

Disponibilidad léxica matemática y estadística: análisis comparativo, perfiles estructurales y predicción de repertorios en la media académica

Lexical availability in mathematics and statistics: Comparative analysis, structural profiles, and repertoire prediction in Secondary Education

José Hernando Ávila-Toscano^{1*}, Orlando Junior Zambrano Yepes¹, María de los Ángeles Páez Pérez², Juan Darío Páez Cáceres¹

¹Universidad del Atlántico, Colombia; ²Universidad de la Costa, Colombia

Recibido:

26 de junio, 2025

Aceptado:

01 de marzo, 2026

Publicado:

31 de marzo, 2026

*Autor de

correspondencia

José Hernando Ávila-Toscano
Universidad del Atlántico

Correo electrónico:

joseavila@uniatlantico.edu.co

Como citar:

Ávila-Toscano, J. H., Zambrano Yepes, O., Páez Pérez, M. de los Ángeles, & Páez Cáceres, J. D. (2026). Disponibilidad léxica matemática y estadística: análisis comparativo, perfiles estructurales y predicción de repertorios en la media académica. *Revista De Estudios Y Experiencias En Educación*, 25. <https://www.rexe.cl/index.php/rexe/article/view/3290>

RESUMEN

La disponibilidad léxica implica la activación de palabras evocadas ante estímulos temáticos. En matemáticas y estadística, disponer de términos especializados contribuye a la comprensión conceptual, especialmente cuando se articula el lenguaje intuitivo con el formal. Este estudio evaluó a 402 estudiantes colombianos de educación media (213 chicas), provenientes de escuelas públicas y privadas, con el objetivo de analizar la disponibilidad léxica y la estructura del repertorio en cinco centros de interés matemático y tres estadísticos. Asimismo, se exploraron perfiles léxicos mediante análisis de clúster y regresión logística. Geometría y Trigonometría fue el centro con mayor riqueza léxica, mientras que en Estadística Inferencial se evidenció un repertorio limitado y con confusión respecto a términos de estadística descriptiva. Se diferencian dos perfiles léxicos diferenciados: uno con alta disponibilidad individual y mayor cohesión en el vocabulario, y otro caracterizado por descentralización léxica y baja cohesión. Se proponen implicaciones pedagógicas para cada perfil, concluyendo que es indispensable enseñar ambas disciplinas con un uso formal del lenguaje técnico y su respectiva contextualización a la realidad del estudiante.

PALABRAS CLAVE

Matemáticas y estadística; vocabulario; lenguaje hablado; conceptos; educación secundaria.

ABSTRACT

Lexical availability refers to the activation of words evoked by thematic stimuli. In mathematics and statistics, having access to specialized terms contributes to conceptual understanding, especially when intuitive language is combined with formal language. This study assessed 402 Colombian high school students (213 girls) from public and private schools, in order to analyze lexical availability and repertoire structure across five mathematical and three statistical domains. Additionally, lexical profiles were explored using cluster analysis and logistic regression. Geometry and Trigonometry was the area with the richest vocabulary, while Inferential Statistics showed limited vocabulary and confusion regarding descriptive statistical terms. Two distinct lexical profiles were identified: one with high individual vocabulary availability and greater cohesion, and another characterized by lexical decentralization and low cohesion. Pedagogical implications are proposed for each profile, concluding that it is essential to teach both disciplines using formal technical language and contextualizing it to the students' reality.

KEYWORDS

Mathematics and statistics; vocabularies; spoken language; concept; secondary education.

1. Introducción

Disciplinas como las matemáticas y la estadística son cognitivamente exigentes para muchos estudiantes, especialmente porque se siguen enseñando con procesos memorísticos y mecánicos (Cerdeña & Jarquín, 2023). Estas disciplinas emplean un lenguaje exacto, basado en símbolos y abstracciones relativamente menos ambiguas que los lenguajes naturales (Mendoza et al., 2024). Los estudiantes resuelven problemas empleando expresiones simbólicas que deben convertir a lenguaje natural y viceversa, lo que implica habilidades lingüísticas asociadas con estos saberes (Martínez 2022).

Todo sistema lingüístico comprende elementos esenciales para construir y comprender significados (Puente et al., 2021). En contenidos matemáticos y estadísticos, esto resulta particularmente relevante, al tratarse de saberes con conceptos y términos complejos, cuyo dominio favorece la comprensión profunda de los significados vinculados a dos campos del conocimiento fundamentales en educación (Zambrano & Agustín, 2023). Esta relación justifica el estudio de la disponibilidad léxica en dichos dominios.

1.1 Disponibilidad léxica: concepto y medida

Aprender implica adquirir términos que representan conceptos, y que ganan complejidad en la medida que avanza la escolarización, exigiendo estrategias adecuadas de los estudiantes para comprender y usar correctamente dicha terminología (Manni, 2022). El léxico adquirido permite que los estudiantes dispongan de mejores bases para cualquier dominio del aprendizaje, dado que es una parte sustancial del conocimiento lingüístico (Crossley, 2017). Allí radica su relevancia en el campo educativo, pues el dominio oportuno de conceptos y términos contribuye con la apropiación del conocimiento disciplinar (Cerdeña et al., 2021).

Los términos aprendidos se incorporan al repertorio lingüístico estudiantil formando parte de sus estructuras cognitivas, para acceder a ellos y emplearlos espontáneamente, lo que se denomina disponibilidad léxica (DL). Permite analizar las palabras evocadas de forma inmediata y natural en torno a estímulos temáticos puntuales (Sandu, 2013), ocupándose de la accesibilidad de las palabras en un contexto específico según las necesidades surgidas en la producción lingüística. Es, por tanto, una manifestación dinámica del léxico mental básico, que incluye el acceso a términos apropiados en una situación comunicativa o área de conocimiento (Ferreira et al., 2014; Ferreira et al., 2019).

Con los estudios sobre DL en educación se busca identificar el horizonte conceptual de los estudiantes en diversos campos del conocimiento, para lo cual se aplican técnicas de análisis que evalúan los términos asociados espontáneamente con tópicos o temas empleados como estímulo para generar la respuesta léxica, llamados centros de interés (Otero et al., 2015; Moreno, 2022). Las medidas (ver Tabla 1) incluyen indicadores lexicométricos enfocados en las propiedades de las palabras, el repertorio individual o la creatividad lingüística, captando características profundas del léxico como su estructura interna, organización y centralidad (Ávila-Muñoz & Sánchez, 2010; Callealta & Gallego, 2016). Otros indicadores tienen fines descriptivos permitiendo retratar el comportamiento léxico del grupo (Rojas et al., 2017).

Tabla 1

Principales indicadores de disponibilidad léxica.

