro de las secciones que integran la tesis, la introducción y la conclusión son dos apartados en que se puede escuchar la voz de emisor. En la introducción, el emisor presenta el fenómeno que va a estudiar, argumenta sobre la relevancia del estudio y su punto de vista en relación con este, expresa sus objetivos y realiza una descripción de las secuencias del proyecto que se elabora. En el caso de la conclusión, se toman de base las reflexiones que se hicieron durante el análisis y se observa si se respondió o no a cada objetivo planteado. La importancia de estas dos secciones se ciñe en el léxico empleado, la voz de quien ha escrito el texto y la cantidad de referencias bibliográficas que se usan para justificar la naturaleza del trabajo.

2. METODOLOGÍA

Esta investigación se enmarca en los estudios léxicos de corte descriptivo contrastivo con enfoque mixto (cuantitativo-cualitativo) de tipo exploratorio, asincrónico y transversal. El foco está en la obtención y análisis del léxico académico y disciplinar que aparece en la introducción y conclusión de las tesis de grado de cinco licenciaturas de Ciencias Sociales y Humanidades de 2007 al 2014 de una universidad pública y estatal de México. Para el análisis, y siguiendo la tradición de análisis de los estudios lexicométricos, se toma el *lemma-type* como la unidad léxica de análisis que cumple la función de organizar la información en las listas de las palabras de las licenciaturas (Lara, 2006; Pérez y Rojas, 2023).

2.1. Tipo de análisis

En este estudio se llevan a cabo dos tipos de análisis: cuantitativo y cualitativo. En el análisis cuantitativo del corpus se examinan las siguientes unidades lingüísticas: *tokens*, *types* y *lema-types*. Para ello, se elaboran tablas que muestran la distribución de estos elementos en las secciones de introducción y conclusión de las tesis de licenciatura. Además, se clasifica el vocabulario del corpus según su pertenencia al léxico académico y disciplinar.

Por su parte, el análisis cualitativo se centra en la identificación y obtención de las diez unidades léxicas más frecuentes en las tesis de grado, tanto a nivel global como por secciones. Estas unidades se contrastan con dos corpus de referencia: *Las 5,000 palabras más frecuentes en los libros de texto oficiales de la educación básica en México* (Varela et al., 2013) y la lista de “A New Academic Word List” elaborada por Coxhead (2000).

De manera específica, se emplea el corpus de Varela et al. (2013) debido a que su léxico académico se extrae de 34 textos oficiales vigentes en el sistema educativo mexicano hasta 2012. Este corpus representa un referente valioso para responder a la premisa central de esta investigación: identificar el léxico académico inicial adquirido en los primeros años de escolarización. Asimismo, se utiliza la lista de Coxhead (2000) como parámetro para determinar si las diez palabras presentes en el corpus de tesis de grado corresponden a dicho léxico especializado.

2.2. Muestra

La muestra de esta investigación está compuesta por los apartados de introducción y conclusión de 135 tesis de la licenciatura. Del total de tesis analizadas, 37 pertenecen a la carrera de Antropología, 5 a Arqueología, 45 a Geografía, 43 a Historia y 5 a Lengua y Literatura. En total se recopilaron y analizaron 270 textos (135 en la sección de introducción y 135 en la sección de conclusión). En este trabajo, se obtuvieron 441,836 *tokens*, 31,978 *types* y 4,170 *lemas-type* para el análisis.

2.3. Obtención y anonimización de la información

Para utilizar la información que sirve de base para este artículo se solicitó por escrito al Centro de Información en Humanidades, Bibliotecología y Psicología (CIHByP) del Campus Oriente de la Universidad Autónoma de San Luis Potosí² el permiso correspondiente. En este escrito, se indicó el objetivo de la investigación y se evidenció que los datos personales que rigen los proyectos de tesis se anonimizaron y quedaron resguardados de acuerdo con el reglamento sobre el uso de datos³.

2.4. Software utilizado

Para analizar los datos se utilizó el programa AntConc⁴, herramienta computacional especializada en el procesamiento de corpus textuales extensos. Este software permite extraer información lingüística que sirve de base para los objetivos de esta investigación. Una vez conformado el corpus, AntConc genera una lista ordenada por frecuencia léxica, lo cual facilita la identificación y lematización de los vocablos más relevantes.

Los vocablos obtenidos se clasificaron en tres categorías: generales, académicos y especializados. El análisis del corpus se llevó a cabo en dos fases: en primer lugar, mediante el uso del software, y posteriormente, a través de una revisión manual detallada, cuya ejecución demandó aproximadamente 18 meses de trabajo continuo.

2.5. Procedimiento

Para obtener el corpus, se llevaron a cabo las siguientes acciones:

- 1) Se solicitó permiso a CIHByP de la Universidad Autónoma de San Luis Potosí. Una vez autorizado el uso de la información, se encontró con que muchas tesis no estaban en formato PDF, por lo que se tuvieron que convertir a este formato.

