

# EVOLUCIÓN DEL LÉXICO DISPONIBLE EN ESTUDIANTES DE EDUCACIÓN BÁSICA: UN ANÁLISIS LONGITUDINAL

## EVOLUTION OF AVAILABLE LEXICON IN BASIC EDUCATION STUDENTS: A LONGITUDINAL ANALYSIS

**Milko Cepeda-Guerra**

milko.cepeda@ucn.cl  
Universidad Católica del Norte

**Paula Urzúa Carmona**

paula.urzua.carmona@edu.udla.cl  
Universidad de las Américas

Recibido: 07/10/2025

Aceptado: 10/12/2025

### Resumen:

Este estudio analiza la progresión del léxico disponible infantil mediante un diseño longitudinal aplicado a estudiantes chilenos desde primero a tercero básico. A través de pruebas de disponibilidad léxica en once centros de interés, se midieron indicadores como promedio de respuestas, número de vocablos, índice de cohesión y frecuencia 1. Los resultados muestran un crecimiento general del repertorio léxico con el avance escolar, aunque no lineal: se observó un estancamiento entre primero y segundo año, seguido de un aumento significativo en tercero. Además, la evolución léxica varió según el centro de interés, con diferencias en estabilidad y cohesión. Se discuten implicancias pedagógicas y se proponen líneas futuras de investigación que incluyan léxico pasivo y producción escrita.

**Palabras clave:** Disponibilidad léxica infantil, crecimiento léxico, estudio longitudinal, cohesión léxica, vocabulario escolar.

### Abstract

This study examines the progression of children's available lexicon through a longitudinal design applied to Chilean students from first to third grade. Lexical availability tests covering eleven semantic categories were administered, measuring indicators such as average word production, number of lexical items, lexical cohesion index, and unique word frequency. Results show a general increase in lexical repertoire with school advancement, although not in a linear fashion: a stagnation was observed between first and second grade, followed by significant growth in third grade. Lexical evolution also varied by semantic field, showing differences in stability and group cohesion. Pedagogical implications are discussed, and future research is encouraged to include passive vocabulary and written production.

**Keywords:** Children's lexical availability, lexical growth, longitudinal study, lexical cohesion, educational lexicon.

## 1. Introducción

La metodología de disponibilidad léxica nace con el afán de complementar los estudios de frecuencia del francés y, desde ahí, abordar la enseñanza de este idioma a ciudadanos migrantes de la época, con el objeto de facilitar su inserción en la sociedad que los acogía (López Morales, 1986). Desde entonces, se ha constatado que, si bien el léxico básico de una lengua permite conocer la proporción real de uso de tipos de palabras, este no resulta suficiente para abordar la totalidad del léxico requerido por una persona para desenvolverse con fluidez en diversos aspectos de la vida cotidiana (Michéa et al., 1953).

Quizás para Michéa y Gougenheim, quienes sentaron las bases de la metodología de la disponibilidad léxica, hubiera sido difícil imaginar la repercusión que sus estudios tendrían en el mundo hispanohablante. Desde esta plataforma, una extensa tradición investigativa en el área se ha centrado durante mucho tiempo en la creación de un diccionario panhispánico de léxico disponible (Hernández & Samper, 2006). Por este motivo, tanto investigadoras e investigadores españoles como latinoamericanos se han abocado a recolectar el léxico disponible de estudiantes preuniversitarios, no solo por la facilidad metodológica para aplicar las pruebas, sino también porque permite conocer este tipo de léxico antes de las influencias de la educación superior.

Esta etapa consolidó la metodología de la disponibilidad léxica en cuanto a centros de interés, tiempos de aplicación y variables sociodemográficas (Samper y Samper, 2006). Sin embargo, también restringió el foco de análisis al grupo preuniversitario, dejando en segundo plano el desarrollo del léxico disponible infantil.

Si bien los estudios de disponibilidad léxica infantil de las décadas de 1970 a 1990 (Echeverría, 1991; López Morales, 1973, 1979; Mena, 1986; Murillo, 1999) sentaron las bases para la creación del diccionario panhispánico de disponibilidad léxica, estos trabajos quedaron rezagados, ya que no contemplaban los centros de interés definidos por el proyecto panhispánico. A pesar de ello, fueron pioneros en estudiar el léxico disponible desde la perspectiva de la lengua materna e identificaron hallazgos que se repetirían en investigaciones posteriores, tales como la restricción del léxico disponible asociada a contextos socioculturales derivados (López Morales, 1973, 1979), las diferencias de género en la producción de vocablos según ciertos centros de interés (Echeverría, 1991; Mena, 1986) y el aumento del caudal léxico a medida que se avanza en la escolaridad (Echeverría, 1991; López Morales, 1973; Mena, 1986; Rojas, 1999).

A mediados de la segunda década del 2000 cuando se retoma con fuerza el estudio de la disponibilidad léxica infantil. Cepeda et al. (2014) recopilaron el léxico disponible de estudiantes de primer año básico en 11 de los 16 centros de interés del panhispánico, destacando diferencias de producción léxica entre niños y niñas. Posteriormente, extendieron su estudio a preescolares (Autor et al., 2017), obteniendo hallazgos similares.

Por otra parte, Escudero Sánchez, R., Santos Díaz, I. C. y Trigo Ibáñez, Ester (2022); Escudero Sánchez, R., Santos Díaz, I. C., y Trigo Ibáñez, E. (2024), y Samper Padilla, J; Hernández Cabrera, C y Samper Hernández, M (2019) presentan estudios donde abordan no solo el estudio de poblaciones similares a la observada en esta investigación, sino que también realizan estudios contrastivos entre España y Latinoamérica, cubriendo diversos centros de interés.