Índices lexicométricos (Callealta & Gallego, 2016)	Descripción
Índice de Disponibilidad Léxica de Palabra (IDLP)	Fuerza o accesibilidad de las palabras por centro de interés. Se pondera por la frecuencia y posición en que se evoca.
Índice de Disponibilidad Léxica Individual (IDLI)	Riqueza léxica individual. Valores altos indican repertorios extensos y frecuentes en posiciones iniciales.
Índice de Descentralización Léxica (IDD)	Creatividad y diversidad léxica. Valores altos indican mayor variedad asociativa y menor previsibilidad.
Indicadores descriptivos (rojas et al., 2017)	Descripción breve
Número de participantes (N)	Total de estudiantes que responde en un campo de interés.
Total palabras evocadas	Total de respuestas dadas por todos los informantes para un centro de interés. Indica volumen léxico general.
Longitud media del repertorio (XR)	Promedio de palabras evocadas por estudiante. Representa la fluidez léxica.
Número de Palabras Diferentes (NPD)	Cantidad de términos únicos registrados por centro de interés. Representa la diversidad léxica bruta.
Índice de Cohesión (IC)	Evalúa la homogeneidad del grupo. Valores altos indican evocación de términos similares; valores bajos, mayor diversidad entre participantes.

Estudiar DL en educación a partir de estos índices, permite conocer las palabras que integran el repertorio mental inmediato del alumnado frente a un área del saber (Rojas et al., 2017), desvelando fortalezas y debilidades en su proceso de aprendizaje.

1.2 Disponibilidad léxica en matemáticas y estadística

El dominio de la terminología matemática es necesario para el entendimiento conceptual (Schuth et al., 2017). Acceder a este vocabulario especializado es desafiante para muchos alumnos al tratarse de estructuras lingüísticas comúnmente ajenas a la vida cotidiana, o con significados técnicos distintos a los usuales (Magallanes & Muñoz, 2022). En la práctica, muchos docentes basan la enseñanza en procesos numéricos, cálculos y métodos que desvinculan el lenguaje matemático del cotidiano, limitando la comprensión profunda de los conceptos y su posibilidad de aplicación contextualizada (Saiz & Torres, 2022).

Los estudiantes deben comprender los enunciados de los problemas y la fundamentación conceptual; por lo tanto, análisis formal y contexto resultan inseparables (Catalán & Montero, 2020). Experimentalmente se ha demostrado, por ejemplo, que la velocidad de los estudiantes para resolver problemas aritméticos depende de la interacción entre el dominio de esta área y elementos lingüísticos conceptualmente relacionados (Daroczy et al., 2020). Por ende, es prioritario que los docentes orienten su enseñanza hacia el significado contextual de los tópicos estudiados, para generar mayor comprensión en el alumnado (Carrillo & Pulido, 2023). Esto implica reconocer y aprovechar el repertorio lingüístico previo de los estudiantes, incluso si es informal, puesto que no necesariamente es un obstáculo sino un punto de partida para la formalización conceptual (Ingram & Andrews, 2018). Al articularlo con los nuevos temas, se fomenta una toma de conciencia que armoniza los conceptos intuitivos con la comprensión formal (Pacheco et al., 2016).

Algunos estudios han explorado el repertorio léxico estudiantil en matemáticas y estadística, aportando evidencia sobre términos disponibles y dominantes de acuerdo con el centro de interés abordado. En estudiantes colombianos se ha reportado que términos como "sumas" y "restas" tienen alta dominancia léxica en aritmética; "ecuaciones" y "letras" en álgebra lineal; "integrales", "derivadas" y "función" en cálculo; y términos sencillos como "números", poseen alta dominancia léxica en varios centros de interés como álgebra, estadística, cálculo y ecuaciones (Magallanes & Muñoz, 2022). Por su parte, en estudiantes chilenos destacan términos como "ecuación"

en el campo del Álgebra, o figuras básicas para Geometría ("triángulo", "cuadrado", "círculo"), junto con términos como "ángulo", "área", "perímetro" y "vértice"; mientras que en el centro de Datos y azar sobresalieron términos propios de la estadística descriptiva ("gráfico", "media aritmética", "moda", "mediana", "porcentaje", "tablas") (Cerdeira et al., 2021).

Por su parte, Puente et al. (2021) evaluaron estudiantes chilenos a partir de los centros de interés Sistema numérico y Valor posicional, encontrando dificultades en el manejo de términos relacionados con más de un dígito. Palabras como "milésima" o "décima" eran mal empleadas, lo que se evidenció cuando debían redactar frases partiendo de la tarea de pie forzado "La mujer maravilla". Los autores identificaron un uso común y erróneo de la palabra "milésima" (ej.: *"La Mujer Maravilla salvó muchos gatitos, creo que era una milésima de gatitos"*), tanto en estudiantes de primaria como de secundaria.

Estos resultados pueden variar según aspectos socioculturales y económicos, dado que el nivel educativo familiar, el uso de textos, e incluso, el entorno escolar y los recursos pedagógicos, pueden incidir positiva o negativamente en la adquisición y uso del vocabulario (Molinares et al., 2014, Trigo et al., 2020). Por ello, algunos estudios incorporan comparaciones entre estudiantes según la escuela de procedencia (privada-pública). Basados en la medición del promedio de respuestas (XR), la cantidad de palabras diferentes (NPD) y el índice de cohesión (IC), Salcedo y del Valle (2013) identificaron que estudiantes chilenos de centros privados disponían de un léxico más amplio y mayor índice de comprensión que sus pares de instituciones públicas. Análogamente, en estudiantes españoles Trigo y Santos-Díaz (2019) identificaron que el promedio de palabras disponibles en escuelas privadas era de 319.50, frente a 308.60 en instituciones públicas.

Aunque estos aportes son valiosos, es común que los análisis se concentren en frecuencias o conteos de palabras, a expensas de la estructura interna del repertorio. Además, aunque se ha planteado que existen desigualdades entre instituciones públicas y privadas en el acceso a vocabulario especializado, estas diferencias suelen basarse en indicadores descriptivos, y poco se ha investigado si dichas brechas se registran al evocar términos relacionados con campos puntuales de la matemática y la estadística. Tampoco es habitual el uso de los indicadores lexicométricos para predecir los desempeños léxicos de los alumnos.

1.3 El presente estudio

Hasta ahora los estudios sobre DL se han centrado principalmente en matemáticas, con menos evidencias en estadística. La mayoría analiza indicadores descriptivos que pierden precisión cuando se comparan grupos, por ejemplo, según escuela pública o privada. Este estudio busca llenar ambos vacíos con una contribución novedosa al combinar análisis lexicométrico descriptivo y contraste de grupos mediante técnicas robustas (test de Yuen), aportando validez a las comparaciones entre grupos. A esto se suma el cálculo de conglomerados (análisis de clúster) para identificar perfiles léxicos y la estimación de modelos predictivos (regresión logística) para determinar la pertenencia a esos perfiles. En este estudio, además de comparar grupos, el objetivo incluye analizar la DL y la estructura del repertorio léxico en contextos matemáticos y estadísticos, explorando perfiles léxicos mediante técnicas de agrupamiento y modelado predictivo.

Para ello, se identifican los términos con mayor índice de disponibilidad léxica (IDL) por centro de interés, se comparan los índices individuales de disponibilidad (IDLI) y descentralización léxica (IDD), junto con otros indicadores léxicos entre estudiantes de colegios públicos y privados, se detectan perfiles léxicos mediante análisis de clústeres, y finalmente se identifican predictores individuales de pertenencia a dichos perfiles usando regresión logística.

2. Método

2.1 Diseño

Se desarrolló un estudio de diseño predictivo transversal (Ato et al., 2013), que permite varios niveles de análisis (descripción, correlación) para seleccionar las mejores variables predictoras para otra variable dependiente. Se emplearon grupos naturales o preexistentes, correspondientes a estudiantes de escuelas públicas y privadas, esto permitió contrastar el comportamiento léxico entre grupos y estimar diferencias.

