

DIVERSIDAD DE VOCABULARIO EN LOS PREESCOLARES. APORTES PARA VALORAR SU COMPETENCIA LÉXICA

Marielos Murillo Rojas

RESUMEN

En este artículo se analiza la riqueza léxica de 101 niños, de cuatro años a seis años y once meses, que asistieron durante el año 2008 a seis instituciones de educación preescolar en el Área Metropolitana de Costa Rica. La riqueza léxica es valorada a partir de la diversidad de palabras de contenido léxico que usan los niños (sustantivos, adjetivos y verbos), en producciones textuales de 50 enunciados, calculada mediante la fórmula Pd50.

Palabras clave: vocabulario, evaluación, educación preescolar

ABSTRACT

This article is analyzing the lexical diversity of 101 children from four to six years old and eleven months that attended during the year 2008 to six institutions of preschool education in the Metropolitan Area, Costa Rica. The lexical richness is valued from the diversity of words of lexical content (nouns, adjectives and verbs) in oral narratives of 50 statements, calculated with the Pd50 formula.

Key words: vocabulary, evaluation, preschool education.

1. Introducción

El conocimiento del perfil de desarrollo normal del lenguaje (DNL) constituye un recurso invaluable para educadores, psicólogos, terapeutas del lenguaje, lingüistas, entre otros profesionales preocupados por la educación y estimulación del niño en sus primeros años de vida.

Dra. Marielos Murillo Rojas. Doctora en Formación del Profesorado con énfasis en didáctica de la lengua y la literatura española, Universidad de Extremadura. Profesora de la Escuela de Formación Docente e investigadora del Instituto de Investigaciones Lingüísticas.

Correo electrónico: marielos.murillo@ucr.ac.cr

Recepción: 19- 1- 2010

Aceptación: 20- 2- 2010

El perfil lingüístico abarca diferentes componentes: léxico-semántico, morfosintáctico, fonológico y pragmático, interdependientes al ofrecer una valoración de la competencia lingüística-comunicativa del niño, según su etapa evolutiva. En el presente artículo, nos ocuparemos del nivel semántico, específicamente en cuanto al léxico, pues se sabe que existe una fuerte relación entre el conocimiento de las palabras, la comprensión lectora y la inteligencia verbal.

En Costa Rica, se ha evaluado el vocabulario mediante técnicas de la léxico-estadística: vocabulario básico, vocabulario disponible y fundamental, con el objetivo de ofrecer inventarios léxicos basados en la competencia real del preescolar, para ponerlos, fundamentalmente, al servicio de la planificación lingüístico-curricular.

Esos inventarios léxicos constituyen también recursos de consulta obligatoria en la elaboración de instrumentos de evaluación, tal es el caso del TEVONIC, Test de vocabulario para niños costarricenses, (Murillo Rojas, en prensa) elaborado a partir de dichos estudios. Con el propósito de complementar las valoraciones realizadas hasta el momento, se propone el análisis del lenguaje espontáneo de los niños mediante la prueba Pd50, la cual da cuenta de la diversidad léxica en textos de 50 enunciados.

En la bibliografía relativa a los estudios de riqueza léxica, se mencionan tres maneras de calcular ese índice. Una de ellas consiste en usar un número fijo de palabras que aparecen consecutivamente en una muestra; tal es el caso del programa Exégesis, del Colegio de México, el cual analiza la densidad léxica en textos de 100 unidades. Otros plantean analizar muestras de lenguaje obtenidas en un tiempo determinado, o sea, contar el número de palabras diferentes e iguales en un período, por ejemplo, de cinco o diez minutos de conversación espontánea. La tercera es realizar el conteo de palabras diferentes e iguales (sustantivos, adjetivos y verbos) en textos de cincuenta enunciados.

Cada una de esas maneras de valorar la riqueza léxica ha sido objetivo de análisis y discusión entre quienes consideran necesario analizar las producciones orales espontáneas de los niños, en pro de contar con una valoración objetiva del “desarrollo normal del lenguaje”.

Aguado (1995) señala las limitaciones de los resultados obtenidos a partir de los estudios hechos con un número fijo de palabras, con el parámetro tiempo y con textos menores a cincuenta enunciados, para concluir que el Pd50 (número de palabras distintas en los primeros 50 enunciados) es la medida idónea al evaluar el vocabulario de los niños en sus aspectos cuantitativos y de diversidad.

Dado lo anterior, nos acogemos a la valoración de la diversidad léxica mediante el Pd50, con una muestra de 101 niños, de cuatro años a seis años y once meses que asistieron durante el año 2008 a seis instituciones de educación preescolar en el Área Metropolitana.

2. Fundamentos teóricos

2.1. Desarrollo del léxico inicial

De acuerdo con Aguado (1995), antes de la aparición del lenguaje propiamente dicho, el niño, inserto en un ambiente sociocultural rico en interacciones, ha acumulado una serie de experiencias intersensoriales y afectivas que le proporcionan herramientas suficientes para emitir, alrededor del primer año de vida, las primeras palabras. Indudablemente, se trata de un gran trabajo compartido entre el niño y el adulto: aprender a comunicarse adecuadamente mediante el

lenguaje. En consecuencia, las actividades del niño durante el primer año de vida son sociales, comunicativas, afectivas y cargadas de significatividad; en primera instancia, es la madre quien asigna significados a los gestos, movimientos, miradas y llantos del recién nacido.