En 2019, Gómez-Devís no solo abordó las estructuras asociativas que emergen en estudiantes de primer año de primaria al evocar vocablos durante una prueba de léxico disponible, dando origen a categorías y subcategorías semánticas, sino que también planteó la necesidad de establecer pautas metodológicas para investigar la configuración del léxico en edades tempranas (Gómez-Devís, 2019, 2021; Gómez-Devís y Herranz-Llácer, 2022).

En paralelo, Samper Padilla et al. (2019) presentaron datos de una muestra de estudiantes de primer año de primaria en Gran Canaria, cuyo objetivo fue identificar el vocabulario disponible a temprana edad. Estos resultados se contrastaron con los de estudiantes costarricenses (Sánchez y Murillo, 2006) y mexicanos (López Chávez y Madrid Guillén, 1993), observando un léxico común predominante con baja presencia de dialectismos.

Pocos años después, Gómez-Devís (2021) comparó resultados cuantitativos de pruebas de disponibilidad léxica aplicadas a estudiantes valencianos, chilenos (Cepeda et al., 2014) y canarios (Samper Padilla et al., 2019), todos en primer año de primaria. Además, analizó las asociaciones semánticas hechas por niñas y niños valencianos en el centro de interés "animales", y más tarde aborda en el centro de interés "colores" (Gómez-Devís y Cepeda, 2022). Ambos estudios reforzaron la idea de que conocer el léxico infantil es el primer paso para expandirlo, al tiempo que posicionan la metodología de disponibilidad léxica como una herramienta valiosa para analizar la progresión cuantitativa del vocabulario en distintos niveles educativos. Recientemente se ha publicado un estudio en el cual se aborda el contacto de lenguas, contraste entre español y catalán, de escolares de 8 años (Gómez-Devís, 2025).

Aunque diversas investigaciones han abordado la evolución cuantitativa del léxico disponible en educación primaria y secundaria (Echeverría, 1991; Gómez-Devís, 2024; López Morales, 1973, 1979; Martí-Contreras, 2024; Mena, 1986; Rojas, 1999), la mayoría han sido estudios transversales que solo permiten observar una proyección del crecimiento léxico con base en grupos de características similares, pero no revelan el proceso real de progresión en el tiempo.

Cabe precisar que no todas las investigaciones encuentran una relación directamente proporcional entre los años de escolaridad y el aumento del léxico disponible. Por ejemplo, Rojas (1999), en su estudio transversal con estudiantes desde kínder hasta tercero básico, observó que entre primero y segundo básico no se registró un incremento léxico, aunque sí lo hubo entre kínder y primero, y entre segundo y tercero básico. Esto sugiere que algunos tramos podrían estar más influenciados por procesos de socialización que por la escolarización misma.

A partir de los antecedentes expuestos, esta investigación se propone aportar evidencia empírica sobre la progresión cuantitativa del léxico disponible infantil a través de un diseño longitudinal. Para ello, se analiza el comportamiento léxico de una muestra de estudiantes durante sus tres primeros años de educación básica. El objetivo es observar si existe un aumento sostenido del caudal léxico evocado en función del avance escolar, y describir el comportamiento de dicha progresión a lo largo del tiempo. El estudio busca, además, sentar una base descriptiva sólida que contribuya al diseño de estrategias pedagógicas orientadas a potenciar el vocabulario disponible en la infancia temprana. En coherencia con este propósito, se plantea como hipótesis que, a medida que las y los estudiantes aumentan sus años de escolaridad, también aumenta su repertorio de léxico disponible.

## 2. Métodos

### 2.1 Participantes

La muestra estuvo compuesta por 35 escolares (24 niños y 11 niñas) provenientes de un colegio particular subvencionado de la región del Maule, Chile. Al inicio del estudio, cursaban primer año de educación básica y tenían entre 6 y 7 años.

Los criterios de inclusión fueron estar inscrito en primer año básico, presentar un desarrollo neurotípico, participar en todas las evaluaciones del estudio, completar íntegramente la prueba de disponibilidad léxica, contar con el asentimiento informado de las y los estudiantes, así como con el consentimiento informado de sus madres, padres o cuidadores.

Los criterios de exclusión incluyeron haber repetido el curso, presentar un diagnóstico que se apartara del desarrollo neurotípico, ausentarse en alguna de las evaluaciones, no completar alguna de las pruebas aplicadas, manifestar rechazo o disentimiento frente a la prueba, o no contar con el consentimiento parental requerido.

Durante el primer año de aplicación participaron 47 estudiantes que cumplían con los criterios de inclusión. En el segundo año, dos niñas se trasladaron de establecimiento, cuatro estudiantes no otorgaron su asentimiento y dos estuvieron ausentes durante la aplicación. En el tercer año, un niño fue diagnosticado con un trastorno del desarrollo, una niña retiró su asentimiento, y otros dos estudiantes se cambiaron de colegio. Así, la muestra final quedó compuesta por 35 estudiantes que participaron en las tres etapas del estudio a partir de la última semana de noviembre de 2017.

Respecto al nivel sociocultural, se puede afirmar que las y los participantes pertenecían a un contexto socioeconómico medio, dado que asistían a un colegio particular subvencionado ubicado en un área urbana. Este dato fue confirmado tanto por la localización del establecimiento como por el perfil de escolaridad e ingreso declarado por las familias.

## 2.2 Procedimiento

Primero se estableció contacto con la dirección del establecimiento, se presentó el proyecto de investigación y se evaluó su factibilidad. Luego, se realizó una reunión informativa con padres, madres y apoderados, en la cual se expusieron los objetivos, características y condiciones del estudio, destacando la voluntariedad de participación.

A continuación, se llevó a cabo un encuentro con los y las estudiantes, donde se les explicó el propósito del estudio en un lenguaje accesible, recalcando su derecho a participar libremente y a desistir en cualquier momento sin consecuencias.

La aplicación de la prueba se realizó durante la última semana de noviembre de cada año, con el fin de asegurar la mayor asistencia posible antes del cierre del año escolar. La administración fue oral, en sesiones individuales de dos minutos cronometrados por cada centro de interés. Los y las evaluadoras eran personas conocidas por el estudiantado, lo que favoreció un ambiente de confianza.