2.2 Participantes

Mediante muestreo accidental se seleccionaron 402 estudiantes colombianos de ambos géneros, entre 15 y 18 años, de dos escuelas públicas y dos privadas, que cursaban la media académica (últimos dos años de escolarización obligatoria) y contaban con autorización familiar para participar en el estudio. La selección no consideró el historial académico en matemáticas o estadística.

La muestra se dividió en 213 chicas y 189 chicos, cuya edad media fue de 15.6 años ($DE = 0.838$). Para las chicas, el promedio de edad fue de 15.6 años ($DE = 0.827$) y para los varones de 15.7 ($DE = 0.846$). Además, 230 estudiaban en escuelas privadas y 162 en colegios públicos.

2.3 Instrumentos

Siguiendo el protocolo internacional de estudios lexicográficos (Santos-Díaz et al. 2020; Santos-Díaz, 2020), los datos se recogieron empleando una tarea de evocación libre de palabras mediante una lista abierta, que permite identificar el léxico estudiantil en temas puntuales gracias a la evocación natural de palabras para ocho centros de interés: cinco para matemáticas y tres para estadística, seleccionados según lo definido en los Derechos Básicos de Aprendizaje (DBA) en Colombia (Ministerio de Educación Nacional [MEN], 2016).

Cinco jueces expertos en enseñanza de matemáticas y estadística evaluaron el instrumento para obtener evidencia de validez de contenido para los centros de interés, según criterios de pertinencia, claridad y representatividad usando una escala de cinco puntos (1 = *inaceptable*, 5 = *excelente*). Los resultados se analizaron con el coeficiente de validez de contenido (CVC) de Hernández-Nieto (2002) (Tabla 2), que obtuvo un valor global de .949, superior al umbral de .80, indicando adecuada evidencia de validez.

Tabla 2

Evaluación experta de centros de interés según DBA para media académica.

Centro de interés	Justificación DBA	Media juez	CVC
Álgebra	Área transversal vinculada al reconocimiento, manipulación y resolución de ecuaciones, representación de situaciones cotidianas y matemáticas mediante expresiones algebraicas.	15	.999
Geometría y Trigonometría	Necesidad de comprender relaciones geométricas, representar figuras, trabajar con ángulos y aplicar trigonometría en contextos escolares y cotidianos.	14.6	.973
Funciones	Herramienta para interpretar fenómenos, representar relaciones y resolver problemas.	14.2	.946
Cálculo	Interpretación de cambios y tasas de variación que enriquece el repertorio léxico avanzado.	13.4	.893
Números y operaciones	Dominio de operaciones numéricas y sus propiedades, incluido en los DBA como base para resolver problemas de variada complejidad.	15	.999
Probabilidad	Comprensión de probabilidad de eventos simples y compuestos, uso de técnicas de conteo y argumentación de respuestas desde la combinatoria.	14.2	.946
Estadística Descriptiva	Interpretación y uso de datos, medidas de tendencia y dispersión, representación gráfica.	14	.933
Estadística Inferencial	Inferencias sobre poblaciones a partir de muestras, uso de software y análisis crítico de resultados estadísticos.	14	.933

2.4 Procedimiento

Se accedió a la muestra con autorización de las instancias administrativas de cada institución, quienes gestionaron el consentimiento informado de las familias de los estudiantes. A estos se les solicitó presencialmente su asentimiento informado cumpliendo con principios éticos internacionales (Declaración de Helsinki) y nacionales (Ley 1090 de 2010, Resolución ministerial 8430 de 1993), que incluyeron: participación libre, anonimato, confidencialidad, protección de datos personales, ausencia de riesgo.

Los datos se recogieron de forma física y colectiva en los salones de clase. A cada estudiante se le proporcionó una lista de registro de palabras por centro de interés, asignando dos minutos para escribir las palabras evocadas por cada centro. El tiempo fue cronometrado y controlado por un investigador. El proceso se cumplió en un tiempo estándar de 16 minutos por grupo.

2.5 Análisis de datos

Las palabras evocadas se cargaron en una base de datos en Excel® que fue sometida a limpieza y depuración. Se normalizaron en minúscula unificando términos similares (ej.: "gráfica", "gráfico", "gráficos" → *gráfica*) para evitar redundancia que pudiese sesgar los indicadores estimados. Inicialmente se realizó análisis descriptivo de los indicadores de DL (IDLP, IDLI, IDD), para caracterizar los repertorios léxicos de los estudiantes, y se incluyeron otros indicadores descriptivos como *N* (número de estudiantes por grupo), *Total palabras* (conteo de palabras evocadas), *XR* (longitud media del repertorio, para medir la fluidez léxica), *NPD* (diversidad léxica bruta), e *IC* (índice de cohesión léxica).

Estos indicadores se compararon de acuerdo con el género de los estudiantes y el tipo de colegio, mientras que los valores de IDD e IDLI se compararon entre escuelas públicas y privadas empleando el test de Yuen, que ofrece una estimación robusta de diferencias entre grupos, adecuada para muestras grandes en condiciones de alta variabilidad y ausencia de normalidad. Para las diferencias identificadas se estimó el tamaño del efecto con el estadístico robusto *g* de Hedges asumiendo varianzas desiguales.

Se complementó el análisis con un procedimiento exploratorio para detectar patrones globales de desempeño léxico mediante análisis de conglomerados (*k-means*) empleando las puntuaciones de los índices IDLI e IDD, y finalmente, se pronosticó la pertenencia a los clústeres mediante un modelo de regresión logística binaria empleando los indicadores léxicos descriptivos (*XR*, *NPD*, *IC*), el tipo de colegio y género como predictores. Los análisis se realizaron con RStudio, empleando librerías especializadas: *dplyr*, *ggplot2*, *stringr*, *nortest*, *car*, *WRS2*.

3. Resultados

3.1 Comparación de DL y estructura del repertorio léxico por tipo de escuela

Inicialmente se evaluó el IDLP para identificar los términos más relevantes por cada centro de interés. Sobresalieron términos sencillos como "números" en los centros de interés Cálculo, Funciones, Álgebra, Estadística inferencial y Probabilidad. Resulta llamativo que cuatro de las cinco palabras más usuales en Estadística descriptiva son idénticas a las registradas en Estadística Inferencial, y en realidad, corresponden a términos propios del razonamiento descriptivo (Tabla 3).

La columna *n_pos* (Tabla 3) reporta el número de posiciones en las que se registran las palabras evocadas, siendo "Cuadrática" una de las comunes al registrarse en 10 posiciones en el centro de interés Funciones. En Álgebra, las palabras con mayor IDLP aparecen consistentemente entre 8 y 9 veces en distintas posiciones, mientras que Geometría y trigonometría, y Números y operaciones, destacan como aquellos en los que se dan las mayores repeticiones de términos, sugiriendo que se trata de palabras comúnmente empleadas por los estudiantes.