Ahora bien, a partir del despliegue de una serie de procesos cognitivos durante el primer año de vida, de la afectividad con que parece venir dotado el niño para expresar sus necesidades básicas y de la dimensión funcional de las interacciones, en la que se resalta el papel de los formatos en terminología de Bruner (1983), se crea la plataforma para el inicio de la etapa lingüística, que coincide con el segundo año de vida.

Aguado (1995) construye la siguiente definición para referirse al formato:

[...] una pauta de interacción estandarizada e inicialmente microcósmica, entre un adulto y un niño, que contiene roles demarcados que finalmente se convierten en reversibles», constituye el procedimiento comunicativo no lingüístico más eficaz que crean el adulto y el niño, antes de que éste adquiera el lenguaje [...] Estos formatos cumplen funciones que luego serán vehiculizados por el lenguaje, en una acción sustitutoria. (Aguado 1995:37-38)

Con el propósito de comprender ampliamente los procesos de adquisición del lenguaje, es necesario puntualizar un poco más en la construcción de esos formatos –procedimientos básicos para poder usar el lenguaje–. En primer lugar, la madre desde el momento del nacimiento actúa como si el recién nacido tuviera creencias, intenciones y deseos, creando un supuesto contexto cognitivo. En segundo lugar, el lenguaje va dirigido al mundo mental del niño (sus deseos y creencias) y se da por supuesto que el niño es capaz de hacer lo mismo, o sea, dirigir su lenguaje al mundo mental del adulto: llora porque quiere solicitar algo.

Esos acontecimientos repetidos permitirán la anticipación de uno u otro elemento de la rutina:

[...] y con ello la construcción en la mente del niño de, ahora sí, supuestos reales sobre el mundo interior del otro, sobre sus intenciones, deseos, etc., en principio solo en secuencias rutinarias. [...] Esta precocidad en la percepción de las intenciones de los demás es de capital importancia para la adquisición del lenguaje porque hace posible que el niño entienda qué hay detrás de unas palabras o de unas frases y ello le va a permitir construir los significados de estas. (Aguado 1995:38)

Pasando a la segunda etapa, desarrollo lingüístico, las primeras palabras forman parte de rutinas específicas en las que cobran sentido, fuera de las que carecen de significado. No obstante, poco a poco, esas mismas palabras se usarán en contextos distintos, lo que indicaría que ya han adquirido el estatus de arbitrariedad y convencionalidad propio de los signos lingüísticos.

Al continuar con el proceso de desarrollo del lenguaje, el niño entra en la etapa de la holofrase, considerada en términos generales, como la expresión de todo un acto verbal mediante una palabra, según las interpretaciones dadas a partir de la asimetría entre la comprensión y la expresión y de la gramática universal con la que todos los seres humanos nacemos. Otras interpretaciones señalan que las palabras del niño expresan de forma clara un mensaje comunicativo que debe ser inferido por el oyente “El niño se comunica con otro y los dos negocian el significado”. (Nino 1992, citado por Aguado 1995). Por otra parte, Nelson (1985), citado por el mismo autor, considera que las palabras expresadas por el niño se refieren más bien a secuencias de acontecimientos completos y que el paso al uso lingüístico de esas palabras se produce cuando el niño es capaz, cognitivamente, de analizar la secuencia de acontecimientos y puede aplicar la palabra a una de las partes resultantes. Desde ese momento, el pequeño podrá separar la emisión de su expresión verbal de ese contexto para aplicarla adecuadamente a otros.

Por otra parte, múltiples investigaciones han puesto de relieve la importancia del *input* lingüístico en la adquisición de los usos del lenguaje y se ha demostrado que una mayor riqueza del vocabulario de los interlocutores adultos y una mayor exposición del niño al *input* adulto, tiene efectos importantes en el crecimiento de su vocabulario.

En suma, el desarrollo funcional del lenguaje se anticipa al estructural. Los adultos ponemos en las primeras conductas del bebé una carga intencional que se da por cierta y se actúa como si el niño estuviera comunicando algo. Ese comportamiento es lo que convierte en estables las interacciones que se construyen a partir de dichas conductas. Por tanto, las funciones comunicativas que el lenguaje va a cumplir empiezan desde los primeros meses y se constituyen en los andamiajes necesarios para la adquisición del lenguaje del niño.

2.2. Las primeras palabras

De acuerdo con Aguado (1995), las primeras palabras propiamente dichas aparecen alrededor de los 12 meses y son puramente referenciales, etiquetas de los objetos. Su función es compartir más la experiencia que el significado; significan la relación del niño con ese objeto determinado. Su crecimiento es lento durante la primera mitad del segundo año, las palabras aparecen fijadas en los contextos de uso y no en otros en que tal vez podrían utilizarse. Los verbos producidos en forma contextualizada se incrementan entre los 14 y 16 meses.

A partir de los 18 meses, la palabra se relaciona con la representación mental de un objeto o una acción, dicho de otra manera, la palabra denota. Esa explosión denominativa responde a las habilidades cognitivas que el niño ha adquirido a través de las experiencias repetidas con ciertos acontecimientos producidos cotidianamente, los que, asumiendo la terminología de Bruner (1983), llamamos formatos más largos y complejos (el ritual del baño, el juego de bola con el padre, la visita al mismo supermercado/pulpería/mercado, etc.). A partir de los dieciocho meses, el incremento del vocabulario se acelera sorprendentemente. La sobreextensión es un fenómeno muy frecuente en este período, el niño generaliza las palabras aprendidas, aquellas que le permiten entresacar alguna característica común con el objeto cuyo nombre desconoce. Las palabras que forman parte del vocabulario durante el segundo año de vida se componen fundamentalmente de sustantivos, verbos y adjetivos.