² Se agradece al CIHByP por el uso de la información que se ha utilizado en este y en otros trabajos.

³ Este se puede consultar en <https://www.psicologia.uaslp.mx/Paginas/Servicios/3737#gsc.tab=0>

⁴ Software de acceso libre, disponible en: <https://www.laurenceanthony.net/software/antconc/>

- 2) Se elaboró el primer conteo de las tesis que estaban a disposición de esta Facultad. De este conteo, se tomaron 135 tesis para el análisis, en ellas se ubicaron las dos secciones que sirven de base para el estudio. Inmediatamente después, se llevó a cabo el proceso editorial que consistió en eliminar citas, nombres, fechas de los contenidos, imágenes utilizadas y parafraseos.
- 3) Se utilizó el software AntConc en dos ocasiones para la organización del corpus. La primera vez fue para buscar los *types* y la segunda para clasificar los *lemas-type* por cada licenciatura.
- 4) Se procesó el corpus y después se ordenaron los listados manualmente en Excel. Esto ayudó a contar con un análisis más preciso de la información. Una vez depurada la base de datos se cargó un nuevo archivo a AntConc para obtener el corpus final. Los resultados de este nuevo procesamiento ayudaron a mejorar la organización inicial de la información.

3. ANÁLISIS

3.1. Análisis cuantitativo

En este trabajo, se obtuvieron 441,836 *tokens*, 31,978 *types* y 4,170 *lemas-type*. En relación con su distribución, se presenta la siguiente Tabla 1:

Tabla 1. *Tokens, types y lemas-type* obtenidos en la introducción y en la conclusión de las tesis por licenciatura.

Lic.	To.	Introducción		Conclusión	
		Ty.	L-T	Ty.	L-T
An.	165,715	5,192	569	3,416	324
Arq.	23,784	1,839	369	2,051	433
G.	88,275	3,788	489	2,958	442
H.	143,150	4,911	440	3,282	389

LyL	20,912	2,409	332	2,132	283
Total	441,836	18,139	2,199	13,839	1,871

Nota: An. Antropología; Arq. Arqueología; G. Geografía; H. Historia; LyL Lengua y Literatura. To. *Tokens*; Ty. *Types* y L-T. *Lema-Type*. **Fuente:** Elaboración propia.

En relación con la información presentada en la Tabla 1, las licenciaturas en Antropología e Historia ocuparon las primeras posiciones en cuanto al número total de *tokens*, ambas superando los 100,000. Este resultado se explica, en parte, por el mayor número de tesis consideradas en estas dos licenciaturas en comparación con las demás. En detalle, la licenciatura en Antropología acumuló un total de 165,715 *tokens*, derivados del análisis de 37 tesis, lo que representa el 37.5 % del total de la muestra. Por su parte, la licenciatura en Historia generó 143,150 *tokens* a partir de 43 tesis, equivalentes al 32.3 % de la muestra. Ambas licenciaturas representan el primer y segundo lugar del corpus. Estos datos reflejan una tendencia clara: a mayor número de participantes, mayor es la cantidad de *tokens* obtenidos.

En cuanto al número total de *types* registrados en el corpus, se contabilizaron 31,978, de los cuales 18,139 corresponden a la sección de introducción y 13,839 a la de conclusión. Esta diferencia representa un 13.4 % a favor de la introducción, equivalente a 4,300 *types* adicionales. La licenciatura en Antropología es la que genera la mayor cantidad de *types* en ambas secciones, mientras Arqueología presenta la cifra más baja. En lo que respecta a los *lema-types*, el corpus totaliza 4,170, distribuidos en 2,199 en la introducción y 1,871 en la conclusión.

Al comparar ambas secciones, la diferencia registrada entre *lemas-type* es de 328 más para la introducción. La licenciatura en Antropología obtuvo más *lemas-type* en la introducción, mientras que Geografía ocupó la primera posición en la conclusión. Un hallazgo importante es que la licenciatura en Arqueología apareció en segundo lugar en *lemas-type* de esta sección. Si comparamos Geografía con Arqueología la diferencia es menor a 20 respuestas.

Sobre la distribución de los vocablos generales (VG), vocablos académicos (VA) y vocablos especializados (VE) se presenta la Tabla 2.

Tabla 2. Distribución de los vocablos (académico o especializado) por sección.

	Antropología		Arqueología		Geografía		Historia		LyL	
	I	C	I	C	I	C	I	C	I	C
VG	2520	2197	630	1258	1935	1906	2500	2336	1000	1182
VA	154	112	25	78	146	107	175	128	114	109
VE	86	57	21	35	61	47	138	119	61	60
4,170	2760	2366	676	1371	2142	2060	2813	2583	1175	1351

Nota: I= Introducción y C= Conclusión. **Fuente:** Elaboración propia.