Tabla 3

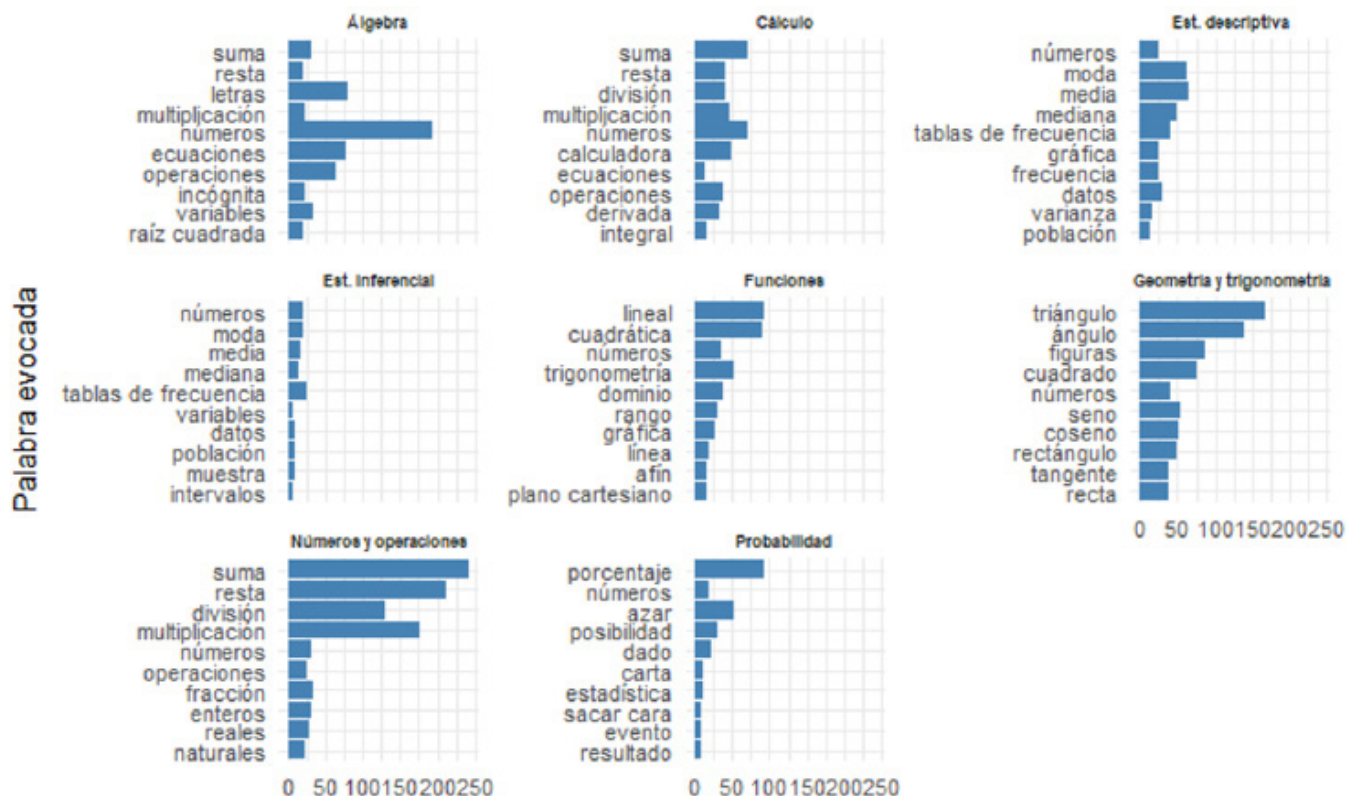
Índice de Disponibilidad Léxica de una Palabra (IDL P) para los cinco términos más usuales en cada centro de interés.

Cálculo	n_pos	IDL P	Funciones	n_pos	IDL P
Números	5	72.5	Lineal	7	94
Suma	5	71.1	Cuadrática	10	90.2
Calculadora	4	50.2	Trigonometría	6	50.9
Multiplicación	7	47.2	Dominio	7	37.2
Resta	6	41.6	Números	5	34.8
Geometría y trigonometría	n_pos	IDL P	Números y operaciones	n_pos	IDL P
Triángulo	12	166	Suma	11	243
Ángulo	13	139	Resta	10	212
Figuras	11	87.4	Multiplicación	12	175
Cuadrado	11	76.5	División	10	131
Seno	11	53.3	Fracción	10	34.1
Álgebra	n_pos	IDL P	Estadística descriptiva	n_pos	IDL P
Números	8	193	Media	7	64
Letras	7	81.5	Moda	7	62.2
Ecuaciones	8	77.4	Mediana	7	48.8
Operaciones	8	63.6	Tablas de frecuencia	6	39.5
Variables	9	34	Datos	6	30.3
Estadística inferencial	n_pos	IDL P	Probabilidad	n_pos	IDL P
Tablas de frecuencia	6	24	Porcentaje	6	93.5
Números	3	19	Azar	6	52.4
Moda	5	18.6	Posibilidad	5	29.4
Media	5	15.7	Dado	5	20.7
Mediana	5	15.4	Números	5	17.8

Una ampliación del análisis de la variabilidad lexicográfica en cada centro de interés se registra en la Figura 1, en donde se agrupan gráficas descriptivas de las 10 palabras más usuales en cada centro. Este análisis visual muestra con mayor precisión el bajo nivel de DL general en Estadística inferencial, que registra las más bajas frecuencias de palabras en toda la muestra. Este resultado es similar al observado en el centro de Probabilidad. Opuestamente, otros campos como Geometría y trigonometría agrupan las palabras con frecuencias más elevadas reportadas.

Figura 1

Las 10 palabras con mayor IDLP por cada centro de interés.



Se realizó análisis comparativo entre los estudiantes de escuelas públicas y privadas a partir del desempeño del IDLI y el IDD. La prueba de Anderson-Darling descartó distribución normal de los datos ($p < .05$) y tampoco hubo homogeneidad de varianzas según los resultados del test de Levene ($p < .05$), optándose por realizar comparaciones robustas con el test de Yuen. El análisis descriptivo por tipo de institución (Tabla 4) demuestra que ambos índices presentaron medias similares para los dos tipos de escuela, lo cual se corroboró con la comparación robusta de Yuen, cuyos resultados no respaldan la existencia de diferencias relevantes ($p > .05$).

Tabla 4

Análisis descriptivo de índices individuales de disponibilidad léxica y prueba de diferencias robustas por tipo de escuela.

Estadísticos	IDLI		IDD*	
	Privado	Público	Privado	Público
Media	0.558	0.553	-61.1	-62.4
IC95% Media	[0.550, 0.566]	[0.544, 0.562]	[-60.5, -57.7]	[-65.6, -59.2]
DT	0.060	0.0677	25.8	24.2
Coeficiente de Variación	0.10	0.12	0.42	0.38
Diferencias robustas (Test de Yuen)				
Índice	Media recortada (tr = 0.2)	IC95% diferencia recortada	Yuen (gl)	p-valor
IDLI	2.676	[-1.9449, 7.2975]	1.1402 (268.77)	0.25 > 0.05
IDD	0.008	[-0.0029, 0.0207]	1.4881 (244.19)	0.138 > 0.05

*Para su interpretabilidad, en este índice los valores menos negativos indican mejor desempeño.

Para complementar el análisis de DL fueron evaluados otros indicadores léxicos considerando el efecto del tipo de colegio en cada centro de interés. Para ello se empleó: número de estudiantes por grupo, total de palabras evocadas por centro de interés, longitud media del repertorio (*XR*), diversidad léxica bruta (*NPD*) e índice de cohesión léxica (*IC*).

