

Sexo y vocabulario: relación en escolares habaneros

Sex and vocabulary: Relationship between elementary students in Havana

Lidia Ester Cuba Vega¹

Para citar este artículo: Cuba, L.E. (2016). Sexo y vocabulario: relación en escolares habaneros. *Enunciación*, 21(1), 80-90.

Recibido: 11-diciembre-2015 /**Aprobado:** 27-marzo-2016

Resumen

Es común escuchar, aun a maestros y profesores experimentados en la enseñanza de lenguas, que las niñas tienen un vocabulario más amplio que los niños. Sin embargo, la mayoría de esas afirmaciones no están avaladas por resultados de investigaciones científicas, sino por el ejercicio de la práctica pedagógica cotidiana. Este trabajo tiene la finalidad de comprobar la veracidad de los criterios acerca de mayores índices de vocabulario en las niñas que en los niños, a partir de la aplicación de pruebas de riqueza léxica, y análisis estadísticos de la significación de los resultados obtenidos (las pruebas estadísticas paramétricas de diferencia de proporciones y la prueba *T de Student*). Para ello, a través de una metodología descriptivo-cuantitativa, propia de estudios léxico-estadísticos, y específica para los análisis de riqueza léxica, se trabajó con dos muestras de textos escritos de escolares de primaria de seis escuelas de La Habana, recogidas en dos momentos diferentes. El estudio presentado corresponde al proyecto de investigación universitario titulado "Proyecto marco sobre enseñanza del léxico", que desarrolla la Facultad de Español para No Hispanohablantes de la Universidad de La Habana, dirigido por la autora del actual artículo.

Los resultados permiten llegar a conclusiones válidas para la muestra analizada.

Palabras clave: discriminación de género, español como lengua extranjera, La Habana, vocabulario.

Abstract

It is common to hear, even from experienced language teachers and professor, that girls have a larger vocabulary than boys. However, most of these claims are not supported by scientific research results, but by the exercise of daily teaching practice. This work aims to verify the truthfulness of the criteria about vocabulary higher rates in girls than in boys, from the application of lexical richness tests and statistical analysis of the significance of the results (parametric difference statistical of proportions tests and Student's T tests). In order to achieve this, following a descriptive –quantitative, own lexical studies– statistical, and specifically for studies of lexical richness, we worked with two samples of written texts by primary school children from six schools in Havana, collected at two different times. The study presented corresponds to university research project entitled under Project teaching vocabulary, developed by the School of

¹ Doctora en Ciencias Lingüísticas de la Universidad de La Habana. Profesora titular. Decana de la Facultad de Español para No Hispanohablantes. Directora del Proyecto de Investigación "Proyecto marco sobre enseñanza del léxico", de la misma facultad en la Universidad de La Habana. Correo electrónico: lidiaacuba@fenhi.uh.cu.

Spanish for Spanish speakers, of the University of Havana, directed by the author of the current article. The results allow valid conclusions for the sample.

Keywords: Gender discrimination, Spanish as a foreign language, La Habana, vocabulary.

INTRODUCCIÓN

En la práctica pedagógica de maestros y profesores de español, se enumeran entre las principales dificultades de los estudiantes, no importa el grado escolar: la pobreza de vocabulario y la falta de un dominio adecuado del léxico de su lengua. Igualmente existen criterios, muchas veces canonizados, acerca de un léxico más *rico* en las niñas que en los niños².

Estos criterios tienen como base, mayoritariamente, apreciaciones dictadas por la experiencia en el aula –y no indagaciones científicas– que permitan establecer indicadores precisos acerca de qué palabras y cuántas debe poseer un estudiante, de acuerdo con su edad, grado escolar, tipo de texto que construyen, etc., y definir así su grado de *pobreza*. Por otro lado, faltan también estudios que posibiliten determinar qué se entiende por *dominio adecuado del léxico*, qué parámetros deben tenerse en cuenta para dictaminar buen o mal *dominio* léxico.

A pesar del tiempo transcurrido, aún subsiste la problemática expresada por Battaner sobre la planificación de la enseñanza del léxico, cuando expone:

[Aunque] el léxico está siendo centro de atención desde todos los puntos problemáticos del estudio de la lengua [sobre todo con] trabajos cualificados de vocabulario disponible (palabras aisladas) en escolares, recogidos por centros de interés, nos falta también un vocabulario recogido en textos producidos por los escolares [...], lo que daría nivel de uso en producción real de los escolares y representaría un conocimiento más exacto que el del vocabulario disponible. (Battaner, 1992).

No abundan las mediciones del vocabulario *que verdaderamente ha adquirido el estudiante*, a partir

de sus textos escritos, especialmente en Cuba, que permitan comprobar la efectividad de lo realizado, y determinar la adecuación o no de la enseñanza del léxico en los escolares.