Se han dado varias explicaciones sobre la adquisición del vocabulario de los niños, pero lo cierto es que esa adquisición no es un trabajo individual. La interacción con las madres y padres, la lectura de libros (técnica efectiva para el desarrollo del vocabulario de los niños de dos años y un buen predictor del aprendizaje de la lectura y la escritura), escenarios multiniños (en especial con edades distintas) y el rol de los hermanos constituyen el *input* necesario para que se dé el aprendizaje de nuevas palabras. En esas interacciones, se hacen converger las formas y las categorías idiosincrásicas del niño con las convencionales.

Por otra parte, se suele considerar que el aprendizaje de las palabras está constreñido o limitado internamente por algún mecanismo para que dicho aprendizaje tenga éxito. La constricción que se suele aceptar para el aprendizaje de nuevas palabras es el contraste léxico: “mecanismo por medio del cual el niño, al oír una nueva palabra, asume que contrasta con las que ya posee en su léxico, porque dos formas tienen significados diferentes, y de esta manera, el niño se ve constreñido a asignar esa nueva palabra a referentes para los cuales no tenga algún nombre” (Aguado 1995:84). Este principio de contraste léxico será efectivo en la medida en que haya cooperación en el acto comunicativo: el niño supondrá que si su interlocutor le dice una palabra que él no conoce, este se referirá a ese objeto para el cual él no tiene nombre.

Otro tipo de restricción para el desarrollo del vocabulario es el procedente de la propia maduración cerebral. Hacia los 18 meses, se produce el máximo pico respecto de la sinaptogénesis, esta crea una mayor capacidad para codificar y almacenar información, pero serán las transacciones del sujeto con el entorno las que permitirán utilizar esa capacidad.

Entre el segundo y el tercer año de vida, se produce un marcado desfase entre comprensión y expresión, normal en todos los niños, y más marcado cuanto menor es, pero más notorio en los verbos que en los nombres. Los adjetivos, por su parte, no han sido ampliamente estudiados desde la perspectiva de la adquisición del lenguaje, pero parece que se incorporan primero los correspondientes a propiedad física de los objetos (forma, tamaño).

Respecto de las estrategias que alrededor del niño se despliegan para que el niño adquiera nuevas palabras, se señalan las siguientes:

- a. Reducción de la densidad léxica: utilización de términos frecuentes con referencia concreta; se prefiere el nombre del objeto completo y no de sus partes.
- b. Estrategia de contraste: los padres, sin darse cuenta de que es más fácil adquirir los pares opuestos, emplean esta estrategia con mucha frecuencia.
- c. Estrategia de sustitución: se introduce el vocabulario de las categorías, mediante la sustitución de un término por el nombre de la categoría a que pertenece: bola/juguete.
- d. Estrategias de proliferación y de dimensión: se introducen las comparaciones: más grande, más pequeño.

Los gestos también forman parte de las ayudas notorias para llevar a cabo la comunicación y ampliar la competencia léxica. En suma, pese a que el niño viene dotado genéticamente para adquirir una lengua, el contexto es de vital importancia para su desarrollo.

2.3. Características generales del vocabulario de los preescolares costarricenses

En Costa Rica, se cuenta con dos estudios de validez nacional sobre el componente léxico en niños preescolares: *Léxico básico preescolar costarricense* –LEBAPRESCO– (2002) y *Léxico disponible del preescolar costarricense* (2006), realizados por Marielos Murillo y Víctor Sánchez; así como una serie de artículos, a manera de cala, que aportan información importante sobre la competencia léxica de este grupo etario.

El léxico, por su condición de componente de la lengua, aumenta como resultado de la edad, la acción del contexto sociocultural en que se desenvuelve y de la afirmación y corroboración de las hipótesis lingüísticas que se va formulando el niño.

Desde el punto de vista cuantitativo, se ha realizado una serie de proyecciones promedio acerca del volumen del vocabulario que posee un niño en diferentes momentos de su etapa evolutiva; se señala que entre la edad del año y medio y los siete años, el vocabulario aumenta hasta llegar a unas 2000 o 2500 unidades, adquisición léxica promedio de un niño normal. Sin embargo, en el *Léxico básico preescolar costarricense* (Murillo y Sánchez, 2002), se logró recuperar 3.191 unidades léxicas o lexemas, emitidos por los niños de cinco y seis años, en sesiones de 30 minutos aproximadamente en las que se proporcionaron discursos relativos a diez mundos léxicos.

Por otra parte, en el estudio *Disponibilidad léxica de los niños preescolares costarricenses* (Sánchez y Murillo, 2006), se recopilan 1.776 sustantivos diferentes, organizados

por universo léxico. Además, se ofrece una distribución léxica por centro de interés o unidad temática, información que permite conocer cuáles son los campos léxicos más productivos para los preescolares, datos valiosísimos al planificar la enseñanza de la lengua en este nivel del sistema educativo. Este estudio también incluye una sección en que se hace una valoración cualitativa del vocabulario obtenido: se agrupan las palabras imprecisas, o sea, aquellos términos que no están en relación jerárquica respecto de cada campo léxico.

Las imprecisiones léxicas también dan cuenta de aquel vocabulario que está en proceso de apropiación semántica, pues el niño todavía no es capaz de abstraer y generalizar algunos rasgos semánticos que le permitan relacionar un significado específico con el significado que le corresponde, por ejemplo, nombrar un referente como heladería con su nombre propio, como es el caso de Pops. La presencia de los términos imprecisos constituye una señal de alerta que indica la etapa en que se encuentra el preescolar en el proceso de adquisición del campo referencial respectivo y, en consecuencia, hacia donde debe orientarse su enseñanza.