Los datos demuestran mayor cantidad de estudiantes de instituciones privadas en cinco de los ocho centros de interés. Únicamente hubo un número similar de estudiantes en los centros de Geometría y trigonometría, Números y operaciones, y Probabilidad. Sin embargo, la diferencia en el número de estudiantes no parece afectar el desempeño general de los indicadores léxicos. En relación con la longitud media del repertorio (*XR*), se identificaron valores ligeramente más altos en cinco centros de interés para colegios públicos, a pesar de agrupar menor cantidad de estudiantes, lo cual indica una leve ventaja en la evocación de palabras en promedio por tema.

También hubo mayor número de palabras diferentes (*NDP*) entre los alumnos de colegios públicos, y valores del índice de cohesión (*IC*) ligeramente más altos en cinco centros de interés. Esta información sugiere que podría haber un patrón de mayor diversidad y convergencia léxica. Sin embargo, al aplicar comparaciones robustas con el test de Yuen para cada uno de estos indicadores, las diferencias observadas no alcanzaron una magnitud suficiente para ser interpretadas como relevantes (ver Tabla 5). Aunque se registran contrastes descriptivos en ambas instituciones, estos no constituyen diferencias robustas desde un enfoque inferencial.

Tabla 5

Indicadores de disponibilidad léxica por centro de interés y comparaciones robustas por tipo de institución.

Centro de interés	Colegio	N	Total palabras	NPD	XR	IC
Álgebra	Privado	222	950	126	4.279.279	.1326316
	Público	175	687	118	3.925.714	.1717613
Cálculo	Privado	159	533	111	3.352.201	.2082552
	Público	150	559	92	3.726.667	.1645796
Estadística descriptiva	Privado	143	495	89	3.461.538	.1797980
	Público	124	405	74	3.266.129	.1827160
Estadística inferencial	Privado	80	235	55	2.937.500	.2340426
	Público	56	142	55	2.535.714	.3873239
Funciones	Privado	190	658	111	3.463.158	.1686930
	Público	146	528	111	3.616.438	.2102273
Geometría y trigonometría	Privado	208	1042	127	5.009.615	.1218810
	Público	209	1227	147	5.870.813	.1198044
Números y operaciones	Privado	199	974	111	4.894.472	.1139630
	Público	199	1123	131	5.643.216	.1166518
Probabilidad	Privado	139	361	67	2.597.122	.1855956
	Público	130	349	74	2.684.615	.2120344
Comparaciones robustas test de Yuen			0.115	0.148	0.421	0.264
p-valor test de Yuen			.91 > .05	.84 > .05	.68 > .05	.79 > .05

Nota. *NPD* = diversidad léxica bruta, *XR* = longitud media del repertorio, *IC* = índice de cohesión.

3.2 Perfiles léxicos mediante análisis de conglomerados y su predicción con regresión logística

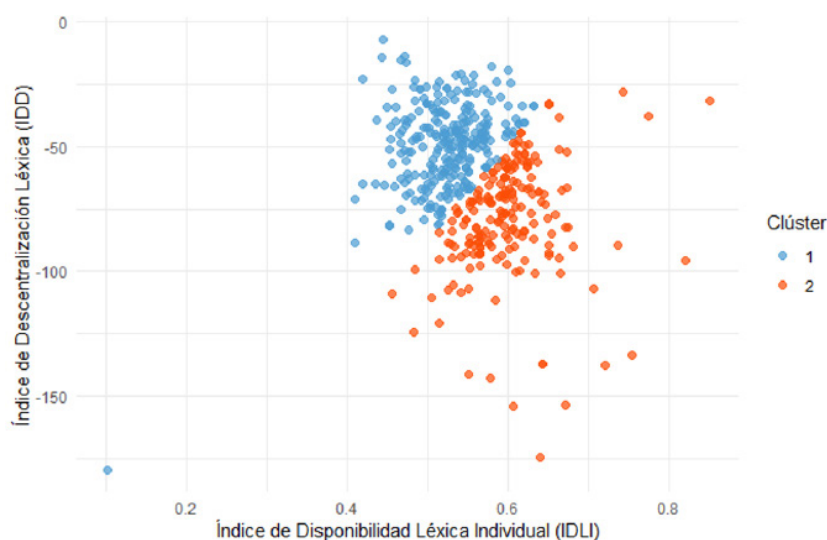
Ante la ausencia de diferencias robustas en la evaluación comparativa de disponibilidad y repertorio léxico, se optó por complementar el análisis mediante un enfoque exploratorio basado en técnicas de agrupamiento, procurando identificar patrones globales en el desempeño léxico de los estudiantes. Este procedimiento se cumplió por medio del análisis de clustering léxico mediante conglomerados *k-means*, lo que permitió identificar perfiles latentes de estudiantes según las puntuaciones en los índices IDLI e IDD. Esta estrategia facilitó la agrupación de los participantes de acuerdo con la similitud estructural de sus repertorios léxicos.

En la Figura 2 se visualizan dos clústeres claramente diferenciados. El Clúster 1 agrupó 266 participantes divididos en 140 de colegios privados y 126 de colegios públicos. Entre los primeros la media IDLI fue de 0.530 y para IDD fue de -48.2, mientras que en los colegios públicos las medias respectivas fueron 0.514 y -50.1. Por su parte, el Clúster 2 fue más pequeño, al agrupar a 193 estudiantes, de los cuales 95 eran de colegios privados (IDLI = 0.596, IDD = -79.3) y 98 de colegios públicos (IDLI = 0.604, IDD = -78.6).

La comparación robusta de los índices IDLI e IDD por clúster, identificó que el Clúster 2 presenta mayor IDLI ($t[242.76] = 17.32$, $p < .001$), con tamaño de efecto grande ($g = -1.50$, IC95% [-1.71, -1.29]). En cambio, el Clúster 1 muestra valores más elevados de IDD ($t[224.45] = 15.29$, $p < .001$), también con efecto grande ($g = 1.49$, IC95% [1.28, 1.70]), lo cual indica menor cohesión en el repertorio léxico.

Figura 2

Clústeres léxicos sobre valores estandarizados de IDLI e IDD.



El clustering léxico sugiere que los estudiantes del Clúster 2 presentan mayor disponibilidad léxica individual, mientras que en el Clúster 1 se agruparon aquellos con un repertorio más descentralizado, es decir, con menor cohesión léxica. El Clúster 2 puede interpretarse como el subgrupo con mejor rendimiento léxico, en la medida en que evocan más términos con mayor rapidez (fluidez) y comparten un léxico más homogéneo a nivel grupal.

Tras definir los clústeres se ajustó un modelo de regresión logística binaria con el fin de predecir la pertenencia a los clústeres en función de variables sociodemográficas (género, tipo de colegio) y características del repertorio léxico (fluidez, diversidad, cohesión). En la Tabla 6 se resumen los resultados del modelo, el cual mostró métricas de ajuste aceptables, explicando entre el 22.6% y el 35.6% de la varianza (véase los pseudo R^2), y, además, alcanza una exactitud global del 73.5% en la clasificación de los casos.