En Cuba, los maestros y profesores de lengua por lo general no emplean instrumentos adecuados para medir el vocabulario que poseen los alumnos, ni para comprobar la adquisición de las nuevas palabras que trabajan en las clases, fundamentalmente por el desconocimiento que tienen de tales instrumentos. Esto se pudo corroborar en un estudio anterior (Cuba Vega, 2006) realizado sobre enseñanza del léxico para el cual se aplicó un cuestionario a una muestra de maestros y profesores de español como lengua materna y como segunda lengua³, en el que se indagaba sobre los instrumentos de que disponían para medir desarrollo y adquisición léxica.

Precisamente, uno de los instrumentos para comprobar y medir la adquisición del léxico lo constituyen las pruebas de riqueza léxica, que cumplen un doble propósito:

1. Actuar como diagnóstico científico.
2. Servir, a partir del diagnóstico, como una vía para la correcta planificación de la enseñanza del léxico en la escuela.

Las pruebas de riqueza léxica buscan el vocabulario activo, pero no en forma de nóminas o listado de palabras, sino en la producción de un texto oral o escrito. Su aplicación posibilita determinar el desarrollo cuantitativo del léxico y la influencia de variables particulares en él, y, por su doble capacidad de diagnosticar y servir como elemento para la planificación de su enseñanza, puede constituir un

2 Algunas investigaciones realizadas han comprobado los mejores resultados para uno u otro sexo indistintamente, como son los casos de Raúl Ávila (1989a), en México, y Ana M. Haché (1991), en República Dominicana.

3 Se entiende aquí por *segunda lengua*, la extranjera que se estudia en el país donde se habla; tal es el caso del español para estudiantes chinos en Cuba.

importante auxiliar para maestros y profesores en su trabajo sistemático.

En el presente trabajo, a partir de la aplicación de las *pruebas de riqueza léxica* y con el empleo de análisis estadísticos de la significación de los resultados obtenidos (las pruebas estadísticas paramétricas de diferencia de proporciones y la prueba *T de Student*), se comprueba la veracidad o no de los criterios acerca de mayores índices de vocabulario en las niñas que en los niños.

El objetivo de la investigación es caracterizar comparativamente la riqueza léxica de niños y niñas, a partir de la utilización de este instrumento de medición, en dos muestras de textos escritos por estudiantes de segundo, cuarto y sexto grados de primaria de La Habana (Cuba) recogidas en dos momentos diferentes, con un intervalo de doce años de diferencia⁴.

Antes de proceder a la explicación de la metodología empleada, se plantean algunos aspectos teóricos sobre riqueza léxica y los indicadores existentes para medirla.

Según expone Humberto López Morales (2010):

[en el] ámbito de la léxico-estadística el estudio de los índices de riqueza léxica ha sido un hito de mucha importancia. Las fórmulas, necesarias para su estudio, se han ido produciendo desde temprano: Giraud (1954), Ham (1979), López Morales (1984), Ávila (1986) y Tesitelová (1992)⁵ (p. 18).

4 Es necesario aclarar que el presente estudio constituye una parte de otro mayor en el que no solo se analizan comparativamente los índices de riqueza léxica de niños y niñas, es decir, la variable sexo, sino también grado escolar (segundo, cuarto y sexto) y tipo de escuela (interna y semiinterna; normal, de arte y deportiva; y de centro y periferia).

5 Al referirse a los datos de que se dispone sobre riqueza léxica, López Morales (2010) considera que son muy escasos todavía, y declara no conocer estudios dedicados a la enseñanza, sino a comprobar los altos índices de riqueza léxica en los medios de comunicación hispánicos de nuestros días. Sin embargo, para esa fecha, la autora de este artículo de investigación había realizado sus primeros estudios (2001) sobre riqueza léxica en escolares de primaria de La Habana, cuyos resultados aparecen publicados en el libro *La lengua en Cuba. Estudios* (2007), compilado por la Dra. Marlen Domínguez. Posteriormente, ha continuado esta línea de investigación en su tesis doctoral y en la aplicación de las pruebas de riqueza léxica a la enseñanza de lenguas extranjeras.

En ese estudio, López Morales (2010) hace un breve recuento de cómo los índices para medir riqueza léxica fueron variando de uno a otro autor, así como las fórmulas empleadas en cada caso para obtener dichos índices. Igualmente, presenta los índices de riqueza léxica obtenidos de diferentes investigaciones sobre el tema, de modo que sirvan para realizar comparaciones entre dichos índices, pues

[lo] que significan estos números y cómo pueden interpretarse necesitan, desde luego, de elementos de comparación. Siempre está a la mano la comparación entre los individuos del grupo, pero eso nos permite hacer comparaciones modestas. Lo ideal es disponer de cifras más generales (p. 23).

Hace énfasis en los estudios del Proyecto DIES-M (Difusión Internacional del Español en los Medios), coordinado por Raúl Ávila (2001), por los datos que presenta, válidos para compararlos con otras investigaciones sobre riqueza léxica.