Es necesario mencionar que, si bien el estudio estaba proyectado para cuatro años con el fin de recolectar antecedentes del total del primer ciclo básico, lamentablemente, debido a la pandemia por COVID-19, la investigación tuvo que ser suspendida.

## 2.3 La prueba

La prueba utilizada se basó en 11 de los 16 centros de interés propuestos por el Proyecto Panhispánico de Disponibilidad Léxica: "partes del cuerpo", "ropa y calzado", "alimentos y bebidas", "medios de transporte", "animales", "juguetes", "ciudad", "casa", "profesiones", "escuela" y "colores". El orden de los centros se mantuvo constante para evitar interferencias asociativas entre dominios semánticos adyacentes.

## 2.4 Edición de los resultados

Las respuestas obtenidas fueron transcritas y sistematizadas en una base de datos siguiendo los criterios convencionales de procesamiento léxico:

a) se escribió todo en minúsculas,

- b) los sustantivos y adjetivos se consignaron en singular,
- c) los diminutivos se normalizaron a su forma base,
- d) las formas verbales se remitieron al infinitivo (excepto gerundios y participios), y
- e) las palabras compuestas se escribieron unidas por guiones.

## 2.5 Procesamiento de los datos

Los datos fueron procesados mediante el software Dispogen II (Echeverría et al., 2005), diseñado para el análisis cuantitativo de estudios de disponibilidad léxica. A través de este programa se obtuvieron tres indicadores principales: el promedio de respuesta por centro de interés (PR), el índice de cohesión (IC) y el número de vocablos (NV). Estos tres indicadores constituyen la base del análisis estadístico y permiten observar la progresión cuantitativa del léxico disponible en función del avance escolar.

## 3. Resultados

Antes de presentar los análisis inferenciales, se exponen los promedios de producción léxica (PR), número de vocablos (NV) e índices de cohesión (IC) por curso en cada uno de los once centros de interés. Estos resultados descriptivos permiten contextualizar la evolución del repertorio léxico disponible desde primero a tercero básico.

Posteriormente, se detalla el análisis de varianza para medidas repetidas (ANOVA-RM), que permite evaluar diferencias significativas en la producción léxica entre los niveles escolares. Se complementa este análisis con indicadores cualitativos como el índice de cohesión léxica (IC) y la proporción de vocablos con frecuencia 1 (F1), los cuales ofrecen una visión más profunda de la estructura interna del léxico disponible infantil.

### 3.1 Promedio de respuesta

La tabla 1 muestra los promedios de respuesta obtenidos por los niños y niñas durante los tres años de estudio en todos los centros de interés. Dichos promedios se dispusieron de acuerdo con el orden de producción léxica desde el más al menos productivo. De esta forma es posible observar que los centros más productivos, según su promedio de respuesta, fueron “Animales”, “Alimentos y bebidas”, “Casa” y “Partes del cuerpo”. En oposición a ellos, los centros menos productivos fueron “Profesiones”, “Juguetes” y “Medios de transportes”. En cuanto a los centros más y menos productivos no hubo variación durante los tres años de investigación.

Tabla 1

*Promedio de respuestas léxicas por centro de interés y nivel escolar, en orden de productividad*

Orden	Centro de interés	1° básico	2° básico	3° básico	Prom. general
1	Animales	16.3	13.4	16.8	15.5
2	Alimentos y bebidas	14.1	13.7	17.4	15.06
3	Casa	14.1	14.5	15.3	14.6
4	Partes del cuerpo	12.2	12.4	16.1	13.5
5	Escuela	13.4	12.2	14.3	13.3

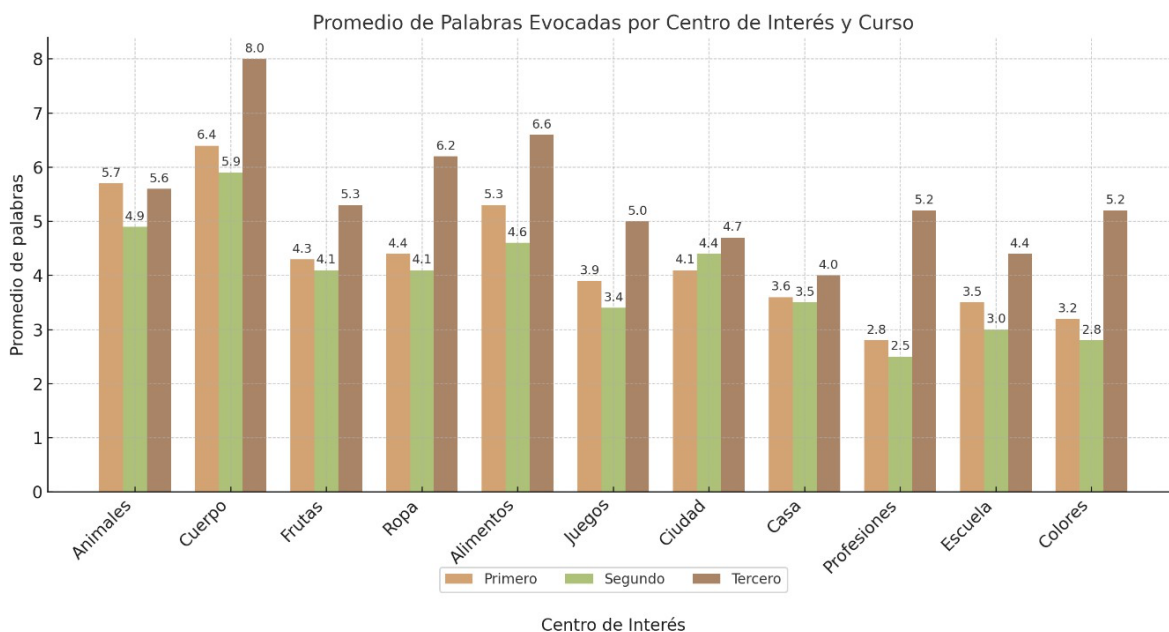
6	Colores	11.6	12	14.1	12.5
7	Ciudad	11.7	11.1	13.5	12.1
8	Ropa y calzado	10.7	10.3	13.2	11.4
9	Medios de transporte	8.8	7.7	10.6	9.03
10	Juguetes	8.7	7.7	10.1	8.8
11	Profesiones	8.3	6.3	8.9	7.8

**Nota.** Los centros de interés están ordenados de mayor a menor según el promedio de respuestas en cada nivel escolar. La columna "Prom. General" corresponde al promedio simple entre los tres niveles. Elaboración propia.