De acuerdo con el LEBAPRESCO, las formas más frecuentes entre los niños son sustantivos (1.892), seguidos de los verbos (610) y de los adjetivos (329). La identificación específica de las palabras por categoría gramatical es básica para planificar la enseñanza de ese tipo de palabras, mediante la programación de secuencias didácticas integradas.

Los adjetivos calificativos constituyen un valioso recurso expresivo en la etapa preescolar y muestran la presencia de un recurso usual en el proceso de adquisición de la lengua, pues es posible agruparlos por pares de antónimos, demostrando de esta manera lo que teóricamente se ha considerado: los padres, sin proponérselo, emplean la estrategia de contraste, es más fácil aprender vocabulario por términos opuestos. La presencia de 28 pares de antónimos en el discurso de los preescolares es una muestra de la manera natural en que adquiere el niño el vocabulario. El uso de conceptos que permiten la asociación es mucho más significativo para el niño que otros procedimientos sin esta característica.

Por otra parte, en el vocabulario de los preescolares también se encuentran voces onomatopéyicas, palabras que tratan de imitar sonidos de los cuales se derivan, tales como: cri cri, guau guau, tin tun, chucuchucuchucu.

En suma, la lectura cuidadosa de los materiales citados ofrece un panorama general del comportamiento del vocabulario expresivo de los preescolares costarricenses.

2.4. Evaluación de las conductas lingüísticas en niños: el léxico

El nivel de conocimiento de palabras y conceptos constituye una medida excelente del desarrollo cognitivo-lingüístico y del potencial de aprendizaje a lo largo de la vida escolar (Gardner 1991; McKeown y Curtis 1987). En los primeros años de vida del niño, el conocimiento del mundo se traduce en el conocimiento de palabras (Crais 1990). En los cursos medios y superiores, el conocimiento de las palabras debe traducirse en conocimiento del mundo.

Miguel Puyuelo, Rondal y Wiig, *Evaluación del lenguaje*.

De acuerdo con Puyuelo, Rondal y Wiig (2000), al valorar los principios de la evaluación de la función lingüística, o sea, los procesos de comprensión y de expresión textuales, es fundamental partir de las preguntas: ¿qué es el lenguaje? y ¿qué evaluar?

Desde el punto de vista del proceso, el lenguaje es el producto de la integración de varios subsistemas: el subsistema de los fonemas, el de los lexemas, el gramatical, es decir la morfosintaxis, las regulaciones pragmáticas y la organización discursiva. Cada subsistema dispone de cierta autonomía con respecto a los otros, como lo demuestran las consideraciones actuales sobre la modalidad neurofuncional del lenguaje y las disociaciones observadas en las patologías del lenguaje y particularmente en las disfasias [...]. (Puyuelo, Rondal y Wiig 2000:18)

La autonomía relativa de los subsistemas lingüísticos obliga evidentemente a evaluarlos por separado.

No podría hacerse una evaluación global del lenguaje sin poder precisar qué aspecto queremos medir. La evaluación de cada subsistema debe realizarse según modalidades específicas correspondientes a las características estructurales y funcionales del subsistema examinado. (Puyuelo, Rondal y Wiig 2000:20).

La evaluación se realiza para averiguar el estado de desarrollo de las conductas lingüísticas o saber cuáles están alteradas y en qué grado, con el objetivo de que la información sirva para planificar las actividades educativas específicas que requieran los sujetos evaluados.

El principio, según el cual no podemos evaluar correctamente más de “una cosa a la vez”, debería tener un valor absoluto en toda evaluación. Inversamente, si proponemos una prueba léxica, presentaremos lexemas aislados o incluidos en enunciados que estamos seguros de que podrán ser entendidos sin dificultad por su estructura morfosintáctica. Los principales test de vocabulario cumplen con esas dos variables.

El análisis del lenguaje espontáneo es una tarea difícil que exige conocimientos profundos de psicolingüística y, en el caso de niños o adolescentes, conocimientos profundos sobre el desarrollo del lenguaje. Por este motivo, la evaluación del lenguaje espontáneo raramente se incluye en las pruebas que proponen los test de lenguaje regulares. Otro aspecto por considerar es el siguiente:

[...] la evaluación de la extensión del lenguaje de producción topa con una limitación drástica. Se trata de la imposibilidad para realizar un inventario exhaustivo del vocabulario expresivo de un sujeto. No es posible plantearse situaciones de evaluación en las que se pida al sujeto que nombre todos los componentes del mundo material. Aun cuando nos lo pudiéramos plantear, deberíamos grabar durante 24 horas al día las emisiones del sujeto durante muchos días. Esta limitación se ve atenuada por el hecho de que el vocabulario de producción es cuantitativamente inferior al vocabulario de recepción del mismo sujeto. Generalmente entendemos más palabras de las que utilizamos en nuestro lenguaje. Sin embargo, la relación es difícil de cuantificar y por esto la evaluación léxica se refiere casi siempre al vocabulario de recepción. (Puyuelo, Rondal y Wiig 2000:137)

En nuestro caso, se evaluará la riqueza léxica mediante la prueba Pd50 (diversidad de sustantivos, adjetivos y verbos en textos de 50 enunciados) y se ofrecerán las normas costarricenses para la comparación de los resultados con otras poblaciones.