Ninguno de los trabajos citados define *riqueza léxica* (RL) y en algunos casos el término aparece entrecomillado⁶, por lo que la autora del presente artículo, luego de una exhaustiva investigación al respecto, lo definió en el año 2001, y un lustro después lo redefinió como: grado de madurez y dominio léxico general de un hablante, medidos en la producción de sus textos, orales o escritos, por índices estadísticos como el porcentaje de vocablos, el intervalo de aparición de palabras de contenido nocional y el porcentaje de vocablos de contenido nocional⁷.

La propia autora (Cuba, 2007) concibió *madurez léxica* como la capacidad que permite al hablante manipular su léxico fundamental (básico + disponible⁸), para utilizar el mayor número posible

6 A juicio de la autora, ello responde al hecho de que los índices que se obtienen en el análisis de algunas muestras son tan bajos que hacen dudar si realmente se trata de riqueza en el sentido exacto de la palabra.

7 Este último índice es incorporado por la Dra. Cuba Vega en el 2001, pues ni Humberto López Morales ni el resto de los autores consultados y citados aquí, lo tienen en cuenta. Más adelante se hará referencia al porqué de su incorporación.

de vocablos, y palabras y vocablos de contenido nocional en un texto, lo que favorecerá una mejor expresión del contenido de su mensaje. El *grado de madurez léxica* es el desarrollo cuantitativo creciente del vocabulario, medido en la comparación de textos, ya sea en el sentido individual o grupal.

Por su parte, el término *dominio léxico general* se refiere a la manifestación del conocimiento que posee el hablante del léxico de su lengua, expresado en la cantidad y crecimiento del vocabulario empleado.

Otros términos que se han tenido en cuenta en la investigación son los de *léxico* y *vocabulario*, empleados indistintamente en otros contextos, pero diferenciados aquí, siguiendo el criterio de Antonio Quilis (1987), cuya distinción entre ambos conceptos resulta interesante:

El léxico del alumno, o su competencia léxica se manifiesta en la cantidad de palabras en potencia que reconoce o que puede utilizar en un momento dado; su vocabulario o su actuación léxica, es el reflejo del léxico en un enunciado dado (p. 254).

Como se aprecia, establece una diferencia entre léxico y vocabulario; cuando se refiere a léxico lo asocia al concepto de *competencia*, y vocabulario lo identifica con *actuación*, términos que, como se conoce, fueron introducidos en lingüística por Noam Chomsky.

METODOLOGÍA

Para la investigación se trabajó con una metodología descriptivo-cuantitativa, propia de estudios léxico-estadísticos, y específica para los estudios de

8 Estos conceptos son definidos por López Chávez de la siguiente forma: "a) *léxico básico*: abarca los vocablos más usuales de una comunidad y que se caracterizan por un alto grado de estabilidad – permanencia– que les permite aparecer con mucha frecuencia y en todo tipo de discursos, independientemente de la temática de que se trate; b) *léxico disponible*: es el conjunto de vocablos que utiliza una comunidad hablante como parte medular de su léxico patrimonial pero que únicamente aparecen en situaciones precisas, por lo que no son muy frecuentes pero sí muy conocidas. Son inestables debido a que es necesario que se toque el tema específico en el discurso para que aparezcan" (López Chávez, 1995, pp. 94-95).

riqueza léxica; se tomaron dos muestras de lengua escrita de escolares de primaria de La Habana, constituidas en dos momentos diferentes con un intervalo entre la primera (muestra A⁹) y la segunda (muestra B) de ocho años. La razón de este lapso entre una y otra obedeció al interés de analizar comparativamente dos muestras de lengua escrita de los escolares en dos momentos diferentes, a fin de comprobar la efectividad o no de la enseñanza del léxico en esas escuelas.

En cada periodo se seleccionaron seis escuelas primarias de La Habana, distribuidas en escuelas semiinternas y externas; de centro y periferia; normales, deportivas y de arte¹⁰.

Las escuelas, atendiendo a las variables independientes, quedaron distribuidas de la siguiente forma:

- *Centro/periferia*: 4 de centro y 2 de periferia.
- *Normal/deportiva/de arte*: 4 normales, 1 deportiva y 1 de arte.
- *Externa/semiinternas*: 5 semiinternas y 1 externa.

En cada escuela la muestra estuvo conformada por escolares de segundo¹¹, cuarto y sexto grados.

9 Esta primera muestra se recogió en 2006.

10 Para el diseño de la muestra B se partió del principio de lograr la mayor similitud posible con la muestra A, siempre que las particularidades del sistema educativo cubano lo permitieran. Sin embargo, por algunos cambios ocurridos en la organización de las escuelas primarias con las que se trabajó en la primera etapa, no fue posible mantener de manera idéntica la selección de las escuelas ni de todas las variables independientes; por eso hubo un cambio de escuela en uno de los municipios, pero manteniendo las mismas características de la anterior. Igualmente, como consecuencia de los cambios operados en los últimos años en el Sistema de Educación Primaria, dirigidos, entre otros objetivos, a lograr una mayor permanencia y atención de los niños en sus centros escolares, la variable escuela externa/semiinterna dejó ser operativa.

