

Densidad léxica y modos de organización del discurso: evaluación de la competencia semántica en el inicio de los estudios secundarios¹

Darío Daniel Delicia

María Victoria Fernández

Eliana Gigena

Área de Lingüística Aplicada, Facultad de Lenguas, UNC

ala@fl.unc.edu.ar

Resumen

La psicolingüística ha demostrado la importancia del acceso competente al significado léxico, tanto para los procesos de comprensión como para los de producción del discurso. Es decir, el dominio del vocabulario se vuelve un factor decisivo para la lectoescritura y, por ello, en el plano didáctico, se torna obligatorio promover su desarrollo, a propósito de expandir las posibilidades lingüístico-comunicativas (y cognitivas) de nuestros alumnos. En la dimensión investigativa, la sistematización de datos relacionados con la adquisición y el desarrollo del léxico proporcionará un diagnóstico minucioso de la situación, el cual, luego, podrá proyectarse a la tarea de enseñanza. En ese marco, en el presente trabajo evaluamos la riqueza léxica en composiciones escritas (narrativas, descriptivas y argumentativas) pertenecientes a niños que aspiran a ingresar a la educación media. Examinamos tales parámetros en relación con las variables sexo y modo de organización del discurso. Así pues, este estudio considera la hipótesis de que existe covariación entre estas variables y la riqueza (o densidad) léxica, noción que entendemos como la consistencia semántica que posee un texto, según la frecuencia con que aparecen en él palabras con contenido conceptual (López Morales 1984, Halliday 1987). Desde el punto de vista metodológico, en el marco de un diseño de investigación descriptivo y desde un enfoque cuantitativo, abordamos un corpus compuesto por 120 fragmentos de textos escritos por los 40 estudiantes que conforman la muestra.

Abstract

Psycholinguistics has demonstrated the importance of having a competent access to the lexical meaning, both for understanding processes and for speech production. That is, the control over vocabulary becomes a critical factor in literacy and therefore, in the educational level, promoting their development becomes compulsory in order to expand the linguistic-communicative (and cognitive) opportunities of our students. In the research dimension, the systematization of data related to the acquisition and development of the lexicon will provide a thorough diagnosis of the situation, which then can be projected to teaching. Within this framework, in this paper we assess the lexical richness in written compositions (narrative, descriptive and argumentative) that belong to children who aspire to join the Middle School. We examine these parameters in relation to sex and mode of discourse organization. Thus, this study considers the hypothesis that there is some covariation between these variables and the lexical richness (or density), a notion understood as the semantic consistency of a text according to the frequency with which conceptual words appear in it (López Morales 1984, Halliday 1987). From the methodological point of view, within the framework of a

¹ Esta comunicación adscribe al trabajo realizado en el marco de la investigación *Gramática y discurso en la educación lingüística cordobesa: índices de riqueza léxica y madurez sintáctica en la producción escrita adolescente* (RS 162/12 - RR 2093/12 - Secretaría de Ciencia y Tecnología, UNC, Argentina).

descriptive research and from a quantitative approach, we address a corpus of 120 fragments of texts written by the 40 students who make up the exhibit.

1. Introducción

En la actualidad, las prácticas de lectura y de escritura se han tornado considerablemente complejas y demandan de los hablantes un dominio cada vez más competente del sistema de la lengua. Sin embargo, en el plano pedagógico, la praxis docente revela que niños y jóvenes atraviesan serias dificultades respecto de su desempeño lingüístico-comunicativo.

Uno de los núcleos de esta problemática se encuentra en la competencia léxica de que disponen los estudiantes para concederle densidad semántica a un texto, a partir de las palabras que la situación comunicativa “establece” que se utilicen. Desde la dimensión investigativa, la sistematización de datos relacionados con la adquisición y el desarrollo del léxico proporcionará un diagnóstico minucioso de la situación, el cual, luego, podrá proyectarse a la tarea de enseñanza.

En ese marco, en el presente estudio evaluamos los índices de riqueza léxica (porcentaje de vocablos nocionales y promedio de cada una de las categorías léxicas) en composiciones escritas pertenecientes a adolescentes, varones y mujeres, que van a comenzar sus estudios secundarios en una escuela pública de la ciudad de Córdoba. En concreto, nos proponemos examinar tales parámetros vinculados a las variables sexo y modo de organización del discurso, a saber: narrativo, descriptivo y argumentativo.