2.5. La riqueza léxica

Respecto de la estimación de la diversidad del vocabulario producido por los niños, existen distintos tipos de medida: valoración de un número fijo de palabras, conteo del número de palabras producidas en un período de tiempo determinado, sumatoria de las palabras emitidas en un número predefinido de enunciados.

- a. **LMEV –longitud media del enunciado–:** valoración de 50 a 100 enunciados producidos por el niño, a partir de los cuales se calcula la longitud media del enunciado (LME) y se analizan algunos componentes sintácticos de las producciones realizadas.

El rango de aplicación es de 1.5 y 5 años o de 5 a 18 años de edad, según el autor. Se calcula dividiendo el número total de morfemas por el número de producciones.

$$\text{LMEV} = \frac{\text{Número de morfemas}}{\text{Número de producciones}}$$

- b. **IDL –Índice de diversidad léxica–:** análisis de la diversidad léxica en 50 o más enunciados. El rango de aplicación aproximada es de 3 a 8 años. Se calcula dividiendo el número de palabras diferentes por el número de palabras producidas.

$$\text{IDL} = \frac{\text{Número de palabras diferentes}}{\text{Número total de palabras producidas}}$$

- c. **Pd50 –promedio de palabras diferentes–:** conteo de las palabras diferentes en textos de 50 enunciados. Con esta medida, se puede tener una idea más precisa de la riqueza del vocabulario de los niños, además, es un cálculo sensible al aumento de la edad. Se calcula dividiendo el número total de vocablos tipo por el número total de palabras.

$$\text{PD50} = \frac{\text{Número de vocablos tipo}}{\text{Número total de palabras}}$$

En Costa Rica, se han realizado tres investigaciones que evalúan la densidad léxica¹; la primera, Barahona Novoa (1996), determina la riqueza léxica de textos producidos en la radio; la segunda, Murillo Rojas (1998), mide la riqueza léxica de niños preescolares de tres a seis años; la tercera, Leoni de León (1997), analiza la prensa escrita. Todas esas investigaciones se realizaron mediante el programa Exégesis².

Exégesis extrae la riqueza léxica por medio de la densidad de un texto. Utiliza el coeficiente que resulta de dividir el número de palabras diferentes de un corpus entre un segmento determinado del texto, que debe tener 100 palabras gráficas. La densidad permite inferir el total de vocablos que se obtendrían de un corpus breve si tuviera una extensión mayor.

Para efectos de este estudio, se entenderá por riqueza léxica la diversidad de palabras de contenido léxico que usan los niños (sustantivos, adjetivos y verbos), en producciones textuales de 50 enunciados, calculada mediante la fórmula Pd50. Cada participante debe narrar, contar una historia, un cuento o una película, según sus intereses, en un tiempo promedio de cinco a diez minutos. Los textos emitidos por los niños se graban para su posterior transcripción. El texto transcrito se divide en 50 enunciados oracionales y se cuentan sustantivos, adjetivos y verbos como vocablos tipo, para identificar las palabras distintas y se hace otro conteo con el total de esas clases de palabras en el texto.

A continuación, se presenta un ejemplo de un texto de cincuenta enunciados, separados mediante una barra oblicua, con el propósito de que sirva de modelo al profesional que desee valorar la riqueza léxica de niños entre cuatro y seis años y once meses.

Sexo: Femenino.

Edad: 5 años y 7 meses.

Tipo de institución: Kínder público, provincia de Alajuela.

Cuento: Los tres cerditos.

Enunciados: 50, ubicados entre dos líneas oblicuas.

Bueno, que /habían tres chanchitos //que siempre jugaban en un parque/, pero /el hermano mayor le dijo a los hermanos menores/ /que no debían de trabajar tanto/ y entonces /le dijo el hermano mayor/ /que fueran a nuestra casa/ pero /no fue/ porque /el hermano mayor le había dicho/. Eh! /Ya hemos jugado bastante hermanitos/ entonces /se fueron a la casa/ y /era una cueva./ Entonces /fueron a bañarse a un tanquecito/ /que tenía mucha agua/ que era limpia/ y /era ... este bonita/ y entonces /fue un pajarito/ /que se llama Petirrojo/ /que es amigo de ellos dos/, /digo de ellos tres/. /Le dijeron/ —/tengan cuidado amiguitos/, /por ahí anda un lobo/ que /quiere comer/ y /está muy hambriento/ y /yo no pienso/ /que él no quiera comérselos a ustedes/, pero /creo que sí/ /se los va a comer/. Entonces /vayan a la casita/, entonces /terminaron jugando en el río/. /Y muy asustados los hermanitos menores/, entonces /los dos menores salieron corriendo/, entonces /el otro dijo —tranquilos hermanitos/ /no nos alcanzará en la cueva de nosotros/. Verdad, entonces /vieron una huella del lobo/, entonces /cuando él fue de nuevo al paso/ /a dejar un montón de pajarillos/ (...) y entonces /le dijeron/ /—vete malo lobo/, entonces /sí se fue/, pero /mañana iba a regresar/, era. Pero entonces /el hermano menor y el del medio preguntaron/ / que estaban muy asustados/ /—ahora qué vamos a hacer/ y /el hermano mayor dijo/ /que todos los hermanitos vamos a ir a construir una casita/, entonces /abrieron las puertas/ y /salieron/, /uno hizo la de paja, el más chiquitito/, /el del medio la hizo de madera/ y /el otro lo hizo de ladrillos/.

Una vez transcrito el texto, con la correspondiente identificación de los 50 primeros enunciados, se procede a identificar las siguientes palabras: sustantivos, verbos y adjetivos; se registran todas sin importar que estén repetidas. Posteriormente, se cuentan para establecer la frecuencia total.