11 Ninguna de las investigaciones sobre riqueza léxica revisada por la autora incluye el segundo grado, pues en la mayoría de los casos sucede como en el estudio de Gloria E. Báez (1989), ya que los textos redactados por los alumnos de ese grado no alcanzan una extensión mínima de 60 palabras, por lo que se comienza a partir de tercer grado. Para este estudio se consideró más acertado tratar de seleccionar segundo, cuarto y sexto grados, para comprobar cómo evoluciona la riqueza léxica en los alumnos de este nivel de enseñanza, desde los primeros años, pasar por uno intermedio y concluir en el último año de la primaria, que es el sexto en Cuba. No se tuvo en cuenta el primer grado, debido a que en esa etapa se está iniciando el trabajo con lengua escrita, por lo que los niños de ese nivel aún no son capaces de redactar textos coherentes.

Teniendo en cuenta que se analiza riqueza léxica en el código escrito, se elaboraron como instrumento de evaluación, las pruebas. Se seleccionó como único modo discursivo para estas pruebas la narración, debido a que al controlar el modo discursivo, se garantiza la confiabilidad de la medición.

Ham Chande (1992) y Ana Ma. Haché (1991) afirman, de acuerdo con los resultados de sus trabajos, que la riqueza léxica varía según el modo discursivo empleado. De igual manera sucede con el tema del texto escrito, por eso se trabajó con dos tipos de narraciones: una con tema libre (texto libre) y otra con tema controlado: "Un día en la playa" (texto controlado).

Cada muestra incluye 102 alumnos, 6 por cada grado, 18 en total para cada escuela, excepto en la escuela de arte que solo contaba con alumnos de cuarto y sexto grados; por eso únicamente tiene 12 alumnos.

Del total de alumnos, 51 corresponden al sexo femenino y 51 al masculino.

En cada muestra se analizaron 202 textos, dos por cada estudiante: uno sobre tema libre y el otro sobre tema controlado.

En resumen, por la composición de ambas muestras se trabajó con un total de 204 alumnos y se analizaron 404 textos producidos por ellos.

Análisis de los textos

Los textos se analizaron en una versión electrónica y corregida ortográficamente. Estos fueron procesados manualmente.

Procedimiento empleado para el análisis de los textos:

1. Conteo del total de palabras de cada uno de los textos. Se trabajó con el siguiente concepto: *Palabra (P)*: unidad gráfica. Desde la perspectiva formal, se entiende por palabra "cada una de las formas gráficas entre espacios en blanco" (López Morales, 2010, p. 22).

2. Conteo de vocablos. Se partió del siguiente concepto de vocablo: *Vocablo (V)*: palabras diferentes en el texto. Se consideran desde el punto de vista formal y el de su significado.
3. Conteo de las palabras de contenido nocional. *Palabras de contenido nocional (PCN)*: aquellas que son portadoras de un contenido léxico: sustantivos, adjetivos, verbos y adverbios.
4. Conteo de los vocablos de contenido nocional. *Vocablos de contenido nocional (VCN)*: palabras de contenido nocional no repetidas.

Descripción de los índices de riqueza léxica

Para la realización de esta investigación, y teniendo en cuenta la posibilidad de aplicación de las pruebas para los docentes, se decidió trabajar con los índices aportados por Humberto López Morales (1991), por ser de más fácil aplicación para maestros y profesores, pues no requieren de una vasta preparación lingüística.

1. *Índice de proporción de vocablos* (porcentaje de vocablos calculado en relación con el total de palabras de cada texto) (*PorcV*). La fórmula para su obtención es $PorcV = V/N$. Con este índice se logra establecer qué porcentaje de las palabras del texto son diferentes, así como precisar en qué medida el estudiante repite o no las mismas palabras, con lo que mostrará su madurez léxica. A medida que aumenta el porcentaje, aumenta la riqueza léxica.
2. *Índice de intervalo de aparición de palabras de contenido nocional* (total de palabras del texto entre la cantidad total de unidades léxicas con contenido nocional). $IndPCN = N/PCN$. Aquí los resultados son valorados en sentido inversamente proporcional. El mejor índice es el que se acerca a 1; a medida que se aleje de ese valor, decrece la riqueza léxica.

3. *Índice de proporción de vocablos de contenido nocional* (porcentaje de VCN, calculado en relación con el total de palabras de cada texto). $Porc_{VCN} = VCN/N$. También es proporcional a la riqueza léxica.