Al llevar a cabo el análisis por sección, en la introducción, las licenciaturas en Historia y Antropología ocuparon los dos primeros lugares con 2,813 y 2,760 vocablos respectivamente. En contraste, las licenciaturas en Arqueología y Lengua y Literatura fueron las que menos vocablos registraron, con 1,175 y 676, respectivamente. Sobre la conclusión, Historia y Antropología también ocuparon las dos primeras posiciones con 2,366 y 2,583 vocablos, mientras que Lengua y Literatura se ubicó en el último lugar con 1,351 vocablos. Al comparar esta última con las dos primeras licenciaturas, se observa una diferencia del 50 %, es decir, Historia y Antropología acumularon un 50% más de vocablos que la licenciatura en Lengua y Literatura.

En el análisis del vocabulario general (VG), Antropología e Historia ocuparon los dos primeros lugares, con 2,520 y 2,500 vocablos, tanto en la introducción como en la conclusión. La licenciatura que generó menos vocabulario general fue Arqueología en la introducción, con 630 vocablos; mientras que, en la conclusión, Lengua y Literatura obtuvo la menor cantidad, con 1,182 vocablos. En cuanto al vocabulario académico (VA), Historia y Antropología obtuvieron las primeras posiciones, con 154 y 146 respuestas,

respectivamente, en la introducción y la conclusión. En contraste, Arqueología se ubicó en el último lugar, con solo 25 respuestas en ambas secciones.

3.2. Análisis cualitativo

Se presenta la información de los diez vocablos más frecuentes del corpus general en la Tabla 3:

Tabla 3. Los diez vocablos más frecuentes obtenidas del corpus general.

	Corpus General	Introducción	Conclusión
1	Conocer	Conocer	Conocer
2	Construir	Construir	Producir
3	Producir	Producir	Construir
4	Migración	Migración	Historia
5	Geografía	Analizar	Migración
6	Antropología	Publicación	Plantear
7	Publicación	Historia	Definir
8	Historia	Definir	Observar
9	Concluir	Plantear	Proponer
10	Agricultura	Geografía	Referir

Fuente: Elaboración propia.

En esta parte del análisis se presentan las 30 palabras divididas en tres grupos: corpus general, introducción y conclusión de la muestra general. Al contrastar y depurar las respuestas, quedaron 17, de las cuales seis pertenecen a vocablos nominales (*migración, geografía, antropología, publicación, historia, agricultura*) y 11 a vocablos verbales (*analizar, definir, plantear, referir, concluir, observar, proponer, etc.*). De entrada, se aprecia entre la

muestra general una mayor aparición de verbos que forman parte del vocabulario académico. Se toman las 10 palabras más frecuentes (50 por las cinco licenciaturas) en la Tabla 4:

Tabla 4. Respuestas obtenidas del corpus académico por licenciatura.

	Antropología	Arqueología	Geografía	Historia	LyL
1	Conocer	Conocer	Producir	Conocer	Conocer
2	Migración	Construir	Geografía	Construir	Narrar
3	Construir	Definir	Migración	Historia	Historia
4	Producir	Producir	Ser	Publicación	Definir
5	Analizar	Geografía	Construir	Analizar	Hispania
6	Definir	Antropología	Ver	Producir	Publicación
7	Nacional	Arqueología	Antropología	Abordar	Referir
8	Documento	Etnia	Emplear	Hispania	Analizar
9	Plantear	Analizar	Nacional	Plantear	Desarrollar
10	Publicación	Aportación	Plantear	Migración	Describir

Fuente: Elaboración propia.

Al clasificar la información de la tabla anterior, hay 15 verbos: *conocer, construir, producir, analizar, definir, plantear, ser, ver, emplear, abordar, conocer, narrar, referir, desarrollar y describir*, que se obtuvieron al depurar la muestra; 10 nominales entre las respuestas frecuentes: *migración, documento, publicación, Geografía, Antropología, Arqueología, etnia, aportación, historia, Hispania*); y un adjetivo: *nacional*. Sobre las diez respuestas frecuentes de tipo académico y especializado se presenta la Tabla 5:

Tabla 5. Respuestas frecuentes de la muestra general, organizadas por vocabulario académico (VA) y especializado (VE) de las licenciaturas.