Con el fin de mostrar un primer acercamiento hacia la progresión del léxico disponible según los años de escolaridad de los y las estudiantes, se confeccionó el gráfico de la figura 1 el cual insinúa un decaimiento del léxico disponible en todos los centros de interés a excepción del centro "Ciudad", desde primero a segundo año básico, repuntando en tercero básico. Es necesario mencionar el caso del centro "Animales", único centro en que el léxico disponible baja su productividad a partir de segundo año, dando nula muestra de un aumento durante el tercer año escolar de los y las estudiantes. Volveremos sobre estas particularidades en el análisis de ANOVA posterior.

Figura 1

*Progresión del promedio de respuesta en todos los centros de interés en todos los cursos*



**Nota.** Los centros de interés muestran un crecimiento no lineal a medida que los y las estudiantes avanzan en años de escolaridad. Elaboración propia.

### 3.2 Número de vocablos

Como se puede apreciar en la tabla 2, los centros de interés con mayor cantidad de vocablos, entendiendo vocablo como una unidad léxica producida por los y las estudiantes, sin considerar repeticiones dentro de un mismo corpus, fueron "Ciudad", "Casa" y "Juguetes", y los centros menos productivos en cuanto a vocablos fueron "Colores", "Medios de transportes" y

“Partes del cuerpo”.

Si se observa la tabla con mayor detención será posible notar que, por ejemplo, el centro de interés “Ropa y Calzado” cuenta con un mayor número de vocablos en primero básico que “Partes del cuerpo”. No obstante, el hecho de que “Ropa y calzado” aumentara el número de vocablos en primero y segundo básico de manera sostenida, al contrario de “Partes del cuerpo”, provocó que el orden de producción de vocablos se hiciera a partir del promedio general de los resultados de los tres años que duró la investigación.

Otro aspecto que es necesario mencionar tiene relación con la evolución de la producción de vocablos durante los tres años de escolarización, dado que a excepción de los centros “Ropa y calzado” y “Medios de transporte”, todos los demás muestran una caída en la producción de vocablos desde primero a segundo básico, aumentando en tercero básico.

Tabla 2

*Cantidad de vocablos según orden de productividad por centro de interés y nivel escolar*

Orden	Centro de interés	1° básico	2° básico	3° básico	Prom. general
1	Ciudad	161	180	196	179
2	Casa	168	174	159	167
3	Juguetes	144	133	183	153.3
4	Alimentos y bebidas	137	131	159	142.7
5	Animales	139	111	147	132.3
6	Escuela	129	133	132	131.3
7	Profesiones	118	102	122	114
8	Ropa y calzado	69	71	81	74
9	Partes del cuerpo	85	65	90	73.7
10	Medios de transporte	51	53	53	52.3
11	Colores	37	36	60	44.3

**Nota.** Los centros de interés están ordenados de mayor a menor según la cantidad de vocablos en cada nivel escolar. La columna "Prom. General" corresponde al promedio simple entre los tres niveles. Elaboración propia.

### 3.3 Índice de cohesión

El índice de cohesión entendido como la medida que expresa el grado de coincidencia léxica entre repertorios individuales de un grupo de hablantes, se relaciona directamente con la cantidad de vocablos diferentes producidos: a menor diversidad léxica, mayor cohesión, y viceversa. La tabla 3 presenta los valores del índice organizados de mayor a menor cohesión promedio, lo que permite observar qué centros de interés tienden a generar mayor coincidencia entre participantes. En esta distribución, los centros con menor cantidad de vocablos distintos presentan mayores niveles de cohesión, como es el caso de “Colores”, “medios de transporte” y “Partes del cuerpo”, mientras que “Juguetes”, “Profesiones” y “Ciudad” exhiben los valores más bajos del índice.

Tabla 3

*Rendimiento del índice de cohesión según centro de interés y nivel escolar*

Orden	Centro de interés	1°básico	2° básico	3° básico	Prom. general
1	Colores	0.1272	0.3341	0.2352	0.2946
2	Medios de transporte	0.1731	0.145	0.173	0.1729
3	Partes del cuerpo	0.1439	0.1492	0.1715	0.1715
4	Ropa y calzado	0.1549	0.1449	0.1541	0.1541
5	Animales	0.1176	0.121	0.1123	0.1176
6	Alimentos y bebidas	0.1028	0.1047	0.1085	0.1055
7	Escuela	0.1039	0.0915	0.108	0.1011
8	Casa	0.0842	0.0833	0.0935	0.0879
9	Ciudad	0.0729	0.0662	0.0688	0.0693
10	Profesiones	0.0707	0.0616	0.0726	0.0683
11	Juguetes	0.0605	0.0576	0.0577	0.0577

**Nota.** Los centros de interés están ordenados de mayor a menor según el valor del índice de cohesión léxica por curso. Este índice se interpreta como un indicador del rendimiento léxico grupal, donde valores más altos reflejan mayor diversidad de vocabulario. Elaboración propia.