El vocablo tipo, que viene a coincidir con la palabra tal como se registraría en un diccionario, se identifica de la siguiente manera, por ejemplo, “hermano” en calidad de vocablo tipo o palabra diferente es una, pero en el texto anterior aparecen las siguientes formas: hermano (siete veces), hermanos (una vez) y hermanitos (cinco veces), para un total de trece; “jugar” aparece con las siguientes formas: jugaban, hemos jugado y jugando, frecuencia tres.

3. Metodología de trabajo

Con el objetivo de conocer la riqueza léxica de los preescolares y establecer normas de valoración con niños costarricenses, se trabajó con una población de 237 niños de 3 años a 6 años y once meses, que asistieron durante el año 2008 a seis instituciones de educación preescolar costarricense: dos públicas, dos privadas y dos CEN-CINAI³. De ese total, se seleccionó una muestra de 101 preescolares, que produjeron textos de 50 enunciados, cantidad necesaria para realizar la valoración estadística de la riqueza léxica, mediante el Pd50.

3.1. Variables

- a. **Riqueza léxica:** se entenderá como la variabilidad de palabras de contenido léxico que usan los niños —sustantivos, adjetivos y verbos—, en producciones textuales de 50 enunciados.

- b. **Tipo de institución:** se trabaja con tres tipos de instituciones: pública, privada y CEN-CINAI, todas del Área Metropolitana.
- c. **Rangos de edad:** se trabaja con tres rangos de edad, en los que se ubican, mayoritariamente, los niños costarricenses que asisten a instituciones preescolares. No se consideró el rango de tres años, dada la dificultad de obtener producciones textuales de cincuenta enunciados durante una sesión de trabajo.

3.2. Procedimientos metodológicos y análisis del corpus

Conviene recordar que para este apartado los textos, por excelencia, deben ser discursos narrativos y la muestra original, debido a las características especiales que se requería del texto producido, se redujo a 101 informantes de los 237 originales, distribuidos de la siguiente manera: kínder público 40 niños, kínder privado 40 niños y 20 de CEN-CINAI. Por esa misma razón, los niños de tres años a tres años y once meses, fueron excluidos de la muestra final.

A continuación, se presenta el procedimiento seguido:

1. Cada participante debía narrar, contar una historia, un cuento o una película, según sus intereses. Los textos emitidos por los niños se grabaron para su posterior transcripción.
2. Se hizo una transcripción grafemática, se escribió cada uno de los enunciados emitidos por el niño con la mayor fidelidad posible; además se transcribieron los enunciados del adulto encuestador para clarificar el contexto lingüístico-comunicativo en que se produjo la emisión textual del sujeto investigado.
3. El texto se dividió en enunciados oracionales.
4. Para el cálculo del número de palabras distintas o vocablos-tipo, se contabilizó solo una vez cada lexema: sustantivo, adjetivo y verbo. Por ejemplo: “casa”, “casita” y “casilla”, cuentan como un solo vocablo-tipo “casa”; “comieron”, “comimos”, “comió”, cuentan como el vocablo-tipo “comer”; “suave”, “suavecito”, cuentan como “suave”.
 - 4.1. Para obtener el número de verbos distintos, se contaron una sola vez las formas infinitivas de cada expresión verbal que aparecían en las muestras individuales. Las perífrasis verbales, entendidas como la unión de dos o más formas lingüísticas que funcionan conjuntamente como una unidad verbal, se contaron como un solo verbo.
 - 4.2. Para obtener el número de sustantivos (comunes, colectivos, abstractos, concretos y compuestos) distintos, se contaron los lexemas de cada nombre una sola vez. No se consideraron los nombres propios.
 - 4.3. Para obtener el número de adjetivos distintos, se contaron los lexemas de cada adjetivo una sola vez.
5. El cálculo del número total de palabras –sustantivos, adjetivos y verbos– de cada texto se obtuvo al contar todas las formas de las palabras objeto de estudio: “casa”, “casita”, “casilla”, son tres palabras.

6. El cálculo del Pd50 se obtiene así:

$$\text{Pd50} = \frac{\text{Número de palabras tipo}}{\text{Número total de palabras}}$$

7. Los datos fueron procesados mediante el programa EXCEL

3.3. Resultados

Los resultados se organizan de la siguiente forma:

- Índice de riqueza léxica total para todos los niños independientemente de la edad cronológica.
- Índice de riqueza léxica parcial, según la variable sexo: niños y niñas.
- Índice de riqueza léxica parcial, según la variable tipo de institución: pública, privada y CEN-CINAI.
- Índice de riqueza léxica parcial, según la variable edad: 4, 5 y 6 años

Las comparaciones se realizaron a partir de las medidas: moda, media, mediana, desviación estándar (DS) y calificación máxima (Cali Max)/calificación mínima (Cali Min).

Tabla 1. Diversidad léxica: datos de la población total

Unidad de medida	Promedio de palabras diferentes	Promedio total de palabras	Índice de diversidad léxica
Moda	78	123	0,62
Media	73,92	114,20	0,65
Mediana	75	114,5	0,65
DS	10,11	16,29	0,09
Cali Max	98	163	0,93
Cali Min	53	79	0,43

Como se observa en la tabla anterior, el índice promedio de diversidad léxica⁴ en nuestra muestra es de 0,65, mayor al considerado como normal en la bibliografía consultada, superior a 0,50.