Licenciatura	VA	VE
Antropología	Conocer	Migración
	Construir	Etnia
	Producir	Nacional
	Analizar	Historia
	Describir	Territorio
	Ser	Antropología
	Referir	Hídrico
	Definir	Indígena
	Estudiar	Biología
	Documento	Corresponsal
Arqueología	Ser	Geografía
	Conocer	Historia
	Construir	Arqueología
	Definir	Antropología
	Describir	Hídrico
	Producir	Región
	Estudiar	Biología
	Abordar	Edafología
	Pictografía	Erosión
	Publicación	Etnia
Geografía	Conocer	Geografía
	Producir	Migración
	Ser	Antropología
	Construir	Hídrico

	Estudiar	Nacional
	Plantear	Población
	Definir	Región
	Describir	Historia
	Publicación	Meteorología
	Abordar	Agrícola
Historia	Conocer	Historia
	Construir	Migración
	Estudiar	Hispania
	Publicación	Nacional
	Ser	Reforma
	Analizar	Regir
	Definir	Indígena
	Abordar	Ley
	Plantear	Revolución
	Producir	Eclesial
Lengua y Literatura	Ser	Historia
	Conocer	Hispania
	Narrar	Mitología
	Definir	Literatura
	Estudiar	Poema
	Publicación	Crónico
	Referir	Indígena
	Analizar	Lingüístico
	Describir	Nacional
	Producir	Personaje

Fuente: Elaboración propia.

En el análisis general, el vocabulario académico asciende a 50 respuestas. De estas, 15 son distintas. Según su clasificación, hay tres nominales (*publicación, pictografía y documento*) y 12 verbos (*conocer, construir, producir, analizar, describir, ser, referir, definir, estudiar, abordar, plantear y narrar*). De estos, *conocer, construir, producir, analizar, describir, ser, referir, definir, estudiar, abordar y plantear* aparecen repetidos en las listas y, de ellos, la voz *definir* es la única respuesta que apareció en las cinco listas. Sobre los nominales, la voz *publicación* no se encontró en las respuestas de Antropología.

En relación con el vocabulario especializado, de las 50 respuestas 27 son diferentes. El 90% de las respuestas en esta sección son nominales (*migración, etnia, historia, territorio, antropología, hídrico, territorio, antropología, indígena, biología, geografía, arqueología, región, edafología, erosión, publicación, meteorología, agrícola, Hispania, reforma, ley, revolución, mitología, literatura, poema, lingüística y personaje*), a diferencia de la lista anterior que eran verbos. En la lista de vocabulario disciplinar, el único verbo que apareció es *regir* en la lista de Historia. *Nacional, corresponsal, eclesial y crónico* son las 4 respuestas de categoría adjetival que surgieron. De esta lista, *historia, nacional, migración, hídrico, indígena y región* aparecieron en más de una ocasión en las respuestas de las licenciaturas.

De nueva cuenta, se toman las 50 palabras frecuentes y generales del corpus de tesis y se comparan con “Las 5,000 palabras más frecuentes en los libros de texto oficiales de la educación básica en México” (Varela et al., 2013) y la lista de Coxhead, “A New Academic Word List” (2000), constituida por las 2,000 palabras más frecuentes del léxico académico inglés.

Los resultados indican que solo 25 palabras aparecieron en el corpus de las 5,000 palabras más frecuentes. En cuanto a su distribución, hay 10 nominales, 14 verbos y 1 adjetivo. *Migración (21), documento (49), sexo (50), geografía (180), historia (411), símbolo (49), método (17), población (628), agricultura (167) y religión (94)* son los nominales que se reiteran en el corpus de 5,000 palabras. Los nominales *método (17)* y *migración (21)* son los vocablos que aparecieron entre las 30 palabras más frecuentes del corpus de las 5,000, mientras

que *agricultura* (167), *historia* (411) y *población* (628) se consideran como las respuestas más alejadas del corpus que se contrasta.

El número de verbos en el corpus de las 5,000 palabras asciende a 14, entre los cuales se encuentran *conocer* (212), *construir* (123), *producir* (104), *analizar* (22), *emplear* (16), *cambiar* (88), *describir* (24), *observar* (127), *ser* (865), *ver* (445), *organizar* (79), *desarrollar* (52), *escribir* (328) y *exponer* (34). Los verbos *emplear* (16), *analizar* (22), *describir* (24) y *exponer* (34) se ubican entre las 50 palabras frecuentes y los verbos *conocer* (212), *escribir* (328), *ver* (495) y *ser* (865) aparecieron alejados de las primeras posiciones del corpus de las 5,000 palabras. Sobre el adjetivo *nacional* (214), aparece entre las primeras posiciones del corpus de tesis y, al compararlo con el corpus de las 5,000 palabras frecuentes, este se ubica fuera de las primeras 50 palabras del corpus de tesis.

Al comparar las palabras frecuentes de cada licenciatura con el corpus de “Las 5,000 palabras” (Varela et al., 2013), la licenciatura en Geografía coincidió con 13 voces (*producir, geografía, migración, ser, construir, ver, emplear, nacional, población, agricultura, analizar, documento, organizar*), Arqueología con 12 (*conocer, construir, producir, geografía, cambiar, describir, historia, observar, símbolo, documento, método y publicación*), Antropología con 9 palabras (*conocer, migración, construir, producir, analizar, nacional, documento, emplear y sexo*) y, en último lugar, se encuentran Historia (*conocer, construir, historia, analizar, producir, migración, organizar, observar*) y Lengua y Literatura (*conocer, historia, analizar, desarrollar, describir, producir, escribir, exponer*) con 8 palabras de 20.