En cuanto a los valores por nivel escolar, en primer año básico el IC fluctúa entre 0.0605 (Juguetes) y 0.1731 (Medios de transporte). En segundo básico, el valor más alto se observa nuevamente en “Colores” (0.3341), y el más bajo en “Juguetes” (0.0576). En tercero básico, los valores oscilan entre 0.0577 (Juguetes) y 0.2352 (Colores). En términos generales, los promedios más altos se registran en los centros “Colores (0.2946), “Medios de transporte” (0.1729) y “Partes del cuerpo” (0.1715), mientras que los valores más bajos corresponden a “Juguetes” (0.0577), “Ciudad (0.0693) y “Profesiones” (0.0683).

A partir de estos resultados descriptivos, se procedió a realizar un análisis inferencial que permite determinar si las diferencias observadas entre niveles escolares eran estadísticamente significativas.

### 3.4 Análisis de Varianza de medidas repetidas

Se realizó un análisis de varianza para medidas repetidas (ANOVA-RM) (López-Roldán & Fachelli, 2015) con el objetivo de comparar la producción léxica de estudiantes de primero, segundo y tercer año básico en los once centros de interés. La variable dependiente fue el número de palabras evocadas por centro, y se utilizó el modelo multivariado con traza de Pillai como estadístico robusto. El cumplimiento del supuesto de esfericidad fue evaluado mediante la prueba de Mauchly, sin requerir correcciones por violación de este supuesto en ningún caso.

Los resultados mostraron diferencias significativas en 8 de los 11 centros analizados. En particular, se observaron efectos estadísticamente significativos en C1 (Cuerpo):  $F(2,33) = 12.331$ ,  $p = .0001$ ,  $\eta^2 = .428$ , potencia = .993, C2 (Frutas):  $F(2,33) = 6.106$ ,  $p = .0055$ ,  $\eta^2 = .270$ , potencia = .857, C3 (Ropa):  $F(2,33) = 10.854$ ,  $p = .0002$ ,  $\eta^2 = .397$ , potencia = .984, C4 (Alimentos):  $F(2,33) = 11.658$ ,  $p = .0001$ ,  $\eta^2 = .414$ , potencia

= .990, C5 (Animales):  $F(2,33) = 3.722$ ,  $p = .0350$ ,  $\eta^2 = .184$ , potencia = .641, C6 (Juegos):  $F(2,33) = 5.653$ ,  $p = .0077$ ,  $\eta^2 = .255$ , potencia = .828, C9 (Profesiones):  $F(2,33) = 9.759$ ,  $p = .0005$ ,  $\eta^2 = .372$ , potencia = .973 y C11 (Colores):  $F(2,33) = 9.751$ ,  $p = .0005$ ,  $\eta^2 = .371$ , potencia = .973.

En estos centros, el patrón fue consistente, mientras las diferencias entre primero y segundo básico fueron en general no significativas o negativas, se observó un incremento sustantivo en tercer año, evidenciando un salto en la disponibilidad léxica a partir de ese nivel.

En contraste, no se encontraron diferencias significativas en los centros: C7 (Ciudad):  $p = .267$ ,  $\eta^2 = .077$ , potencia = .275, C8 (Casa):  $p = .481$ ,  $\eta^2 = .043$ , potencia = .166 y C10 (Escuela):  $p = .070$ ,  $\eta^2 = .149$ , potencia = .526. Estas categorías presentaron aumentos leves o estancamientos, con potencias insuficientes para detectar diferencias sutiles.

En el centro de interés “Animales” (C5), las comparaciones por pares con ajuste de Bonferroni (Field, 2013) revelaron una caída significativa en segundo año respecto de primer año ( $p = .044$ ) y 2014 ( $p = .007$ ), mientras que no se observaron diferencias entre primer y tercer año ( $p = .751$ ). Este comportamiento sugiere una desviación puntual del patrón general.

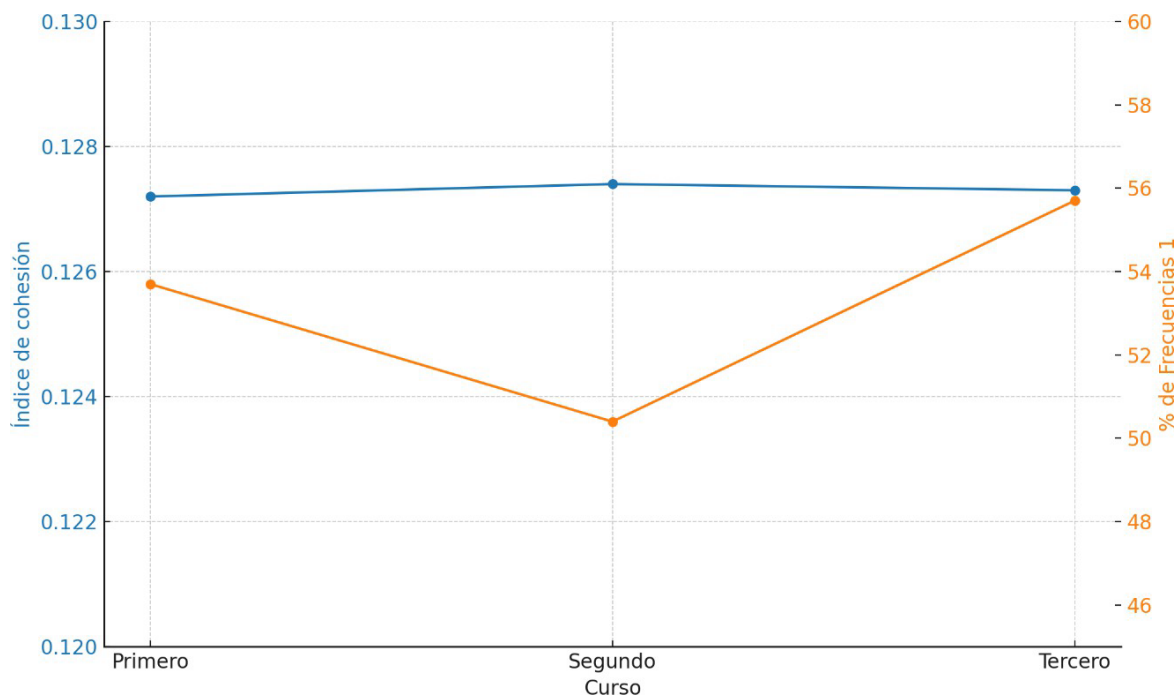
### 3.5 Análisis complementario de cohesión léxica y frecuencia 1