El modelo indicó que la pertenencia al Clúster 2 (mejor desempeño léxico) se asocia con una mayor longitud media del repertorio ($OR_{XR} = 1.37$) y una menor diversidad léxica bruta ($OR_{NPD} = 0.61$). Entre tanto, no se encontraron efectos estadísticos relevantes del género, el tipo de colegio y el índice de cohesión (IC).

Tabla 6

Regresión logística binaria para predecir pertenencia a clústeres léxicos.

Variable	OR	Intervalo de confianza 95%		p
		Inferior	Superior	
(Intercepto)	3.50	0.025	530.00	.620
Colegio (Público)	1.13	0.734	1.74	.580
Género (chico)	0.87	0.563	1.34	.521
XR	1.37	1.12	1.69	.003 < .01
NPD	0.61	0.471	0.773	< .001
IC	3.20	0.012	880.00	.683

Resumen del ajuste del modelo McFadden = .226; Cox & Snell (r2ML) = .264; Nagelkerke (r2CU) = .356

Nota. OR = Odds Ratio; XR = longitud media del repertorio; NPD = número de palabras diferentes; IC = índice de cohesión léxica.

4. Discusión

Este estudio aporta evidencia novedosa sobre la DL en torno a ocho centros de interés en matemática y estadística de estudiantes de escuelas privadas y públicas, además, a partir de la estructura del repertorio explora perfiles léxicos y emplea modelos de relaciones predictiva para comprender dichos perfiles. Los resultados muestran un léxico disponible ampliamente mayor en los centros de interés matemáticos en comparación con los estadísticos. Esto respalda los hallazgos de estudios previos donde una alta disponibilidad en centros como Geometría, contrasta con registros inferiores en otros como Datos y Azar (Salcedo & Del Valle, 2013; Cerda et al., 2021). En efecto, Geometría y trigonometría destaca como el centro de interés de mayor DL registrado en la muestra, seguido de Números y operaciones y de Funciones, sin embargo, es importante reconocer que los términos disponibles para los participantes se caracterizan por su simpleza y transferibilidad léxica.

La palabra “suma” sobresale en los centros de interés Álgebra, Cálculo y Números y Operaciones, mientras que “números” aparece en siete de los centros estudiados. Si bien esto sugiere que existe transferencia léxica y conceptual, también denota que el lenguaje general de los estudiantes no dispone de términos más especializados (Urzúa et al., 2006) que informe de un dominio más profundo de los tópicos estudiados.

Esta evidencia indicaría un bajo nivel de dominio conceptual en los diferentes centros de interés que son cruciales en el currículo matemático colombiano. La insuficiencia lexical disminuye la efectividad en la generación de redes semánticas que optimicen el aprendizaje de la matemática (Ferreira et al., 2014) y advierte la presencia de fórmulas de enseñanza prototípicas donde el exceso de formalismo y el enfoque en los contenidos sobrepasa la aplicación contextualizada de los conceptos en la vida real (Cerda et al., 2021), reduciendo la posibilidad de que los alumnos otorguen significado.

Con la intención de ampliar los hallazgos relacionados con la estadística, en este estudio se emplearon tres centros de interés para captar mejor la variabilidad de las palabras frente a tópicos más precisos, sin embargo, se ha registrado una alta coincidencia en los términos usados para los centros Estadística descriptiva y Estadística inferencial. En ambos se evocan indistintamente palabras relativas a medidas de tendencia central y tablas, indicando que los estudiantes no diferencian los términos especializados de cada rama, lo cual sugiere un conocimiento superficial que conduce a confundirlas. A esto se suma que los indicadores lexicométricos para Estadística inferencial y Probabilidad son los más bajos, de hecho, los tres centros estadísticos registran la menor cantidad de participantes (*N*) y el índice más bajo de palabras diferentes (*NPD*), con solo 80 en escuelas privadas y 56 en públicas para el campo inferencial, y 67 y 74 respectivamente para tópicos de probabilidad.

Esta evidencia demuestra que el bajo conocimiento estadístico es una debilidad en la educación media. Desde hace tiempo la literatura viene denunciando debilidades sistemáticas en la enseñanza de la estadística que involucran desde cuestiones curriculares, hasta limitaciones en el saber disciplinar del propio profesorado (Avilez et al., 2017). En muchas escuelas la estadística se da con una baja intensidad horaria, se omite por completo de la matemática

escolar (Álvarez et al., 2020), o se imparte empleando recursos —como libros— no especializados, que desatienden contenidos relevantes como problemas de probabilidad (Vásquez et al., 2020) o carecen de tareas contextualizadas (Vargas-Delgado et al., 2021). Esto demuestra la necesidad de atender la enseñanza de la estadística como un campo fundamental de formación escolar, aplicando mejoras curriculares sustanciales y estrategias innovadoras que faciliten articular el saber formal con problemas contextualizados, desde lo cual se favorezca una comprensión más profunda de la disciplina (Carrillo & Pulido, 2023).

Por otro lado, los índices de DL permitieron identificar dos perfiles léxicos diferenciados, confirmando que el repertorio lingüístico de los estudiantes evaluados no es homogéneo, sino que se da en función de las diversas trayectorias formativas y las experiencias construidas dentro de sus contextos de desarrollo escolar y sociofamiliar (Molinares et al., 2014; Trigo et al., 2020). El Clúster 2 revela un perfil de mayor rendimiento léxico, que implica más disponibilidad individual y mayor cohesión en el vocabulario de los estudiantes, combinación que sugiere una mejor fluidez asociativa y el uso de términos disciplinares compartidos con el grupo. El análisis de regresión logística refuerza este hallazgo, al mostrar la fluidez léxica (XR) como el predictor principal de pertenencia al clúster, confirmando que los alumnos que lo integran evocan más palabras de forma fluida y eficiente.

En términos pedagógicos, los estudiantes de este perfil se beneficiarían de estrategias didácticas como el uso sistemático de glosarios técnicos contruidos progresivamente, es decir, tareas que partan de la definición del concepto para luego establecer ejemplos y contraejemplos, y finalmente describir diversos usos típicos del término en contextos específicos. Asimismo, otras estrategias útiles son las tareas de “traducción” de un texto matemático o estadístico a lenguaje natural, garantizando que no se pierda la precisión. Por ejemplo, que las tareas mantengan condiciones, cuantificadores, unidades y relaciones entre variables; con ello se respeta el rigor conceptual. Otra estrategia útil consiste en realizar reportes de informes de proyectos con lenguaje formal para referirse a conceptos disciplinares (ej. “dispersión”, “varianza”), y otras que estimulen la profundización léxica y el uso de términos matemáticos y estadísticos en tareas argumentativas y de resolución de problemas. A partir de estas acciones se puede facilitar que el estudiante reconozca el término, lo defina con sus propias palabras, lo aplique correctamente en diversas situaciones contextualizadas, y lo conecte con conceptos cercanos para evitar un uso “memorístico” o aislado del vocabulario. Este tipo de tareas permite armonizar el uso de nuevos términos con los ya adquiridos por los estudiantes (Pacheco et al., 2016).

Por su parte, el Clúster 1 se asocia con mayor descentralización léxica, caracterizado por una cohesión baja de los términos disponibles. Comparativamente, la descentralización no sugiere baja DL sino aumento de la dispersión semántica, lo que puede implicar que el léxico escolar no es suficientemente sistematizado por estos estudiantes. El modelo de regresión también respalda esta interpretación, dado que la diversidad léxica bruta (NPD) se asoció con mayor probabilidad de pertenecer al Clúster 1, indicando que los estudiantes con repertorios dispersos y variados evocan términos diferentes, pero menos consolidados en el vocabulario disciplinar.