Se toman otra vez estas 50 palabras frecuentes del corpus de tesis y se contrastan con la lista de Coxhead (2000) sobre lenguaje académico y especializado. Los resultados indican que solo 15 palabras aparecieron en esta lista: 10 verbos (*analizar, cambiar, citar, concluir, conocer, construir, definir, exponer, seleccionar y ser*) y 5 nominales (*método, publicación, región, sexo y símbolo*). En cuanto a la distribución de las respuestas, los verbos *analizar* y *construir* aparecieron en todas las licenciaturas, mientras que los verbos *conocer* y *definir* se ubicaron en cuatro licenciaturas: Antropología, Arqueología, Historia y LyL. El resto de los verbos aparecieron en una sola ocasión, el verbo *cambiar* solo apareció como

respuesta de Arqueología, los verbos *concluir* y *seleccionar* forman parte de las respuestas de Historia, el verbo *ser* apareció en Geografía y *exponer* formó parte de LyL. Sobre los nominales, los vocablos *método* y *símbolo* aparecieron como respuesta de Arqueología, los vocablos *publicación* y *región* se ubicaron como respuestas de Geografía y el vocablo *sexo* apareció en Antropología.

De manera puntual, se presentan las tablas de distribución de las diez palabras frecuentes que aparecieron en cada sección de la tesis y clasificadas por vocabulario académico y disciplinar. En la Tabla 6 se muestran los resultados de Antropología, en la Tabla 7 los de Arqueología, en la Tabla 8 los de Geografía, en la Tabla 9 los de Historia y en la Tabla 10 los de Lengua y Literatura:

Tabla 6. Respuestas del vocabulario académico y especializado en la introducción y conclusión de las tesis de Antropología.

Antropología			
Introducción		Conclusión	
VA	VE	VA	VE
Conocer	Migración	Conocer	Migración
Analizar	Etnia	Ser	Etnia
Ser	Nacional	Construir	Nacional
Construir	Historia	Producir	Antropología
Describir	Antropología	Estudiar	Corresponsal
Producir	Hídrico	Describir	Ecología
Referir	Indígena	Educación	Historia
Definir	Territorio	Referir	Indígena
Estudiar	Biología	Símbolo	Población
Publicación	Geografía	Definir	Territorio

Fuente: Elaboración propia.

Tabla 7. Respuestas del vocabulario académico y especializado en la introducción y conclusión de las tesis de Arqueología.

Arqueología			
Introducción		Conclusión	
VA	VE	VA	VE
Ser	Geografía	Ser	Geografía
Describir	Historia	Conocer	Arqueología
Conocer	Arqueología	Construir	Biología
Construir	Región	Definir	Prehispánico
Estudiar	Prehispánico	Describir	Antropología
Abordar	Antropología	Producir	Caxcan
Definir	Erosión	Concluir	Dendrítico
Producir	Hídrico	Difusión	Edafología
Cuestionar	Lítico	Estudiar	Erosión
Pictografía	Topografía	Pictografía	Etnia

Fuente: Elaboración propia.

Tabla 8. Respuestas del vocabulario académico y especializado en la introducción y conclusión de las tesis de Geografía.

Geografía			
Introducción		Conclusión	
VA	VE	VA	VE
Conocer	Geografía	Conocer	Migración
Producir	Migración	Producir	Geografía
Ser	Hídrico	Construir	Población
Estudiar	Nacional	Plantear	Hídrico

Plantear	Antropología	Oscilamiento	Nacional
Construir	Región	Concluir	Agrícola
Definir	Historia	Recabar	Forestal
Publicación	Meteorología	Cuestionar	Región
Abordar	Población	Imperar	Demográfico
Describir	Biología	Implementar	Ejido

Fuente: Elaboración propia.

Tabla 9. Respuestas del vocabulario académico y especializado en la introducción y conclusión de las tesis de Historia.

Historia			
Introducción		Conclusión	
VA	VE	VA	VE
Conocer	Historia	Conocer	Historia
Construir	Migración	Construir	Hispania
Analizar	Hispania	Plantear	Revolución
Estudiar	Nacional	Publicación	Imperio
Publicación	Reforma	Analizar	Reforma
Ser	Regir	Observar	Regir
Abordar	Revolución	Producir	Borbón
Plantear	Eclesial	Cuestionar	Eclesial
Definir	Geografía	Organizar	Hegemónico
Organizar	Indígena	Concluir	Indígena

Fuente: Elaboración propia.

Tabla 10. Respuestas del vocabulario académico y especializado en la introducción y conclusión de las tesis de Lengua y Literatura.