Además del análisis inferencial de varianza, se desarrolló un análisis complementario orientado a explorar la estructura interna del léxico disponible a partir de dos indicadores: el índice de cohesión y la proporción de palabras con frecuencia 1 (F1). La pertinencia de este análisis radica en la necesidad de examinar la calidad del repertorio léxico más allá de la mera cantidad de palabras evocadas. En efecto, el aumento en la cantidad de vocablos no siempre implica un avance cualitativo si estos se distribuyen de forma altamente dispersa o son únicos dentro del grupo. La presencia elevada de vocablos con frecuencia 1 —es decir, palabras mencionadas por un solo estudiante— puede indicar una falta de consolidación léxica compartida, especialmente en niveles donde se espera progresión. Por ello, ambos indicadores permiten observar con mayor precisión las trayectorias de evolución o estancamiento del léxico disponible a nivel grupal.

Para la comparación entre cursos, se aplicó la prueba de Friedman (Siegel & Castellan, 1988), dado el carácter no paramétrico de los datos y la dependencia entre las mediciones. Este enfoque permitió determinar que no existían diferencias estadísticamente significativas en los valores de IC y F1 a lo largo de los tres años escolares: Índice de cohesión (IC):  $\chi^2(2) = 3.09$ ,  $p = .213$  Frecuencias 1 (F1):  $\chi^2(2) = 1.64$ ,  $p = .441$ . Sin embargo, ambas medidas mostraron trayectorias similares: una leve caída en segundo básico seguida de una recuperación en tercero. Esta tendencia se representa gráficamente en la Figura 2.

Figura 2.

*Evolución del índice de cohesión (IC) y porcentaje de palabras con frecuencia 1 (F1) por curso*



**Nota.** Ambos indicadores muestran una leve caída en segundo básico y una posterior recuperación en tercero, evidenciando un patrón de evolución no lineal del léxico disponible infantil. Elaboración propia.

#### 4. Discusión

Los resultados obtenidos permiten observar una progresión general en el léxico disponible a medida que avanza la escolaridad, lo que se manifiesta en un aumento del promedio de palabras evocadas, así como en un crecimiento del repertorio léxico no repetido (vocablos). No obstante, esta progresión no es completamente lineal. El análisis muestra un estancamiento entre primer y segundo año básico en la mayoría de los indicadores, seguido de un aumento significativo en tercero básico, fenómeno que se replica incluso en centros de interés muy productivos, como “Alimentos”, “Partes del cuerpo” o “Animales”. Esta irregularidad pone en cuestión la idea de un desarrollo léxico puramente acumulativo y plantea la posibilidad de transiciones intermedias complejas entre niveles, particularmente en el segundo año escolar.

La comparación entre centros permite constatar que, si bien existe una tendencia general al aumento del repertorio léxico, algunos campos semánticos presentan mayor estabilidad y cohesión que otros. Por ejemplo, “Colores” y “Medios de transporte” son centros con menor cantidad de vocablos y mayor índice de cohesión, mientras que “Juguetes”, “Profesiones” y “Ciudad” muestran mayor dispersión. Este comportamiento se vincula con la naturaleza de los campos semánticos activados, ya que los centros más cerrados tienden a generar mayor coincidencia entre hablantes, mientras que los más abiertos provocan una mayor diversidad de respuestas. Como lo han señalado Gómez-Devís (2021) y Gómez-Devís y Cepeda (2022), cada centro de interés activa un universo temático particular, con diferentes grados de amplitud, familiaridad y estructuración, lo que repercute en la cohesión léxica del grupo.

En este contexto, el análisis inferencial mediante ANOVA de medidas repetidas permitió establecer diferencias significativas en la mayoría de los centros de interés, confirmando el

aumento del léxico disponible a lo largo de la escolaridad. Sin embargo, hubo excepciones notables. En el centro de interés “Animales”, las comparaciones por pares revelaron una caída significativa en segundo año respecto de primer año y de tercero, lo que sugiere una desviación puntual del patrón general observado en otros centros.

Además del análisis inferencial de varianza, se desarrolló un análisis complementario orientado a explorar la estructura interna del léxico disponible a partir de dos indicadores: el índice de cohesión y la proporción de palabras con frecuencia 1 (F1). Este análisis se justifica en la necesidad de profundizar en la calidad del repertorio léxico más allá del número de palabras evocadas. En efecto, el aumento en la cantidad de vocablos no siempre implica un avance cualitativo si estos se distribuyen de forma altamente dispersa o son únicos dentro del grupo. La presencia elevada de vocablos con frecuencia 1 —es decir, palabras mencionadas por un solo estudiante— puede indicar una falta de consolidación léxica compartida, especialmente en niveles donde se espera progresión. Por ello, ambos indicadores permiten observar con mayor precisión las trayectorias de evolución o estancamiento del léxico disponible a nivel grupal.

El análisis de estos indicadores mediante la prueba de Friedman no reveló diferencias significativas entre los tres cursos, sin embargo, sí permitió visualizar una trayectoria regular: una leve caída en segundo básico seguida de una recuperación en tercero, lo cual coincide con lo observado en el análisis de vocablos y promedio de respuestas. Esta convergencia sugiere que los fenómenos de estancamiento y posterior progreso se manifiestan en distintos niveles del léxico disponible y no solo en el total de palabras evocadas.