Así pues, este estudio considera la hipótesis de que existe covariación entre la riqueza léxica y las variables antes mencionadas. La riqueza léxica, también denominada “densidad léxica” por Halliday (1987), se entiende aquí como la consistencia semántica que posee un texto según el número de palabras nocionales que lo componen (López Morales 1984). Básicamente, es una medida estadística que señala la frecuencia con que aparecen, dentro de un discurso, palabras conceptuales en relación con palabras funcionales.

Son numerosos los trabajos desarrollados dentro del mundo hispánico que han abordado el estudio de la riqueza léxica en niños y adolescentes (Hayé de Yunén 1991, Valencia Espinoza *et al* 1992, Torres González 2003, Quesada Murillo *et al* 2001, Reyes Díaz 2007, Cuba Vega 2007, Murillo Rojas 2009, Delicia y Fernández 2011, entre muchos otros). En la mayoría de estas investigaciones, las variables consideradas fueron, principalmente, el sexo y la edad; asimismo, el origen geográfico, el nivel socioeconómico, el tipo de centro educativo y el modo discursivo aparecen como factores que pueden incidir en el desarrollo y la competencia léxico-semántica.

En relación con las variables que manipularemos en este estudio, el sexo de los participantes parece comportarse como un factor heterogéneo respecto de la densidad léxica; pero, en general, la mayoría de las investigaciones concluye en que no se asocia a los índices de riqueza léxica. En cuanto al tipo discursivo, se ha reportado que la densidad semántica de un texto varía de acuerdo con el modo discursivo en que el hablante produce un texto.

De acuerdo con el propósito planteado más arriba, los interrogantes que guiaron esta investigación son los siguientes. ¿Cuáles son los índices de riqueza léxica alcanzados por adolescentes cordobeses que ingresan al nivel secundario de escuelas públicas de la ciudad de Córdoba? ¿Hay diferencias significativas entre estos índices según el sexo de los participantes y el modo discursivo producido en la escritura? ¿Cuál es la clase de palabra con contenido conceptual de uso frecuente, de acuerdo con esas variables?

2. Marco teórico

La base conceptual de esta investigación coloca la pieza léxica en un lugar central dentro de la comprensión y producción de discursos, ya que constituye “la puerta de acceso” al significado de los textos. Así pues, la relación entre la noción de competencia léxica y procesos de interpretación/formulación de enunciados inscribe este estudio en el marco teórico de la psicolingüística.

De acuerdo con Giammatteo y Albano (2009: 26), la noción de competencia léxica puede describirse integrando dos enfoques: el intralingüístico y el extralingüístico. Desde la perspectiva intralingüística, se estima que el hablante reconoce el léxico a partir de interrelacionar información proveniente de los diferentes niveles de la lengua. Por su parte, la perspectiva extralingüística supone que para el conocimiento de las palabras “se necesita dominar la dimensión que estas pueden adquirir en sus diferentes contextos de aparición” (Giammatteo y Albano 2009: 28). Es decir, existen usos específicos y exclusivos de significados léxicos, cuya determinación está dada por coordenadas discursivas, pragmáticas, contextuales (socioculturales), entre otras.

La competencia léxica, sistémica y contextual puede evaluarse a través del establecimiento de los índices de riqueza léxica, parámetros propuestos por H. López Morales en 1984. Siguiendo a este autor, la noción de riqueza léxica se asume como el caudal de vocabulario conceptual o nocional que se actualiza conforme al tema del discurso y a las circunstancias comunicativas.

Según Cuba Vega (2007: 290), la riqueza léxica es “el grado de madurez y dominio léxico general [...] de un hablante, medido en la producción de sus textos orales o escritos por índices estadísticos”. Para establecer estos índices, se vinculan, dentro de un discurso dado, el número de palabras léxicas (piezas léxicas) y el número de palabras funcionales.

El porcentaje de vocablos con contenido léxico o conceptual constituye uno de los índices por medio de los cuales se puede establecer la riqueza léxica de un hablante. Otro parámetro es el del promedio de cada una de las categorías nocionales (sustantivos, adjetivos, verbos y adverbios). Torres González (2003) ha analizado estos indicadores en el desarrollo léxico de estudiantes de primaria y secundaria, y, en el presente trabajo, se tendrán en cuenta como variable lingüística de estudio.

Operacionalmente, la definición de riqueza léxica articula una serie de nociones en torno del concepto de palabra. La palabra, formalmente, es el elemento que se reconoce por estar separado por blancos en la escritura y, gramaticalmente, es la unidad constituida por una o por varias palabras. En este caso, se usa el concepto de pieza léxica o unidad léxica (RAE 2009: 11).