Lengua y Literatura			
Introducción		Conclusión	
VA	VE	VA	VE
Ser	Historia	Conocer	Historia
Narrar	Hispania	Narrar	Mitología
Conocer	Literatura	Construir	Hispania
Estudiar	Crónico	Producir	Crónico
Publicación	Lingüístico	Referir	Demográfico
Definir	Poema	Transgredir	Literatura
Describir	Biología	Cuestionar	Poema
Analizar	Bélico	Definir	Bélico
Producir	Personaje	Argumento	Indígena
Citar	Arquetipo	Concluir	Nacional

Fuente: Elaboración propia.

Del análisis de las tablas anteriores, tanto en la introducción como en la conclusión el número de respuestas ascendió a 50 respuestas, 10 por cada licenciatura. En el análisis de la sección introducción, hay 19 respuestas distintas pertenecientes al vocabulario académico: 17 verbos y 1 nominal; mientras que, en la sección de la conclusión, hay 27 respuestas de 50. En esta parte, 21 son verbos y 6 nominales. Las respuestas verbales que se repiten en las listas en la introducción del vocabulario académico fueron: *conocer, analizar, ser, construir, producir, definir, estudiar, describir y abordar*. En las respuestas nominales, solo apareció *publicación* en varias ocasiones. En la conclusión, *conocer, construir, producir y concluir* fueron los verbos que aparecieron en las listas, mientras que *argumento, publicación, oscilamiento, difusión, pictografía y educación* pertenecen a esta sección.

En el análisis del vocabulario disciplinar de la introducción, aparecieron 25 respuestas de 50 posibles (10 por cada lista). Esto representa el 50% de la muestra. De las 25 respuestas, 19 son nominales (*historia, Hispania, literatura, poema, biología, personaje, arquetipo, indígena, geografía, revolución, reforma, migración, antropología, región, meteorología y población*) y 5 adjetivales (*crónico, lingüístico, bélico, eclesial e hídrico*). Las respuestas de la sección de la conclusión ascendieron a 38 respuestas de 50 (diez por cada licenciatura). De estas, hay 27 nominales (*migración, etnia, antropología, ecología, historia, indígena, población, territorio, geografía, arqueología, biología, Caxcan, edafología, erosión, publicación, agrícola, región, ejido, Hispania, revolución, imperio, reforma, borbón, mitología, literatura y poema*), 10 adjetivos (*nacional, corresponsal, prehispánico, dendrítico, hídrico, forestal, demográfico, eclesial, hegemónico, crónico*) y un verbo (*regir*).

4. DISCUSIÓN Y CONCLUSIÓN

El objetivo de este trabajo fue identificar las diez palabras más frecuentes del léxico académico y disciplinar presentes en las secciones de introducción y conclusión de 135 tesis de grado, con el propósito de evidenciar el grado de dominio y la variación léxica que manifiestan los estudiantes universitarios de Ciencias Sociales y Humanidades.

Para ello, se conformó un corpus compuesto por las secciones mencionadas de tesis pertenecientes a cinco licenciaturas dentro de esta área, partiendo del supuesto de que el léxico académico y el léxico disciplinar se adquieren en etapas diferentes del proceso formativo.

Los resultados indican que el léxico académico comienza a incorporarse desde los primeros años de escolarización, mientras que el léxico disciplinar se consolida de forma progresiva a lo largo de la trayectoria universitaria. La comparación entre ambos tipos de léxico permite arribar a una conclusión preliminar. Sin embargo, para ahondar en el análisis, sería necesario contrastar estos hallazgos con el léxico presente en relatos infantiles, lo que permitiría establecer una línea de desarrollo más amplia en la adquisición del vocabulario académico.

Por lo que respecta al análisis, al seleccionar las 50 palabras más frecuentes y generales del corpus correspondiente a las cinco licenciaturas, y compararlas con dos corpus de referencia (Varela et al., 2013; Coxhead, 2000), los resultados evidencian que los vocablos *analizar*, *cambiar*, *conocer*, *construir*, *exponer*, *método*, *región*, *ser*, *sexo* y *símbolo* están presentes desde los primeros momentos del ámbito académico y se mantienen en el discurso especializado de las tesis de grado. Este hallazgo resulta revelador, ya que la ampliación de los listados de palabras permite identificar los vocablos adquiridos durante los niveles académicos iniciales (primaria, secundaria y bachillerato).

A partir del análisis de los *tokens* en la muestra general, se concluye que el número de participantes –en este caso, tesis de grado– sí incide significativamente en el volumen de respuestas. Los datos evidencian que con 80 tesis de grado (37 de Antropología y 43 de Historia) se ha obtenido más del 70% del total del corpus. En cuanto a los *types*, al compararlos con los *tokens*, se observa una reducción del 30.5% del total. Esta misma tendencia se refleja en los *lemas-type* del corpus, los cuales representan el 25% del total de *tokens*. Este comportamiento coincide con lo reportado en estudios sobre frecuencia léxica, donde se establece que los *lemas-type* tienden a ser menos numerosos que los *types*, y estos, a su vez, son considerablemente inferiores en cantidad respecto a los *tokens* (Cisneros y Olave, 2019).