Humberto López Morales (1991) solamente trabaja con las dos primeras medidas: porcentaje de vocablos e intervalo de aparición de palabras de contenido nocional. Sin embargo, en una prueba de cala realizada por la autora de la investigación, se pudo comprobar que existen estudiantes con un intervalo de aparición de PCN relativamente bajo, que repiten con frecuencia estas palabras, y otros con intervalos más altos que prácticamente no repiten las PCN¹². Ello llevó a la autora del presente estudio a tratar de establecer un indicador que tuviera en cuenta este último aspecto: determinar el número de palabras de contenido nocional no repetidas en los textos¹³.

ANÁLISIS DE LOS RESULTADOS

Para la descripción y análisis de los resultados se llevaron los datos obtenidos por el procesamiento manual de los textos, a hojas de cálculo en documentos de *Excel*.

12 Por ejemplo, un estudiante A obtiene una $Porc_v$ de 50% y un Ind_{PCN} 1,84 (considerado bueno para una muestra de estudiantes), alcanza solo un 31,4 % en $Porc_{VCN}$, lo que evidencia que una parte considerable de PCN se repiten a lo largo del texto, o sea, su producción de VCN es baja.

13 Precisamente, al revisar la bibliografía existente en Cuba, se pudo comprobar que paralelamente en el trabajo de Gloria E. Báez (1989) se había introducido este mismo índice para complementar los dos anteriores. Este indicador resulta importante, porque como bien explica Gloria E. Báez, los vocablos de contenido nocional son los que posibilitan obtener una mayor riqueza y variedad en la producción léxica, ya que los otros vocablos (nexos y actualizadores: conjunciones, preposiciones, artículos, pronombres, que cumplen funciones de conectores, relatores o determinantes) son adquiridos desde edades tempranas del desarrollo léxico y por ser paradigmas cerrados no tienden a incrementarse a través del aprendizaje escolar, y lo que se enseñará acerca de ellos, en la medida en que se asciende en grado, es el uso correcto; de ahí que no incidan esencialmente en la medición de la riqueza léxica.

Se realizaron las pruebas estadísticas paramétricas de *diferencia de proporciones* y la prueba *T de Student*, para comprobar la significación de las diferencias. Ambas permiten analizar si dos muestras difieren entre sí de manera significativa, pero la primera trabaja con proporciones o porcentajes y la segunda con intervalos o razones.

Por tales requisitos, la primera prueba se aplicó para comprobar la significación de las diferencias de los resultados obtenidos en porcentaje de vocablos y porcentaje de VCN, al finalizar el análisis interno de cada muestra y al realizar la comparación de las muestras entre sí.

La prueba *T de Student* se aplicó al análisis del intervalo de aparición de PCN. Para efectuarla se trabajó con el programa computacional *SPSS*, versión 11.5, luego de asentarlos datos obtenidos en el análisis de los textos de cada grado escolar, en un documento de *Excel*.

Porcentaje de vocablos

En la muestra A, el sexo femenino obtiene un porcentaje mayor en el empleo de vocablos (62,6 %) que el masculino (62,0 %), superior en +0,6 %. Los resultados de la prueba de diferencia de proporciones indican que esta diferencia no es significativa.

En la muestra B, el sexo masculino obtiene un porcentaje más alto que el femenino, superior en +0,3%; sin embargo, los resultados de la prueba de diferencia de proporciones indican que aunque el sexo masculino presenta mejores resultados, las diferencias no son significativas.

Si analizamos cada sexo por separado en los dos periodos para determinar el crecimiento o decrecimiento del porcentaje de vocablos en niñas y niños, observamos que la muestra A supera a la B (62,6 % frente a 61,9 %) en el sexo femenino, con un decrecimiento de -0,7 %. Por su parte, en el masculino, en la muestra B presenta mejor porcentaje con un incremento de +0,2, como se evidencia en la tabla 1.

Tabla 1. Porcentaje de vocablos por sexo.

Etapa	Sexo	
	Femenino	Masculino
Muestra A	62,6	62,0
Muestra B	61,9	62,2

Fuente: elaboración propia.

Según estos resultados, de una etapa a otra, los niños presentan mayor desarrollo léxico.

Al confrontar los resultados de uno y otro sexo, se aprecia que en la muestra A el femenino obtiene mejor porcentaje de vocablos, y en 2006, el masculino, con un ventaja de +0,25 % (figura 1).

Este resultado fue comprobado con la prueba de diferencia de proporciones al comparar los dos sexos en ambos periodos como muestras independientes. La diferencia obtenida no es significativa por lo que no favorece a ninguno de los dos sexos, en ninguno de los dos periodos.

Intervalo de aparición de palabras de contenido nocional (PCN)

En la muestra A, los mejores resultados van de uno a otro sexo, aunque el menor intervalo favorece

al sexo masculino. El sexo femenino obtiene un intervalo de aparición de palabras de contenido nocional de 1,92, mientras que el intervalo para el masculino es 1,91 PCN por texto, lo que significa que los niños tienen un intervalo mejor en -0,01 %. Sin embargo, esta ventaja no es significativa, lo que fue comprobado con la aplicación de la prueba *T de Student*.