En nuestro caso, se considerará como rango de normalidad a los preescolares que se ubican entre una desviación estándar inferior y una superior de la media:

Superior: mayor a 0,75.

Normal: 0,56 a 0,74.

Inferior: menor a 0,55.

Los textos tuvieron un promedio de 73,92 palabras diferentes o sea vocablos-tipo y un total promedio de 114,20 palabras de discurso, habiéndose tomado en cuenta, se hace la observación una vez más, solo palabras de contenido léxico: sustantivos, adjetivos y verbos, los cuales no se registraron por separado.

Vale la pena resaltar los puntajes máximos y mínimos en la valoración individual de la diversidad léxica: 0,93 y 0,43. De los cuales se deduce que el niño con una densidad de 0,43 debe necesariamente recibir un programa de apoyo para enriquecer su vocabulario.

Respecto de la variable sexo, la muestra con que se trabaja es muy homogénea, considera el aporte de 55 mujeres y de 56 varones, factor que favorece la comparación al tratarse de grupos del mismo tamaño.

Los resultados indican que la variable sexo no es pertinente al valorar la riqueza léxica de los preescolares costarricenses, pues el comportamiento de niños y niñas es prácticamente el mismo.

Por otra parte, la distribución por grupos de edad presenta el siguiente comportamiento: 21 niños de 4 a 4,1; 58 niños de 5 a 5,11 y 22 niños de 6 a 6,11, esta conformación muestral responde a la variable “presencia” de 50 enunciados o más. El resultado de cada grupo de edad se presenta en las tablas 2, 3 y 4.

Tabla 2. Diversidad léxica: variable edad, 4 años

Unidad de medida	Promedio de palabras diferentes	Promedio total de palabras	Índice de diversidad léxica
Moda	78	121	0,60
Media	74,47	114,68	0,65
Mediana	75	115	0,65
DS	9,97	15,49	0,08

Tabla 3. Diversidad léxica: variable edad, cinco años

Unidad de medida	Promedio de palabras diferentes	Promedio total de palabras	Índice de diversidad léxica
Moda	78	123	0,62
Media	73,91	114,53	0,65
Mediana	75	115	0,65
DS	10,16	16,03	0,09

Tabla 4. Diversidad léxica: variable edad, seis años

Unidad de medida	Promedio de palabras diferentes	Promedio total de palabras	Índice de diversidad léxica
Moda	78	123	0,62
Media	73,72	113,06	0,66
Mediana	75	112,5	0,65
DS	9,68	14,28	0,09

Los preescolares más productivos fueron los niños de cinco años, al ofrecer un mayor número de textos completos. No obstante, al comparar los índices de diversidad léxica promedio en cada grupo no se observan diferencias, excepto por un punto en la media del último grupo.

Un resultado similar obtienen Álvarez Durán, Barraza Parraguez, Heredia Tudela, Jonquera Fernández y Olea Ortega (2006), al comparar el índice de diversidad léxica de 30 niños

con desarrollo normal del lenguaje (DNL) y 30 preescolares diagnosticados con la patología “Trastorno específico del lenguaje” (TEL), en el rango de edad cuatro años a cuatro años y once meses, residentes de Santiago de Chile. Para efectos nuestros, tomaremos únicamente los niños con desarrollo normal del lenguaje.

En ese estudio, se trabaja con textos de cincuenta enunciados, pero para los cálculos estadísticos consideraron solamente del enunciado 21 al número 50. Metodología distinta a la nuestra. No obstante lo mencionado, interesa resaltar que la media de diversidad léxica de los niños con DNL es de 0,64, muy próxima a la nuestra 0,65, aspecto que en cierta medida permite proyectar nuestros datos a poblaciones mayores.

Ahora bien, respecto de la variable tipo de institución (CEN-CINAI, kínder público y kínder privado), la distribución de los grupos es la siguiente: 40 preescolares de kínder público, 41 de kínder privado y 20 de CEN-CINAI. Respecto de la longitud de los textos producidos, fueron más elocuentes los niños de kínder público y privado que sus pares que asisten a CEN-CINAI.

La valoración cuantitativa de los textos es muy similar en los tres tipos de institución (0,65 y 0,66), las medias son casi las mismas. Sin embargo, al revisar los totales de palabras producidas en cada tipo de institución, podrían encontrarse otras valoraciones que permiten inferir la presencia de enunciados más largos en unas instituciones respecto de las otras: kínder público, 120,53 palabras; kínder privado, 111,26 palabras y CEN-CINAI, 103,95 palabras.

Aguado (1995) reflexiona en torno a la relación entre longitud promedio de los enunciados y riqueza léxica: enunciados con mayor número de palabras se verán reflejados en los índices de riqueza léxica. Por supuesto que al hacer valoraciones individuales se reflejará mejor esta relación, la cual lamentablemente se desdibuja al trabajar con poblaciones mayores.

En suma, con el objetivo de ubicar a los preescolares en pruebas individuales, se considerarán los siguientes rangos:

Superior: mayor a 0,75.

Normal: 0,56 a 0,74.

Inferior: menor a 0,55.

Por tanto y de conformidad con los parámetros, los niños que obtengan un rango inferior a 0,55 deberán recibir una atención educativa especial en la producción de discursos narrativos orales, a fin de promover su enriquecimiento léxico.

4. Conclusión

Los rangos promedio para valorar la riqueza léxica en los preescolares de cuatro, cinco y seis años son los mismos, según el comportamiento de la muestra analizada en este estudio.