En cuanto al análisis de los *lemas-type* del corpus, se muestra una clara prevalencia en el uso de los sustantivos (N) sobre los verbos (V) y los adjetivos (Adj). Esto es consistente con estudios previos (Cisneros y Olave, 2019; Parodi, 2005) sobre el lenguaje académico. En general, los sustantivos tienden a ser la categoría con mayor frecuencia de aparición, dado su papel en la construcción de conceptos y nociones clave en textos especializados.

Un aspecto particularmente relevante del análisis comparativo es que las tesis en Antropología e Historia registran un mayor uso de vocablos en la sección de introducción, mientras que Geografía e Historia destacan en las conclusiones. En cuanto a los verbos, Historia es la licenciatura con mayor presencia verbal en la sección de la introducción, mientras que Antropología presentó mayor uso verbal en las conclusiones. Respecto a

los adjetivos, Antropología muestra un uso superior tanto en la introducción como en la conclusión de esta categoría gramatical.

Sobre los datos léxicos analizados del corpus, se ha obtenido un total de 35,677 *tokens*, distribuidos en 270 textos (135 por sección). Al comparar estos resultados con los obtenidos por Cisneros y Olave (2019) se observa una diferencia significativa: este estudio presenta 20,641 *tokens* más, a pesar de haber utilizado un menor número de textos (81 textos en el estudio comparado). En lo que se refiere al comparativo de *lemas-type*, se obtuvieron un total de 4,740, mientras que el estudio de Cisneros y Olave generó 4,005 *lemas-type* a partir de 81 textos. Este dato también es revelador para los estudios de este tipo porque el número de *lemas-type* no depende del volumen del texto analizado, sino que está relacionado con la diversidad y riqueza léxica de los propios textos.

Además, si observamos los resultados por categoría gramatical, el número alto de verbos aparecidos entre las primeras diez respuestas refleja la tendencia a establecer las acciones por encima del nombramiento o la descripción que se da en el adjetivo. Este es otro hallazgo revelador en cuanto a las categorías gramaticales que aparecen en las tesis de grado. Inicialmente, se podría decir que primero aparecerán los procesos-acciones y después otras categorías gramaticales. Estos resultados no solo evidencian las tendencias en el uso y selección de las categorías gramaticales dentro cada una de las secciones de la tesis, sino que también sugieren posibles implicaciones discursivas en la manera en que se construyen y comunican los contenidos académicos a través de un determinado tipo de léxico.

En particular, el uso destacado de adjetivos en Antropología podría interpretarse como una estrategia para enriquecer la caracterización de los fenómenos estudiados, diferenciándose así de otras áreas del conocimiento. Así pues, este hallazgo subraya la relevancia en cuanto a su aprendizaje entre el léxico académico general, que se adquiere a lo largo de la formación educativa, y el léxico disciplinar, que se especializa en el contexto universitario y se asocia con el conocimiento específico de cada disciplina.

En cuanto a la pregunta que se propuso responder, se afirma que sí existe tanto un vocabulario académico como disciplinar en ambas secciones de la tesis que, de cierta manera, nutre y fortalece el aparato crítico y metodológico del área, en este caso, de las Ciencias Sociales y Humanidades y, en cuanto a la diferencia por conteo de palabras, no es significativa por lo menos a las diez palabras frecuentes. En otro momento se tendría que ampliar la muestra a 50 palabras frecuentes por licenciatura para justificar mejor esto.

Sobre los resultados, las tesis de Arqueología, Geografía y Lengua muestran un número de términos académicos y especializados inferior a los 100 *tokens*. En contraste, Antropología e Historia superan ese umbral, lo que indica un mayor uso de vocabulario en estas licenciaturas. Al analizar específicamente las secciones de introducción y conclusión, Arqueología es la única licenciatura que presenta un número total de términos académicos y disciplinares por debajo de las 100 respuestas, lo que tal vez se deba al poco número de tesis encontradas en el periodo de muestra (2007-2014). Por otro lado, aunque Antropología es la licenciatura que con mayor frecuencia aparece en los primeros lugares en cuanto al uso de vocablos generales y por categoría gramatical, muestra un menor número de términos especializados en la sección de conclusión.