En el segundo periodo analizado, el sexo femenino obtiene un intervalo por texto superior, con una diferencia de -0,03 respecto del masculino. El análisis general de los dos sexos en el periodo, muestra que las diferencias, en su mayoría favorables al femenino, no son significativas, de acuerdo con los resultados de la prueba *T de Student*.

El análisis de cada sexo por separado arroja que el uso de PCN por texto en la muestra A es superior en las niñas al de la muestra B, ya que el intervalo de aparición es menor (1,92 frente a

**Figura 1.** Vocablos por sexo.

Fuente: elaboración propia.

1,96), lo que significa que en el primer periodo la RL en la producción de PCN es mayor.

En el caso de los niños, el intervalo de aparición de PCN es menor en la muestra A que en la B (1,91 frente a 1,99).

En resumen, en la muestra A los mejores intervalos corresponden al sexo masculino y en la muestra B al femenino (tabla 2).

El análisis de los dos periodos revela mejores resultados para la muestra A en los dos sexos. (figura 2).

Los resultados según la prueba *T de Student* para la comparación de los dos sexos en uno y otro año, indican que las diferencias obtenidas son no significativas en los dos periodos.

Porcentaje de vocablos de contenido nocional (VCN)

En la muestra A, el sexo masculino obtiene mayor porcentaje de VCN que el femenino en +0,7; sin embargo, la aplicación de la prueba de proporciones muestra que las diferencias entre los dos sexos en el periodo no son significativas.

En la muestra B, el sexo femenino es superior al masculino en +0,5 %. De acuerdo con estos datos, la RL es mayor en las niñas que en los niños, pero los resultados de la prueba de diferencia de proporciones informan sobre la no significación de estas diferencias.

Tabla 2. Intervalo de aparición de PCN

Etapa	Sexo	
	Femenino	Masculino
Muestra A	1,92	1,91
Muestra B	1,96	1,99

Fuente: elaboración propia.

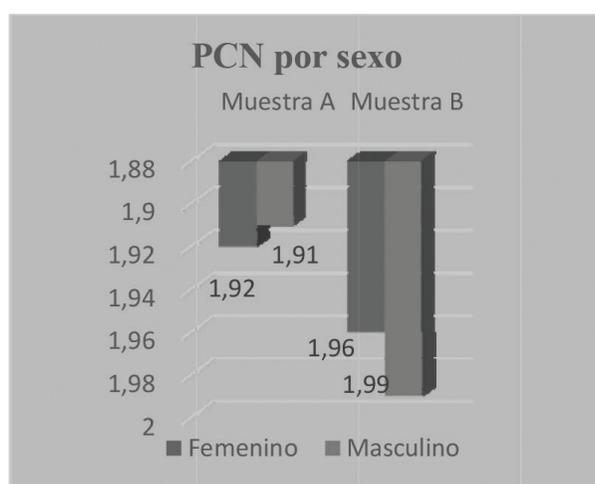


Figura 2. PCN por sexo.

Fuente: elaboración propia.

El análisis de cada sexo por separado indica que la utilización de VCN por texto en las niñas, en la muestra A, es superior en +0,3 %, en relación con la B (42,9 % y 42,6 %, respectivamente).

En los varones, el porcentaje de empleo de VCN es mayor en la primera etapa, como ocurrió con el femenino. La diferencia favorable a la muestra A es de +1,5 % (43,6 % frente 42,1 %) (tabla 3).

La confrontación de los dos años presenta con mejores resultados a la muestra A, lo que significa que los escolares en la primera etapa poseen una mayor madurez léxica, es decir, presentan un índice superior de RL (figura 3), pero las diferencias entre los dos años no son significativas, según se recoge en los resultados de la prueba estadística aplicada.

Los resultados de la variable sexo en los tres indicadores son heterogéneos, pues unas veces favorecen al femenino y otras al masculino.

Ante la heterogeneidad de los datos en esta variable independiente, se analizaron comparativamente los resultados de cada sexo en los tres indicadores de los dos periodos, con el fin de comprobar si esta variabilidad afectaba directamente la significancia de dicha variable.

Aunque en el porcentaje de vocablos el sexo femenino supera al masculino en 0,01 %; en el intervalo de PCN también las niñas aventajan a los varones en 0,01 y en el porcentaje de VCN el sexo masculino obtiene +0,25 % respecto del femenino, ninguna de las tres variables dependientes, de acuerdo con las pruebas estadísticas aplicadas, presenta diferencias significativas.