La norma obtenida, presentada en tres rangos –superior, normal e inferior–, ofrece información básica para complementar una valoración global del desempeño lingüístico de los preescolares y, por tanto, focalizar la atención en aspectos específicos, tales como la riqueza léxica en la producción narrativa.

Este estudio aporta información básica para realizar una valoración objetiva del desarrollo normal del lenguaje de los preescolares costarricenses.

Notas

1. Medida que determina cuántas palabras diferentes aparecen en un texto de 100 unidades.
2. El programa Exégesis, donado por el Colegio de México, a la Escuela de Filosofía, Lingüística y Literatura, Universidad de Costa Rica, analiza la densidad de textos de 100 unidades. Este programa determina la riqueza léxica con base en los siguientes rangos:
 - a) superior: de 66,5 a 70,5
 - b) central: de 60,5 a 64,5
 - c) inferior: de 54,5 a 58,5
3. La categoría CEN-CINAI corresponde a las guarderías del Ministerio de Salud, centros de educación y nutrición que atienden niños de dos a los seis años.
4. El índice máximo de diversidad léxica es de 1.

Bibliografía

- Álvarez Durán, Barraza Parraguez, Heredia Tudela, Jonquera Fernández y Olea Ortega. 2006. *Desempeño semántico en niños con TEL: utilización del índice de diversidad léxica*. Seminario de investigación: Universidad de Chile.
- Aguado, Gerardo. 1995. *El desarrollo del lenguaje de 0 a 3 años: bases para un diseño curricular en la educación infantil*. Madrid: CEPE.
- Aguado, Gerardo. 1999. *Trastorno específico del lenguaje: retraso del lenguaje y disfasia*. Málaga: Aljibe.
- Bruner, Jerome. 1983. *El habla del niño*. Barcelona: Paidós.
- Chiat, Sula. 2001. *Mapping theories of developmental language impairment: premises, predictions and evidence*. En *Language and cognitive processes in developmental disorders*. University of Oxford: Dorothy Bishop.
- Dunn, Lloyd y Leota Dunn. 2006. *PPVT-III Peabody Test de Vocabulario en imágenes*. Madrid: TEA
- Echeverría, Max *et al.* 1997. "Estructura de un perfil de competencia lingüística". *Revista REALE*. 7: 39-55
- Galeote, Miguel. 2004. *Adquisición del lenguaje: problemas, investigación y perspectivas*. Madrid: Pirámide.
- Gallardo, José Ramón y José Luis Gallego. 2003. *Manual de logopedia escolar: un enfoque práctico*. Málaga: Aljibe.

- Herrera, María; María Mathiesen y Ana Pandolfi. 2000. "Variación en la competencia léxica del preescolar: algunos factores asociados" *Facultad de Filosofía y Humanidades de la Universidad Austral de Chile*. 35: 61-70. http://www.scielo.cl/scielo.php?pid=S0071-17132000003500004&script=sci_arttext. Consulta:14 de agosto de 2009.
- Murillo, Marielos. 1998. "Riqueza léxica en niños preescolares costarricenses". *Revista Educación*. 22 (1): 131-140.
- Murillo, Marielos y Víctor Sánchez. 2002. *Léxico básico de los niños preescolares costarricenses*. San José: Universidad de Costa Rica.
- Nippold, Marilyn A. 1998. *Later language development. The school-age and adolescent years*. Austin: PRO-ED, Inc.
- Pandolfi, Ana; María Herrera y María Mathiesen. (1994). "El lenguaje infantil: procedimientos analíticos para su descripción". *Revista Latinoamericana de Psicología*. 26 (1): 15-33.
- Pardo, Néstor. 2006. *Insuficiencia en vocabulario, terapia del lenguaje y educación*. <http://espanol.geocities.com/taouniv/vocabulario.html>. Consulta: 1 de marzo de 2009.
- Peralbo, Manuel *et al.*1998. *Desarrollo del lenguaje y cognición*. Madrid: Pirámide.
- Perelló, Jorge. 1977. *Trastornos del habla*.(3ed.). España: Científico-Médica.
- Puyuelo, Miguel y Jean-Adolphe Rondal. 2003. *Manual de desarrollo y alteraciones del lenguaje: aspectos evolutivos y patología en el niño y el adulto*. Barcelona: Masson.
- Puyuelo, Miguel; Jean-Adolphe Rondal y Elisabeth Wiig. 2000. *Evaluación del lenguaje*. Barcelona: Masson.
- Rondal, Jean-Adolphe. 1990. *La interacción adulto-niño y la construcción del lenguaje*. México: Trillas.
- Sánchez Corrales, Víctor y Marielos Murillo Rojas. 2006. *Disponibilidad léxica de los niños preescolares costarricenses*. San José: Universidad de Costa Rica.
- Serra, Miguel y Encarna Pérez. 2003. *Análisis del retraso del lenguaje(A-RE-L)*. Barcelona: Ariel.
- Serra, Miguel *et al.* 2000. *La adquisición del lenguaje*. Barcelona: Ariel.
- Siguan, Miguel; Rosa Colomina e Ignasi Vila. 1990. *Método para el estudio del lenguaje infantil*. Barcelona: Abril.

- Terré, Orlando. 2007. *Trastornos y alteraciones en el lenguaje infantil*. (3ed.). San José: Santa Paula.
- Triadó, Carmen y María Fornés. 1992. *La evaluación del lenguaje: una aproximación evolutiva*. Barcelona: Anthropos.
- Villegas, Francisco (2004). *Manual de logopedia: Evaluación e intervención de las dificultades fonológicas*. Madrid: Pirámide.