Desde la perspectiva del significado y a los fines del análisis, las palabras se clasifican en léxicas y funcionales. Las palabras léxicas son aquellas que llevan información conceptual, nocional o descriptiva; las palabras funcionales, por su parte, son las que “aportan significaciones abstractas [...], tales como la referencia, la pluralidad, las marcas sintácticas de función” (RAE 2009: 43).

Respecto de los modos de organización del discurso, hemos adoptado el concepto de Charaudeau (2009: 74), según quien, designan “los procedimientos que consisten en utilizar determinadas categorías de la lengua para ordenarlas en función de las finalidades discursivas del acto de comunicación”.

Siguiendo al mismo autor, los modos narrativo, descriptivo y argumentativo, poseen una función de base (relacionada con la intención del enunciador) y un principio organizacional tanto del mundo como de la escena de enunciativa.

Así, la función de base del modo narrativo es construir temporalmente una sucesión de hechos con el objeto de relatar una historia; la del descriptivo identificar y calificar a los seres; y la del modo argumentativo construir explicaciones sobre aseveraciones acerca del mundo, a partir de la demostración y la persuasión (Charaudeau 2009).

Es importante destacar que, cuando al interior del corpus se observó la realización de un texto en tal o cual modo de organización del discurso, tomamos en cuenta esta noción homologada con el concepto de secuencia textual dominante (Adam 1991). Esto significa que no desestimamos la heterogeneidad estructural de los textos analizados.

3. Metodología

El diseño de esta investigación alcanza un nivel no experimental, transeccional, descriptivo (Hernández Sampieri *et al* 2003: 273). El análisis de los datos se efectuó desde un enfoque cuantitativo.

Una vez establecidos los índices de riqueza léxica (IRL), se procedió a calcular, sobre esos parámetros, medias aritméticas (\bar{X}), porcentajes (%) y desviaciones típicas (S). Los valores obtenidos fueron luego comparados en función de las variables sexo y modo de organización del discurso.

Con el propósito de evaluar la significatividad estadística de la comparación efectuada, se realizaron las pruebas paramétricas del *t* de Student y de ANOVA, según el caso. Todas estas mediciones se llevaron a cabo por medio del programa informático *IBM SPSS*.

Los participantes fueron alumnos (varones y mujeres) que se preparan para iniciar sus estudios secundarios en una escuela pública nacional en la ciudad de Córdoba. A partir de esta muestra, conformamos un corpus de 120 textos escritos, correspondientes a cada una de las modalidades discursivas objeto de interés. Este material se recolectó por medio de la aplicación de tres pruebas de producción escrita (siempre sobre los mismos sujetos y en momentos diferentes) y sobre él se efectuó un análisis por segmentos discursivos cuya extensión alcanzó las 100 piezas léxicas.²

4. Resultados obtenidos

4. 1. Índices de riqueza léxica según la variable sexo. Porcentaje de vocablos nocionales

SEXO	Media	N	Desv. típ.
MASCULINO	64,44	60	6,79
FEMENINO	62,96	60	6,20

² Los comandos didácticos presentados a los estudiantes, tanto en forma oral como escrita, se describen sintéticamente a continuación:

- *Modo de organización narrativo*: Contame... ¿qué hiciste este verano durante las vacaciones?
- *Modo de organización descriptivo*: Vamos a recorrer el colegio, ¿cómo es tu colegio?
- *Modo de organización argumentativo*: ¿Qué opinás acerca de tener que venir al colegio todos los días?

SEXO	Media	N	Desv. típ.
MASCULINO	64,44	60	6,79
FEMENINO	62,96	60	6,20
Total	63,70	120	6,52

Tabla N° 1: índices de riqueza léxica según la variable sexo

El reporte de datos que se muestra en la tabla N° 1 permite observar que la variable sexo marca diferencias en cuanto a la densidad léxica. Se advierte que el grupo de varones obtiene un IRL de 64.44% y que el femenino llega a un 62.96%. Esto señalaría que el grupo masculino puede emplear más palabras con contenido léxico o nocional que el femenino, considerando las modalidades discursivas en su conjunto.

Con el objeto de verificar si las diferencias entre los índices calculados son estadísticamente significativos, se realizó una prueba paramétrica *t* de student en un nivel alfa (α) 0.05. Por medio de esta prueba, se pudo corroborar que las diferencias entre los IRL de masculinos y femeninos no son significativas, dado que el valor que asume la probabilidad (*p*) de F es de 0.215, esto es, superior al nivel alfa.