Los resultados favorecedores a niñas o niños indistintamente se presentan de manera similar en las tres variables independientes restantes: grado escolar, tipo de texto (controlado o libre) y tipo de escuela, en todas las variables dependientes

Tabla 3. Porcentaje de VCN

Etapa	Sexo	
	Femenino	Masculino
Muestra A	42,9	43,6
Muestra B	42,6	42,1

Fuente: elaboración propia.

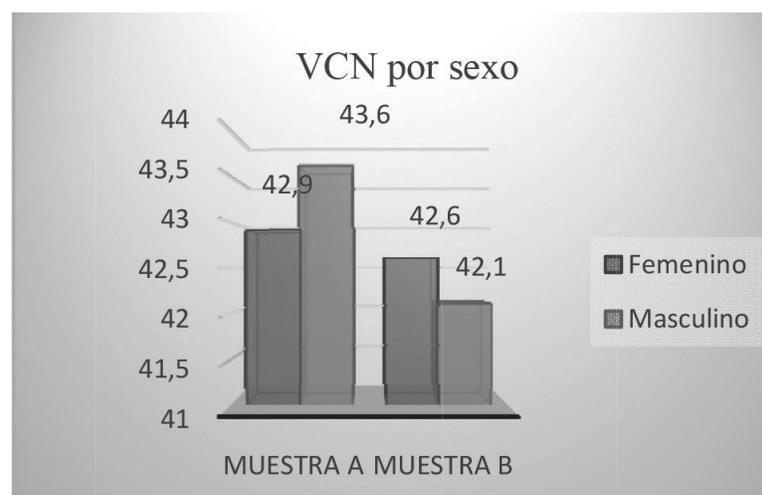


Figura 3. VCN por sexo

Fuente: elaboración propia.

(índice de proporción de vocablos, índice de intervalo de aparición de palabras de contenido nocional e índice de proporción de vocablos de contenido nocional).

Esta situación es contraria a la observada en otras investigaciones que miden densidad o riqueza léxicas en textos escritos por escolares de primaria, en las que se comprueba mejores resultados de uno u otro sexo. Tal es el caso de "Las palabras de los niños: tercer grado de primaria", investigación efectuada por Raúl Ávila (1989a) con niños mexicanos, donde el número de palabras diferentes por texto y el coeficiente promedio son mayores en las niñas (Ávila, 1989b, pp. 31-35), así como el caso de la investigación titulada "Aportes de las pruebas de riqueza léxica a la enseñanza de la lengua materna", de Ana M. Haché de Yunén (1991), en la que trabaja con textos escritos por niños dominicanos, ya que tanto en el porcentaje de vocablos como en el intervalo de aparición de PCN, los varones superan a las niñas (pp. 54-56).

Más allá de la comparación de los indicadores por sexo, vale anotar que los resultados obtenidos en los dos primeros índices que miden RL ($Porc_v$ e Ind_{PCN}), están por encima de los logrados en las investigaciones realizadas en México por Gloria Báez (1989) y en República Dominicana por Ana Margarita Haché (1991).

Partiendo de los tres rangos indicativos establecidos por Raúl Ávila (1989a, 1989b) para la riqueza léxica, y teniendo en cuenta que se trata de escolares de primaria, los porcentajes de proporción de vocablos obtenidos en los dos periodos analizados, pueden considerarse altos para la muestra, pues corresponderían al nivel central (entre 60,5 % y 64,5 %).

Si se tiene en cuenta que el rango superior de densidad corresponde generalmente a personas con estudios universitarios y de amplia información, el hecho de que los resultados en el actual estudio se ubiquen en el rango central es altamente satisfactorio, ya que el nivel de instrucción de los escolares de la presente muestra corresponde a la enseñanza primaria.

Por otra parte, comparando con porcentajes obtenidos por el mismo investigador en programas de radio y televisión, y en el ensayo de Octavio Paz, *Tiempo nublado*, que constituyen muestras de lengua para un estrato de habla culta, en los que la densidad léxica oscila entre 67,7 % y 69,5 %, así como el índice que establece para el habla popular (lengua hablada) que es de 56,5 %, se ratifica la idea de que los porcentajes de los estudiantes de la muestra son altos para su nivel y madurez.

CONCLUSIONES

Los resultados del análisis de los textos escritos por niños y niñas habaneros en los dos periodos comparados también permiten afirmar que, al menos en esta muestra, el tópico de que las niñas poseen un vocabulario más amplio que los niños y mejor riqueza léxica, no es más que eso, un tópico.

Los resultados indistintos en otras muestras con niños y niñas de otros países llevan a considerar la validez de estos resultados en cuanto a que no hay nada preestablecido en relación con la riqueza léxica y el sexo.

Igualmente, los resultados obtenidos permiten calificar de altos los índices de riqueza léxica, aun en el segundo grado, comparados con los obtenidos en otros estudios, tanto en las niñas como en los niños.

Las pruebas estadísticas aplicadas posibilitaron valorar la significación de las diferencias obtenidas, aspecto novedoso para este tipo de estudio, al menos en los que conoce la autora de la investigación.