Esta consistencia en los patrones identificados da mayor fuerza al hallazgo del estancamiento entre primero y segundo año básico, ya anticipado por Murillo (1999) en un estudio transversal, pero ahora confirmado en un diseño longitudinal. La ventaja metodológica de esta investigación radica precisamente en que el seguimiento a un mismo grupo de estudiantes elimina la posibilidad de que las diferencias se deban a variaciones interindividuales o cohortes distintas, lo que fortalece la hipótesis de que existe un punto de inflexión en el desarrollo léxico infantil que merece mayor atención.

Esta evidencia plantea preguntas fundamentales para futuras investigaciones: ¿Qué factores podrían estar incidiendo en este estancamiento? ¿Se relaciona con el proceso de adquisición de la lectoescritura, con cambios curriculares o con características cognitivas propias de esa etapa escolar? ¿Se manifiesta también en otras dimensiones lingüísticas, como la comprensión, la producción escrita o el léxico pasivo? Estos interrogantes abren la puerta a estudios complementarios que permitan profundizar en la comprensión del desarrollo léxico infantil.

Finalmente, es importante señalar que este estudio, en coherencia con la metodología clásica de disponibilidad léxica, se ha centrado exclusivamente en la evocación de sustantivos. Si bien esto ha permitido observar con claridad el repertorio nominal disponible en niños y niñas de los primeros años de escolaridad, deja fuera otras dimensiones fundamentales del desarrollo lingüístico. Como advierte López Morales (1986), la riqueza expresiva de los hablantes no se reduce a los sustantivos disponibles, sino que también se construye mediante la expansión gramatical, el vocabulario pasivo y especializado, y el uso funcional del lenguaje en diversos contextos.

Como lo han señalado Gómez-Devís (2021, 2024) y Gómez-Devís y Cepeda (2022), cada centro de interés activa un campo semántico con propiedades diferentes. Algunos presentan repertorios acotados y estables —como “colores” o “medios de transporte”— mientras que otros, como “animales” o “profesiones”, se caracterizan por una gran apertura y variabilidad contextual, lo que puede generar una mayor dispersión léxica y menor cohesión grupal. Esta diversidad semántica plantea la necesidad de incorporar análisis más finos por campo léxico, con el fin de comprender en profundidad cómo y dónde se produce la progresión —o el estancamiento— del léxico disponible.

En este sentido, la línea futura de investigación no solo debería continuar observando longitudinalmente la progresión nominal, sino también integrar pruebas complementarias que exploren el léxico pasivo, la comprensión lectora, la producción escrita y otros indicadores del desarrollo lingüístico general. Solo así será posible delinear una trayectoria más completa del crecimiento léxico en la infancia escolarizada y entender cómo se entretajan los múltiples factores que inciden en ella.

## 5. Conclusiones

Los resultados de este estudio confirman que el repertorio de léxico disponible infantil tiende a expandirse con el avance escolar, aunque dicha progresión no es uniforme ni constante. El crecimiento más significativo se observa entre segundo y tercer año de educación básica, mientras que entre primero y segundo año se aprecia un estancamiento, e incluso retrocesos puntuales en algunos centros de interés. Este patrón matiza la idea de un desarrollo léxico lineal, y sugiere que ciertos hitos curriculares —como la formalización de la lectoescritura— pueden reorientar transitoriamente los recursos cognitivos, afectando la fluidez y disponibilidad del vocabulario.

A nivel grupal, el análisis de la cohesión léxica y la proporción de palabras con frecuencia 1 refuerza esta lectura: la dispersión léxica y la baja coincidencia entre hablantes son indicios de que el aumento en vocablos no siempre implica una consolidación léxica compartida. Asimismo, los datos evidencian trayectorias dispares entre centros de interés, lo que refuerza la necesidad de incorporar análisis cualitativos que consideren la naturaleza semántica, la carga cultural y la experiencia de uso asociadas a cada categoría.

Entre las principales limitaciones se encuentran el carácter acotado y homogéneo de la muestra, la pérdida de participantes en el seguimiento y la imposibilidad de extender la observación a niveles superiores debido a la pandemia. Estas restricciones limitan la generalización de los hallazgos, pero no invalidan su valor como aporte empírico original al estudio longitudinal de la disponibilidad léxica infantil.

Futuros estudios debieran ampliar la cobertura etaria y sociocultural, incorporar centros de interés verbales y desarrollar análisis complementarios sobre léxico pasivo, comprensión lectora y producción escrita. Solo así será posible comprender con mayor profundidad cómo se construye, transforma o detiene el léxico disponible en los primeros años del desarrollo lingüístico escolar.

## Referencias

- Cepeda, M; Granada, M. y Pomés, M. (2014). Disponibilidad léxica en estudiantes de primero básico. *Literatura y lingüística*, 30,166-181. Universidad Católica Silva Henríquez, Chile. <http://dx.doi.org/10.4067/S0716-58112014000200010>
- Cepeda, M. et al. (2017). Índices de disponibilidad léxica en alumnos de prekínder en un contexto rural de la comuna de Maule, Chile. *Revista Interpretos* 10(18), 183-199. <https://repositorio.ucm.cl/handle/ucm/1303>
- Escudero Sánchez, R., Santos Díaz, I. C., y Trigo Ibáñez, E. (2022). Evolución del léxico disponible sobre la “Escuela” del alumnado de Educación Primaria según el tipo de centro educativo. *Cultura, Lenguaje y Representación*, XXVIII, 61-81. <https://doi.org/10.6035/clr.6376>