4. 2. Índices de riqueza léxica según la variable modo de organización del discurso. Porcentaje de vocablos nocionales

MODO DE ORGANIZACIÓN DEL DISCURSO	Media	N	Desv. típ.
NARRATIVO	68,63	40	4,25
DESCRIPTIVO	59,36	40	6,75
ARGUMENTATIVO	63,10	40	4,64
Total	63,70	120	6,52

Tabla N° 2: índices de riqueza léxica según la variable modo de organización del discurso

Según la información que se expone en la tabla N.º 2, la riqueza léxica para producir textos narrativos, descriptivos y argumentativos presenta también variaciones. En un orden descendente, puede observarse que el discurso narrativo, en el cual los participantes obtienen un IRL de 68.63%, favorece una mayor consistencia de vocablos léxicos. Le sigue el modo argumentativo con un 63.10%. Finalmente, la afluencia de palabras conceptuales es menor en el texto descriptivo donde el parámetro alcanza un nivel de 59,36%.

Aquí también se verificó si las diferencias halladas son relevantes desde el punto de vista estadístico. En efecto, de acuerdo con el estudio de ANOVA, cuyo nivel alfa (α) de significación se estableció en 0.05, se pudo confirmar que son significativos los contrastes entre los IRL establecidos para cada modo de organización del discurso, pues el valor que asume la probabilidad (*p*) de F es inferior al nivel alfa: 0.000.

4. 3. Promedio de categorías con contenido conceptual

Nos propusimos en este estudio verificar los promedios de uso de cada clase de vocablo conceptual. Así, calculamos la media (\bar{X}) de sustantivos, adjetivos, verbos y adverbios en relación con las dos variables secundarias consideradas. Los valores asumidos por estas categorías se destacan en la tabla N° 3.

Primero, llevamos a cabo este conteo a fin de observar en qué clase de palabra con contenido descriptivo se configuraría, en particular, la diferencia entre el conjunto de varones y de mujeres. En efecto, si se desglosa el porcentaje general de palabras conceptuales, alguna de las categorías que lo conforman podría señalar diferencias específicas entre los dos grupos.

En este caso, también corroboramos tales diferencias mediante el estudio de ANOVA en un alfa (α) de 0.05 y, en ningún caso, se obtuvieron valores inferiores a este nivel. Por lo tanto, las diferencias entre varones y mujeres respecto de los promedios de cada categoría léxica no son estadísticamente significativas.

En segundo término, efectuamos el conteo de cada categoría gramatical descriptiva, tomando en consideración los modos de organización del discurso. Como se observa en la tabla N° 3, las diferencias son más marcadas y el comportamiento de las léxias es bastante heterogéneo según el discurso de que se trate.

La narración parece favorecer el uso de verbos (16.38%). El sustantivo es la clase de palabra más usada en la descripción y en la argumentación, donde los valores ascienden a 18.52% y 14.78%, respectivamente. Nótese el bajo porcentaje de adjetivos en la descripción, categoría que típicamente se señala como uno de los rasgos que define a este modo del discurso.

El estudio de ANOVA realizado para comparar las diferencias entre estos promedios reportó, en todos los casos, un valor inferior al nivel alfa (α), establecido en 0.05. Así, tales diferencias son significativas desde el punto de vista estadístico.

SEXO	SUSTANTIVOS	ADJETIVOS	VERBOS	ADVERBIOS
MASCULINO	5,65	1,95	4,21	1,01
FEMENINO	4,97	1,57	4,49	1,22
MODO DISCURSIVO	SUSTANTIVOS	ADJETIVOS	VERBOS	ADVERBIOS
NARRATIVO	14,53	5,57	16,38	5,17
DESCRIPTIVO	18,52	4,48	8,90	1,30
ARGUMENTATIVO	14,78	5,82	13,90	3,60

Tabla N° 3: promedio de categorías léxicas según la variable sexo y modo de organización del discurso

5. Conclusiones

La evidencia empírica recogida a través de esta investigación permite formular las siguientes afirmaciones (no generalizables):

- La variable sexo no incide sobre la riqueza léxica de los adolescentes cordobeses aspirantes a ingresar al 1° año de la escuela media. Aun cuando se hallaron diferencias entre ambos grupos, se comprobó que no son estadísticamente relevantes. Los resultados obtenidos revelan paridad con los alcanzados por la mayoría de los estudios previos realizados sobre este tema (*cf.* Introducción).

- En cuanto a la densidad del vocabulario presente en textos organizados bajo la modalidad narrativa, descriptiva y argumentativa, la comparación entre los porcentajes calculados alcanza valores más elevados y, efectivamente, las diferencias entre ellos poseen significación estadística. Se constata una marcada asociación entre la variable lingüística estudiada y el modo de organización del discurso. Según este dato, existe coincidencia parcial con los resultados obtenidos por otras investigaciones que demostraron que el modo discursivo narrativo posee mayor densidad de vocabulario y que, correspondientemente, la afluencia léxica es menor en el modo argumentativo y en el descriptivo (*cf.* Valencia Espinoza et al. 1992).