El trabajo efectuado permite validar la pertinencia de las pruebas de riqueza léxica para medir el vocabulario de los escolares en textos escritos, ya sea atendiendo a la variable independiente sexo, como a las tres restantes. Igualmente, se comprueba su utilidad para medir el desarrollo léxico y la efectividad de la enseñanza, al ser posible su aplicación a una misma muestra en momentos diferentes.

Por último, vale apuntar a que con esta investigación se da respuesta no solo a una necesidad pedagógica, sino a la preocupación expresada por Humberto

López Morales (2010) en el artículo mencionado cuando lamentaba que no se hubiera seguido con esta línea de investigación en Puerto Rico. En tal sentido refería:

Aquel episodio terminó allí, pero el interés por el tema no desapareció. Debo y puedo decir que los resultados alcanzados entonces en cuanto a riqueza léxica fueron muy favorables. Vale la pena rescatar aquella experiencia, actualizar esos materiales y proseguir con la tarea (p. 27).

RECONOCIMIENTOS

El estudio presentado en este trabajo corresponde al proyecto de investigación universitario titulado "Proyecto marco sobre enseñanza del léxico", que desarrolla la Facultad de Español para No Hispanohablantes de la Universidad de La Habana, dirigido por la autora del actual artículo.

REFERENCIAS

- Ávila Cell, R. (1989a). Las palabras de los niños: tercer grado de primaria. En: *Estudios sobre español de América y Lingüística Afroamericana* (pp. 19-39). 45 Congreso Internacional de Americanistas. Bogotá: Publicaciones del Instituto Caro y Cuervo.
- Ávila Cell, R. (1989b). Léxico infantil de México: palabras, tipos, vocablos. *Revista Varona*, 21, 118-132.
- Báez Pinal, G.E. (1989). Índices de riqueza léxica en escolares de primaria del D.F. Material sin publicar. México.
- Battaner, P. (1992). *La investigación en enseñanza del español como lengua materna: vocabulario y léxico*. Congreso de Sevilla. Recuperado de: http://cvc.cervantes.es/obref/congresos/sevilla/unidad/ponenc_battaner.htm
- Cuba Vega, L.E. (2001). Índices de riqueza léxica en escolares de primaria de ciudad de La Habana. En: *Actas del VII Simposio Internacional de Comunicación Social* (pp. 263-265). Enero. Santiago de Cuba: Centro de Lingüística Aplicada. Tomo II.
- Cuba Vega, L.E. (2004). "Las pruebas de riqueza léxica y su aplicación en Español como Lengua Extranjera (ELE)". En Actas del VIII Congreso Sociedad Española de Didáctica de la Lengua y la Literatura, 5-9 de diciembre de 2004. La Habana. Recuperado de http://sedll.org/es/admin/uploads/congresos/8/act/250/cuba_vega_lidia.pdf
- Cuba Vega, L. (2006). *El tratamiento del léxico en las clases de español. Apuntes para una propuesta metodológica*. Sin publicar. La Habana: Universidad de La Habana.
- Cuba Vega, L. (2007). Riqueza léxica en escolares de primaria de Ciudad de La Habana. En: M. Domínguez Hernández (ed.) (2007). *La lengua en Cuba. Estudios* (pp. 283-308). Santiago de Compostela: Universidad de Santiago de Compostela.
- Haché de Yunén, A.M. (1991). Aportes de las pruebas de riqueza léxica a la enseñanza de la lengua materna. En: H. López Morales. *La enseñanza del español como lengua materna* (pp. 49-60). Actas del II Simposio Internacional sobre Aportes de la lingüística a la enseñanza del español como lengua materna. Río Piedras: Editorial de la Universidad de Puerto Rico.
- Ham Chande, R. (1992). Del 1 al 100 en lexicografía. En: L. Vallés Calaña (pp. 45-80). La Habana: Universidad de La Habana.
- López Chávez, J. (1995). Léxico y enseñanza de la lengua materna. En: M. Arjona Iglesias et al. (eds.). *Actas del I Encuentro sobre Problemas de la Enseñanza del Español en México, Zacatecas* (pp. 93-105). Mayo. México: Universidad Autónoma de México.
- López Morales, H. (1991). La enseñanza del español como lengua materna. En: *Actas del II Simposio Internacional sobre Aportes de la Lingüística a la Enseñanza del Español como Lengua Materna*. Río Piedras: Editorial de la Universidad de Puerto Rico.
- López Morales, H. (2010). Los índices de "riqueza léxica" y la enseñanza de lenguas. En: *Actas del XXI Congreso Internacional de la ASELE' 2011* (pp. 15-28). Salamanca. Recuperado de: http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/asele_xxi.htm
- Quilis, A. (1987). La enseñanza de la lengua materna. *Boletín de la Academia Puertorriqueña de la Lengua Española*, 1, 214-261.