El desajuste entre el vocabulario de uso cotidiano del alumnado y el léxico curricular esperado representa dificultades en la comprensión de las matemáticas (Strasser et al., 2024), por lo cual es necesario enseñar vocabulario matemático y estadístico de forma explícita. En tal sentido, producto de la baja consistencia en el uso de términos, los estudiantes con este perfil léxico podrían beneficiarse de tareas que promuevan la fijación de términos disciplinares esenciales, como el uso de palabras clave o mapas de palabras inherentes a los temas a tratar, o ejercicios comparativos para identificar polisemia mostrando cómo palabras cotidianas adquieren significado especializado en matemática y estadística (ej. “media”).

Finalmente, los resultados no mostraron efecto de variables como género y tipo de escuela en la DL o la pertenencia a los perfiles identificados. Aunque investigaciones previas han reportado diferencias a favor de estudiantes de instituciones privadas (Salcedo & del Valle, 2013; Trigo & Santos-Díaz, 2019), los datos de este estudio no respaldan ese resultado. Además, se utilizaron métodos estadísticos robustos que refuerzan la solidez de esta conclusión, al ofrecer estimaciones más estables del contraste entre grupos. En este sentido, las diferencias observadas en DL y perfiles obedecen principalmente a características del repertorio léxico, más que a condiciones sociodemográficas. Esto tiene implicaciones relevantes, al sugerir que la dificultad para acceder a un vocabulario técnico especializado no se restringe a estudiantes de un solo contexto socioeconómico, sino que podría ser un fenómeno de diversos escenarios escolares.

Trabajos futuros pueden ampliar el análisis del tipo de escuela en escenarios diversos, e incluso ampliar el alcance a la educación rural. También se requiere subsanar algunas limitaciones propias de los estudios transversales, proponiendo seguimientos que permitan identificar si el léxico de los estudiantes varía en riqueza y fluidez en función de la edad o el progreso académico. A esto se suma que es necesario distinguir entre usar palabras para comunicar una idea o emplearla solo como una denominación, es decir, el uso técnico del lenguaje requiere analizar también la intención comunicativa (Ingram & Andrews, 2018), procurando con ello un análisis más preciso del nivel de comprensión real de los estudiantes frente a conceptos matemáticos y estadísticos, cuya profundidad puede no ser capturada en estudios enfocados en la evocación libre de términos sin aplicación en tareas explícitas.

5. Conclusión

El léxico disponible es fundamental para lograr una comprensión eficaz de contenidos matemáticos, pues el aprendizaje de esta área exige el uso de un lenguaje conceptualmente preciso que articule coherentemente términos y símbolos. Muchos estudiantes presentan dificultades para apropiarse del vocabulario formal, sugiriendo una brecha entre su comprensión intuitiva y el uso académico de términos matemáticos y estadísticos. Esta situación pone en evidencia la necesidad de fortalecer la DL en diversos campos de la matemática, y especialmente en estadística, donde los estudiantes de la media académica muestran debilidades significativas. Ampliar el repertorio lingüístico formal puede favorecer mejores trayectorias de aprendizaje, y puede mejorar la comprensión de conceptos y su aplicación contextual.

Referencias

- Álvarez, I., Guerrero, Y., & López, Y. (2020). Taxonomía de errores y dificultades en la construcción e interpretación de tablas de frecuencia. *Zetetike*, 28, 1-22. <https://www.doi.org/10.20396/zet.v28i0.8656553>.
- Ato, M, López, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059. <https://dx.doi.org/10.6018/analesps.29.3.178511>.
- Ávila-Muñoz, A., Sánchez, J. (2010). La disponibilidad léxica. Antecedentes y fundamentos. En Ávila-Muñoz & Villena Ponsoda (Eds.). *Variación social del léxico disponible en la ciudad de Málaga: diccionarios y análisis* (pp. 37-81). Editorial Sarriá.
- Avilez, A., Ordaz, M., & Reyna, L. (2017). Conocimiento y actitudes acerca de la Estadística, de los profesores de secundaria del estado de Yucatán. *Unión: revista iberoamericana de educación matemática*, 52, 46-72. <https://dialnet.unirioja.es/servlet/articulo?codigo=6392993>.
- Callealta, F., & Gallego, D. (2016). Medidas de disponibilidad léxica: comparabilidad y normalización. *Boletín de Filología*, 51(1), 39-92. <http://dx.doi.org/10.4067/S0718-93032016000100002>.
- Carrillo A., & Pulido, R. (2023). Literatura y educación estadística crítica para el abordaje de problemas sociales en el aula: una propuesta innovadora en estudiantes de educación media. *Discimus. Revista Digital de Educación*, 2(2), 141-157. <https://doi.org/10.61447/20231211/art8>.
- Catalán, R. & Montero-Saizaja, A. (2020). Conceptualization of the 'school' in the english available lexicon of spanish adolescents. *Miscelánea: A Journal of English and American Studies*, 61, 13-32. <https://dialnet.unirioja.es/descarga/articulo/7663175.pdf>.
- Cerda, J., & Jarquín, R. (2023). Importancia de la comunicación para la educación en el aprendizaje de la Matemática. *Revista Torreón Universitario*, 12(34), 17-22. <https://doi.org/10.5377/rtu.v12i34.16337>.
- Cerda, G., Pérez, W., & Chandía E. (2021). Disponibilidad léxica en centros de interés asociados a ejes curriculares de matemáticas en estudiantes de contextos de alta vulnerabilidad social. *Revista de Estudios y Experiencias en Educación*, 20(42), 33-51. <https://doi.org/10.21703/rexe.20212042cerda2>.
- Crossley, S., Liu, R., & McNamara, D. (2017). Predicting math performance using natural language processing tools. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 339-343). Association for Computing Machinery. <https://doi.org/10.1145/3027385.3027399>.
- Daroczy, G., Meurers, D., Heller, J., Wolska, M., & Nürk, H-C. (2020). The interaction of linguistic and arithmetic factors affects adult performance on arithmetic word problems. *Cognitive Process*, 21, 105-125. <https://doi.org/10.1007/s10339-019-00948-5>.
- Ferreira, R. A., Garrido, J. I., & Guerra, A. (2019). Predictors of lexical availability in English as a second language. *Onomázein*, 46, 18-34. <https://doi.org/10.7764/onomazein.46.03>.
- Ferreira C., Salcedo L, P., & del Valle L. (2014). Estudio de disponibilidad léxica en el ámbito de las matemáticas. *Estudios Filológicos*, 54, 69-84. <https://doi.org/10.4067/s0071-17132014000200004>.
- Hernández-Nieto, R. (2002). *Contributions to Statistical Analysis: The coefficients of pro-portional variance, content validity and Kappa*. Universidad de Los Andes.
- Ingram, J., & Andrews, N. (2018). Use and meaning: What students are doing with specialised vocabulary. En E. Norén, H. Palmér & A. Cooke (Eds.), *Proceedings of the IV ERME Topic Conference: Classroom-based research on mathematics and language* (pp. 81-88). Dresden, Germany. <https://hal.science/hal-01856481>.