- El análisis del promedio de cada una de las categorías léxicas no arrojó ningún caso particular de vocablo que estableciera diferencias importantes, entre varones y mujeres, en cuanto a la competencia para usar sustantivos, adjetivos, verbos y adverbios.

- Finalmente, en relación con los modos del discurso, los verbos son las lexías más frecuentes en la narración, frente a otras clases de palabras. Por su parte, los sustantivos predominan en la argumentación y en la descripción. Las diferencias halladas fueron comprobadas mediante prueba estadística y se reportaron significativas.

Para terminar, es importante destacar una de las principales implicancias que, en opinión de los autores, se desprende de la evaluación de la riqueza léxica, sobre todo para la enseñanza de la lengua materna. La medición de léxico activo, de aquel que los estudiantes usan en una situación concreta de habla, dice en parte del vocabulario que ‘no usan’; en consecuencia, el docente en su tarea áulica deberá promover el incremento del caudal léxico, a propósito de impactar directamente en el mejoramiento de las prácticas de interpretación y producción discursivas de los estudiantes.

6. Bibliografía

Adam, Jean Michel. *Les textes: types et prototypes*. París: Nathan, 1991.

Charaudeau, Patrick. *Linguagem e discurso. Modos de organização*. São Paulo: Contexto, 2009.

Cuba Vega, Lidia Esther. “Índices de riqueza léxica en escolares de primaria de ciudad de la Habana”. En Domínguez Hernández, Marlen. A. (ed.), *La lengua en Cuba. Estudios*. Santiago de Compostela: USC, 2007.

Delicia, D. y María Victoria Fernández. “Medición de la riqueza léxica en el discurso descriptivo de niños aspirantes a ingresar al nivel de enseñanza secundaria”. Ponencia leída en el *II Congreso de la Lengua y la Literatura en la escuela secundaria: “Rumbos y cartografías: las lenguas y las literaturas en las aulas hoy”*. La Plata: Colegio Nacional Rafael Hernández, UNLP, 2011.

Giammatteo, Mabel e Hilda Albano. *Lengua. Léxico, gramática y texto. Un enfoque para su enseñanza basado en estrategias múltiples*. Buenos Aires: Editorial Biblos, 2009.

Halliday, Michael Alexander Kirkwood. "Spoken and written modes of meaning". En Horowitz, Rosalind y Samuels S. Jay (eds.), *Comprehending Oral and Written Language*. New York: Academic Press, 1987, pp. 55-82.

Hayé de Yunén, Margarita. "Aportes de las pruebas de riqueza léxica a la enseñanza de la lengua materna". En López Morales, Humberto (ed.), *La enseñanza del español como lengua materna*. Río Piedras: Universidad de Puerto Rico, 1991.

López Morales, Humberto. *La enseñanza de la lengua materna. Lingüística para maestros de español*. Madrid: Playor, 1984.

Murillo Rojas, Marielos. "Diversidad de vocabulario en los preescolares. Aportes para valorar su competencia léxica". *Revista de Filología y Lingüística*, N° 35, 1, 2009, pp. 123-138.

Quesada Murillo, María Rosario; Miguel Ecurra Mayaute y Ana Delgado Vasquez. "Riqueza léxica en estudiantes de educación básica (nivel secundaria) de Lima metropolitana". *Revista de Investigación en Psicología*, N° 1, 4, 2001, pp. 110-116.

Real Academia Española [RAE]. *Nueva gramática de la lengua española*. Madrid: Espasa, 2009.

Reyes Díaz, M. J. "Apuntes para la enseñanza del vocabulario". *Revista de Filología*, N° 25, 2007, pp. 529-538.

Torres González, Antonia Nelsi. "Riqueza léxica en textos narrativos escritos por estudiantes de Tenerife". En Moreno Fernández, Francisco; Francisco Gimeno Menéndez; José Antonio Samper; María Luz Gutiérrez Araus; María Vaquero y César Hernández (eds.), *Lengua, variación y contexto. Estudios dedicados a Humberto López Morales*. Madrid: Arco / Libros, 2003, pp. 435-449.

Valencia Espinoza, Alba; Emilio Avila; Gloria Muñoz Rigollet; Nicolás Núñez; Max Sergio Echeverría y Mónica Véliz. "Evaluación de la riqueza léxica de estudiantes de último año de enseñanza media". *Estudios Filológicos*, 27, 1992, pp. 59-71.