- Escudero Sánchez, R., Santos Díaz, I. C., y Trigo Ibáñez, E. (2024). El léxico bilingüe del alumnado de educación primaria según el tipo de centro educativo. *Lenguas Modernas*, 64, 121-146.  
<https://lenguasmodernas.uchile.cl/index.php/LM/article/download/77278/78372/302594>
- Echeverría, M. S. (1991). Crecimiento de la disponibilidad léxica en estudiantes chilenos de nivel básico y medio. En H. López Morales (Ed.). *La enseñanza del español como lengua materna* (pp. 61-78). Universidad de Puerto Rico.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage Publications.
- Gómez-Devís, M-B. (2019). A propósito de las redes semánticas en el léxico disponible de escolares de primero de Educación Primaria. *Ogigia: Revista electrónica de estudios hispánicos*, 25, 165-183. <https://doi.org/10.24197/ogigia.25.2019.165-183>
- Gómez-Devís, M. B. (2021). Disponibilidad léxica en niños de 6 años. Alcance y proyección didáctica del corpus léxico infantil. *Cultura, Lenguaje y Representación*, 25, 169-181. <https://dialnet.unirioja.es/servlet/articulo?codigo=8099346>
- Gómez-Devís, M. B. (2024). *Léxico disponible de los escolares valencianos. Alcance, control y evolución en educación primaria*. Octaedro.
- Gómez-Devís, M. B. (2025). El lèxic català i castellà en els escolars valencians de 3r d'educació primària. Caracterització des de la disponibilitat lèxica. *Studia Romanica Posnaniensia*, 52(3), 5-31. <https://doi.org/10.14746/strop.2025.52.3.1>
- Gómez-Devís, M. B., y Herranz-Llácer, C. V. (2022). Léxico disponible de escolares de la etapa primaria o básica: bases y propuesta metodológicas. *Pragmalingüística*, 30, 183-204. <https://doi.org/10.25267/Pragmalinguistica.2022.i30.09>
- Gómez-Devís M. B., Cepeda, M. (2022). Bases para la enseñanza del léxico: mecanismos de asociación y configuración de redes en el léxico disponible infantil. *TEJUELO. Didáctica De La Lengua Y La Literatura. Educación*, 35(3), 105-134. <https://doi.org/10.17398/1988-8430.35.3.105>
- Gómez-Devís, M. B., Cepeda, M y Herranz-Llácer C. V. (2023). Léxico disponible y redes asociativas de escolares de seis años. La configuración del lexicón en las etapas iniciales de adquisición. *Círculo de Lingüística Aplicada a la Comunicación*, 96, 215-227. <https://doi.org/10.5209/clac.81903>
- Hernández Cabrera, C., y Samper Hernández, M. (2006). Nivel educativo e índices de disponibilidad. En J. Blas Arroyo, M. Casanova Ávalos y M. Velando Casanova (Eds.). *Discurso y sociedad: Contribuciones al estudio de la lengua en contexto social* (pp. 791-809). Universitat Jaume I. <https://dialnet.unirioja.es/servlet/libro?codigo=10465>
- Herranz-Llácer, C. V., y Gómez-Devís, M. B. (2022). La investigación en disponibilidad léxica infantil: aplicaciones para la enseñanza de ELE. *Cultura, Lenguaje y Representación*, 28, 83-101. <https://dialnet.unirioja.es/servlet/articulo?codigo=8693911>
- López Chávez, J. y Madrid Guillén, M. T. (1993). *Léxico disponible de primer grado de primaria*. UNAM y Alhambra Mexicana.
- López Morales, H. (1973). *Disponibilidad léxica en escolares de San Juan*. MS.

- López Morales, H. (1979). Disponibilidad léxica y estratificación socioeconómica. En H. López Morales (Ed.). *Dialectología y Sociolingüística. Temas Puertorriqueños* (pp. 173-181). Hispanova de Ediciones
- López Morales, H. (1999). *Léxico disponible de Puerto Rico*. Arco/Libros.
- López-Roldán, P., y Fachelli, S. (2015). *Metodología de la investigación social cuantitativa*. Universitat Autònoma de Barcelona.
- Martí Contreras J. (2024). Disponibilidad léxica en niños en Partes del cuerpo y la alineación con estándares de aprendizaje de lenguas extranjeras. *Círculo de Lingüística Aplicada a la Comunicación*, 97, 245-259. <https://doi.org/10.5209/clac.88441>
- Mena, Mónica. 1986. *Disponibilidad léxica infantil en tres niveles de enseñanza básica*. [Tesis de maestría]. Universidad de Concepción.
- Michéa, R. (1953). Most fréquents et mots diponibles. Un aspecto Nouveau de la statintique du langage. *Les langues modernes*, 47, 338-344.
- Murillo Rojas, M. (1993). Disponibilidad léxica en los preescolares: estudio de cinco campos semánticos. *Káñima, Revista de Artes y Letras*, 27(2), 117-127.
- Murillo Rojas, M. (1994). Comidas y bebidas: estudio de la disponibilidad léxica en preescolares. *Káñima, Revista de Artes y Letras*, 28(2), 117-133.
- Murillo Rojas, M. (1999). Crecimiento de la disponibilidad léxica: Niños de preescolar y primer ciclo de la educación básica costarricense. *Revista de filología y lingüística de la Universidad de Costa Rica*, 25(2), 187-204.
- Rojas, M. M. (1999). Crecimiento de la disponibilidad léxica: niños de preescolar y primer ciclo de la educación básica costarricense. *Revista de filología y lingüística de la Universidad de Costa Rica*, 25(2), 187-204.
- Samper Padilla, J. A., y Samper Hernández, M. (2006). Aportaciones recientes de los estudios de disponibilidad léxica. *Lynx: Panorámica de estudios lingüísticos*, 5, 5-95.
- Samper Padilla, J; Hernández Cabrera, C y Samper Hernández, M (2019). "Disponibilidad léxica en niños grancanarios de 6 años. Comparación con México y Costa Rica", *Ogigia. Revista electrónica de estudios hispánicos*, 25, 139-163. <https://doi.org/10.24197/ogigia.25.2019.139-163>
- Sánchez, V. y Murillo, M. (2006). *Disponibilidad léxica de los niños preescolares costarricenses*. Universidad de Costa Rica.
- Siegel, S., y Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences*. McGraw-Hill.