REFERENCIAS BIBLIOGRÁFICAS

- Bazerman, Charles. (2012). *Géneros textuales, tipificación y actividad*. Benemérita Universidad Autónoma de Puebla.
- Castello, Montserrat, Mariona Corcelles, Anna Iñesta, Gerardo Bañales y Norma Vega. (2011). “La voz del autor en la escritura académica: una propuesta para su análisis”. *Revista Signos*, vol. 44, núm. 76, pp. 105-117.
- Cisneros, Mireya y Giohanny Olave. (2019). “El léxico disciplinar en la educación universitaria”. *Panorama Económico*, vol. 27, núm. 1, pp. 249-266.
- Coxhead, Averil. (2016). “Acquiring academic and disciplinary vocabulary”. *The Routledge Handbook of English for Academic Purposes*, editado por Ken Hyland y Philip Shaw. Routledge, pp. 177-190.

- Coxhead, Averil. (2000). "A New Academic Word List". *TESOL Quartely*, vol. 34, núm. 2, pp. 213-238.
- Da Cunha, Iria y M. Amor Montané. (2019). "Textual genres and writing difficulties in specialized domains". *Revista Signos*, vol. 52, núm. 99, pp. 4-30.
- Fernández-Silva, Sabela y Nelson Rojas. (2015). "La variación terminológica en la comprensión y producción de textos académicos". *Ibérica*, núm. 30, pp. 183-208.
- Fregoso Peralta, Gilberto y Luz Eugenia Aguilar González. (2019). "Disponibilidad, diversidad y complejidad de vocabulario técnico en estudiantes de ingeniería agroindustrial". *Revista de Educación y Desarrollo*, núm. 50, pp. 27-36.
- Gómez Molina, José Ramón. (2004). "Las unidades léxicas en español". *Carabela*, núm. 56, pp. 27-50.
- Lara, Luis Fernando. (2006). *Curso de lexicología*. El Colegio de México.
- López Ferrero, Carmen. (2002). "Aproximación al análisis de los discursos profesionales". *Revista Signos*, vol. 35, núm. 51-52, pp. 195-215.
- Meza, Paulina, Fernando Lillo-Fuentes y Felipe González-Catalán. (2022). "Descripción de rasgos léxicos de textos profesionales producidos por estudiantes de derecho y medicina". *Literatura y Lingüística*, núm 45, pp. 333-366.
- Nazar, Rogelio. (2011). "Estudio diacrónico de la terminología especializada utilizando métodos cuantitativos: ejemplos de aplicación a un corpus de artículos de lingüística aplicada". *Revista Signos*, vol. 44, núm. 75, pp. 8-67.
- Parodi, Giovanni. (2005). "Discurso especializado y lingüística de corpus: hacia el desarrollo de una competencia psicolingüística". *Boletín de Lingüística*, vol. 17, núm. 23, pp. 61-88.
- Pérez, Marco. (2025). "Análisis de la riqueza léxica en tesis de grado de Ciencias Sociales y Humanidades (2007-2014)". *Lenguaje*, vol. 53, núm. 1, pp. 1-25.
- Pérez, Marco y Guadalupe Rojas. (2023). "Análisis de las frecuencias léxicas verbales de redacciones universitarias de estudiantes asesorados de la Clínica de Escritura de la Facultad de Psicología de la UASLP, México". *Lengua y Sociedad*, vol. 22, núm. 2, pp. 357-376.
- Porrás-Garzón, Jorge y Laia Vidal-Sabanés. (2019). "Les abreviacions en l'informe mèdic: abreviatures, sigles i símbols". *Comunicació, llenguatge y salut*, pp. 59-65.

- Rojas Corona, Ma. Guadalupe, Gabriela Silva Maceda y Silvia Romero Contreras. (2024). “Relación del lenguaje académico y la comprensión lectora con la escritura académica en educación superior en pandemia”. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 14, núm. 18, <https://doi.org/10.23913/ride.v14i28.1813>
- Torres Perdigón, Andrea. (2018). “¿Escritura disciplinar? Una experiencia a partir de la lectura crítica y la habilidad argumentativa en la universidad”. *Revista Mexicana de Investigación Educativa*, vol. 23, núm. 76, pp. 95-124.
- Tutin, Agnés, Juan Felipe Zuluaga Molina y Alejandro Arroyave Tobón. (2017). “*En este artículo queremos mostrar que...* Léxico verbal y posicionamiento del autor en los artículos en ciencias humanas”. *Lingüística y Literatura*, núm. 72, pp. 247-271.
- Varela Barraza, Julio Agustín, Felipe Cabrera González, Daniel Zarabozo Enríquez de Rivera, Yuridia Larios Villa y Miguel González Ortiz. (2013). “Las 5000 palabras más frecuentes en los libros de texto oficiales de la educación básica en México”. *Revista electrónica de investigación educativa*, vol. 15, núm. 3, pp. 114-120.
- Venegas, René, Paulina Meza Guzmán y Juan Martínez Hincapié. (2013). “Procedimientos discursivos en la atribución del conocimiento en tesis de lingüística y filosofía en dos niveles académicos”. *RLA. Revista de Lingüística Teórica y Aplicada*, vol. 51, núm. 1, pp. 153-179.