- Magallanes, E., & Muñoz, M. (2022). Disponibilidad de terminología matemática en alumnos de la UPZ. Una explicación de los resultados académicos. *Revista Educación Las Américas*, 11(2). <https://doi.org/10.35811/rea.v11i2.174>.
- Manni, J. (2022). La disponibilidad léxica en tres centros de interés en una escuela vulnerable en Santiago de Chile. *Nueva Revista del Pacífico*, 76, 111–136. <https://doi.org/10.4067/s0719-51762022000100111>.
- Martínez, R., Cervantes, G., & Jiménez, G. (2022). Quantitative reasoning, language and mathematics. *Zona Próxima*, 36, 76–92. <https://doi.org/10.14482/zp.36.510.71>.
- Mendoza, O., Beato, O., Guillen, J., Pacheco, J., & Palma, G. (2024). La influencia del del lenguaje matemático en el proceso de aprendizaje de las matemáticas. *Ciencia Latina Revista Científica Multidisciplinar*, 8(3), 5980–5992. https://doi.org/10.37811/cl_rcm.v8i3.11795.
- Ministerio de Educación Nacional [MEN] (2016). *Derechos Básicos de Aprendizaje V2. Matemáticas*. MEN. <https://goo.su/nlrBIZ>.
- Molinarés, S., Meriño, A., & Barro, M. (2014). Factores pedagógicos relacionados con el rendimiento académico en estudiantes de cinco instituciones educativas del distrito de Santa Marta, Colombia. *Revista intercontinental de psicología y educación*, 16(2), 151–169. <https://psicologiayeducacion.uic.mx/index.php/1/article/view/163>.
- Moreno, J. (2022). Aplicaciones de la disponibilidad léxica a la lexicografía de aprendizaje en ELE. *Tejuelo*, 35(3), 233–262. <https://doi.org/10.17398/1988-8430.35.3.233>.
- Otero, I., Vizcaino, A., & Carmenates, D. (2015). Creencias epistemológicas de profesores y alumnos sobre la Matemática. *Unión - Revista Iberoamericana de Educación Matemática*, 11(42). <https://union.fespm.es/index.php/UNION/article/view/638>.
- Pacheco, A., Castañeda, S., & Sánchez, S. (2016). Disponibilidad léxica matemática en estudiantes de ingeniería y ciencias. *Unión: Revista Iberoamericana de Educación Matemática*, 47, 44–61. <https://dialnet.unirioja.es/servlet/articulo?codigo=5847525>.
- Puente, A., González, X., & Calderón, J. (2021). Vocabulario matemático: análisis del uso del sistema numérico en estudiantes adolescentes chilenos. *Revista Educa UMCH*, 18, 32–54. <https://doi.org/10.35756/educaumch.202118.203>.
- Rojas, D., Zambrano, C., & Salcedo, P. (2017). Metodología de análisis de disponibilidad léxica en alumnos de pedagogía a través de la comparación jerárquica de lexicones. *Formación Universitaria*, 10(4), 3–14. <http://www.redalyc.org/articulo.oa?id=373552294002>.
- Saiz, M., & Torres, J. (2022). El lenguaje utilizado en la enseñanza y el aprendizaje de las matemáticas en educación rural. *Encuentro Educacional*, 29(1), 113–131. <https://doi.org/10.5281/zenodo.8122054>.
- Sandu, B. (2013). Romanian students' lexical availability in Spanish as a foreign language. *Estudios Interlingüísticos*, 1, 121–133. <https://dialnet.unirioja.es/servlet/articulo?codigo=4515443>.
- Salcedo, P., & del Valle, L. (2013). Disponibilidad léxica matemática en estudiantes de enseñanza media en Concepción, Chile. *Atenas*, 1(21), 1–16. <https://atenas.umcc.cu/index.php/atenas/article/view/606>.
- Santos-Díaz, I. (2020). *El léxico bilingüe del futuro profesorado. Análisis y pautas para estudios de disponibilidad léxica*. Peter Lang.
- Santos-Díaz, I., Trigo, E., & Romero, M. (2020). Propuesta de una taxonomía de los centros de interés en los estudios de disponibilidad léxica. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 36(4), 2020360404. <https://doi.org/10.1590/1678-460X2020360404>.

- Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction*, 49, 157–165. <https://doi.org/10.1016/j.learninstruc.2017.01.005>.
- Strasser, K., & Zamora Mansilla, R. (2024). Vocabulario crítico para recuperación y aceleración de aprendizajes en ciencias y matemáticas en 2° y 3° básico. *Calidad en la Educación*, 60, 29–61. <https://doi.org/10.31619/caledu.n60.1459>.
- Trigo, E., & Santos-Díaz, I. (2019). *El tipo de centro educativo (público o privado) sobre el léxico disponible: análisis de cada subgrupo*. <https://api.core.ac.uk/oai/oai:riuma.uma.es:10630/17813>.
- Trigo, E., Romero, M., & Santos-Díaz, I. (2020). Disponibilidad léxica y dominio de la ortografía: un estudio empírico basado en la influencia de los factores sociales. *Onomázein*, 47, 27–45. <https://doi.org/10.7764/onomazein.47.02>.
- Urzúa, C., Sáez, K., & Echeverría, M. (2006). Disponibilidad léxica en matemáticas: Un análisis cuantitativo y cualitativo. *RLA. Revista de Lingüística Teórica y Aplicada*, 44(2), 59–76. <https://dx.doi.org/10.4067/S0718-48832006000200005>.
- Vargas-Delgado, L., Yepes, D., & Ávila-Toscano, J. (2021). Significado de las medidas de tendencia central en textos de matemática para secundaria. *Praxis*, 17(1). <https://doi.org/10.21676/23897856.3575>.
- Vásquez, C., Díaz-Levicoy, D., & Arteaga, P. (2020). Objetos matemáticos ligados a la estadística y la probabilidad en Educación Infantil: un análisis desde los libros de texto. *Bolema: Boletim de Educação Matemática*, 34(67), 480–500. <https://doi.org/10.1590/1980-4415v34n67a07>.
- Zambrano, C., & Agustín, M. (2023). Un estudio de disponibilidad léxica asociado a las áreas de aprendizaje y motivación por el aprendizaje. *Lengua y Sociedad*, 22(2), 473–492. <https://doi.org/10.15381/lengsoc.v22i2.23855>.

Contribución de los autores

José Hernando Ávila-Toscano: Conceptualización – Análisis formal – Investigación – Metodología – Redacción, revisión y edición

Orlando Junior Zambrano Yepes: Investigación – Metodología – Redacción, revisión y edición

María de los Ángeles Páez Pérez: Investigación – Metodología – Redacción, revisión y edición

Juan Darío Páez Cáceres: Investigación – Metodología – Redacción, revisión y edición

Implicaciones éticas

No existen implicaciones éticas por declarar en la elaboración o publicación de este artículo.

Financiación

Los autores no recibieron apoyo financiero para la elaboración ni para la publicación de este artículo.

Conflictos de interés

Los autores declaran no tener conflictos de interés en relación con la elaboración o publicación de este artículo.

Agradecimientos

Agradecemos la colaboración de las autoridades educativas que facilitaron el acceso a los participantes. También agradecemos a los estudiantes que diligenciaron el instrumento y a sus familiares que consintieron su participación.