

Incidencia de los años de escolaridad y cantidad de lectura en el léxico disponible de un grupo de estudiantes universitarios del área de pedagogía

José Alejandro Martínez-Lara¹
Universidad Adolfo Ibáñez, Chile

Resumen

En este trabajo, se analiza el léxico disponible (LD) de un grupo de universitarios chilenos según variables relacionadas con el nivel de escolaridad y las prácticas lectoras. Partimos de las hipótesis de que, por un lado, los estudiantes de cuarto año tendrán mayores índices de disponibilidad léxica (IDL) que los del primer año. Por otro lado, los alumnos que dedican más horas semanales a la lectura tendrán mayores IDL que aquellos que dedican poco o ningún tiempo a dicho consumo cultural. En este sentido, los objetivos de esta investigación son: 1) describir el caudal léxico de un grupo de alumnos de la carrera de pedagogía de las menciones de: Castellano y Filosofía, Ciencias biológicas y Parvularia, de una institución adscrita al Consejo de Rectores de Universidades Chilenas, respecto a tres centros de interés (CI): *La lectura*, *La educación* y *La escuela: muebles y materiales* y 2) describir la incidencia de los años de escolaridad y cantidad de horas de lectura en el léxico. Para esto, hemos analizado una

¹ Para correspondencia, dirigirse a: José Alejandro Martínez-Lara (jose.martinez.l@uai.cl), Av. Padre Hurtado 750, Ofic. D-319, Viña del Mar, Chile. ORCID 0000-0002-9755-1241.

muestra, recogida y editada conforme a los criterios generales del Proyecto Panhispánico de Léxico Disponible, y procesada mediante el programa Dispogen. Los resultados han mostrado que el CI con mayor número de palabras fue *La escuela*, mientras que el CI con mayor número de vocablos fue *La educación*. Respecto a los años de escolaridad, hubo un incremento del caudal léxico de los estudiantes de último año en comparación con los del primer año. En cuanto a la cantidad de lectura, pudo apreciarse que a mayor cantidad de horas de lectura mayor es el promedio de palabras.

Palabras clave: léxico-estadística; disponibilidad léxica; lectura; años de escolaridad; pedagogía.

INCIDENCE OF YEARS OF SCHOOLING AND AMOUNT OF READING
IN THE LEXICAL AVAILABILITY OF ONE GROUP OF THE UNIVERSITY
STUDENTS OF THE AREA OF PEDAGOGY

Abstract

In this research, we analyze the lexical availability (LD) of a group of Chilean university students according to variables related to their level of schooling and reading practices. We start from the hypothesis that, on the one hand, fourth-year students will have higher indexes of lexical availability (LDI) than first-year students. On the other hand, students who dedicate more hours per week to reading will have higher LDIs than those who dedicate little or no time to such cultural consumption. Thus, the objectives of this article are: 1) to describe the lexical flow of a group of students of the pedagogy career of the mentions of: Spanish and Philosophy, Biological Sciences and Kindergarten, of an institution attached to the Consejo de Rectores de Universidades Chilenas, with respect to three centers of interest (CI): Reading, Education and School: furniture and materials and 2) to describe the incidence of years of schooling and number of hours of reading in the lexicon. For this purpose, we have analyzed a sample, collected and edited according to the general criteria of the Proyecto Panhispánico de Léxico Disponible, and processed using the Dispogen program. The results have shown that the IC with the highest number of words was The school, while the IC with the highest number of words was Education. Concerning the years of schooling, there was an increase in the lexical stock of the students in their last year compared to those in their first year. Regarding the amount of reading, it could be seen that the more hours of reading, the higher the average number of words.

Key words: lexical-statistical; lexical availability; pedagogy; reading.

Recibido: 01/11/20

Aceptado: 15/07/21

INTRODUCCIÓN²

En el presente trabajo léxico-métrico, analizamos el léxico disponible (LD) de un grupo de estudiantes universitarios de la carrera de pedagogía de las menciones de: Castellano y Filosofía, Ciencias biológicas y Parvularia, de una universidad estatal adscrita al Consejo de Rectores de Universidades Chilenas (CRUCH), sobre tres centros de interés, dos novedosos (*La lectura y La educación*) y uno tradicional (*La escuela: muebles y materiales*).

Los estudios de disponibilidad léxica (DL) se iniciaron en Francia en la segunda mitad del siglo XX durante los trabajos preparatorios de *Le français élémentaire* de Gougenheim *et al.* (1954). Desde entonces, este campo lexicológico ha despertado el interés de investigadores de distintas latitudes. En el ámbito hispánico, esta línea de investigación léxico-estadística ha alcanzado una notable importancia, gracias al impulso que le diera el profesor Humberto López Morales con su investigación acerca del español de Puerto Rico, en 1973, y, años más tarde, con la creación y coordinación del Proyecto Panhispánico de Léxico Disponible (PPHLD)³ (López Morales y Trigo Ibáñez 2019, Serfati 2021).

En un primer estadio de este terreno disciplinar en español, los estudios se enfocaron en el lenguaje de los preuniversitarios, generalmente desde una perspectiva sociolingüística, dialectal, psicolingüística y de enseñanza de lengua (Gómez Molina y Gómez Devís 2004; Trigo Ibáñez 2011; Mahecha y Mateus Ferro 2017; Trigo Ibáñez *et al.* 2019; Cerda *et al.* 2021; Perdomo y Alfonso 2021). Empero, este campo lexicológico se amplió hacia otras realidades lingüísticas, como el léxico de niños (Mesa 1989; Gómez Devís 2019; Jiménez 2019), adultos mayores (Urzúa 2018), comunidades de hablas estratificadas sociolingüísticamente (Ávila Muñoz y Villena 2010), universitarios (Rojas *et al.* 2017; Blanco *et al.* 2020; Ávila *et al.* 2020) e incluso hacia el área de la literacidad en salud (Castillo Fadić y Pino Castillo

² Este trabajo es parte de mi proyecto de tesis doctoral, financiado por la Agencia Nacional de Investigación y Desarrollo (ANID) de Chile, a través de la Beca de Doctorado Nacional, folio n° 21192227.

³ Para más detalles acerca del PPHLD, diríjase a <http://www.dispoplex.com/>

2020), con la realización de estudios comparados entre el léxico disponible de profesionales de la salud y personas con condiciones cardiovasculares crónicas (Castillo Fadić y Santos Díaz en prensa). Estas investigaciones en español como lengua materna impulsaron también los estudios de DL de español como lengua extranjera y han influido en los estudios de enseñanza de lenguas (Lin 2012; Sefarti 2016, 2021; Mendoza 2018; Herranz y Marcos 2019; Aabidi 2020; Santos Díaz *et al.* 2020). Además, están emergiendo nuevas aplicaciones lingüísticas basadas en los datos de DL, como la creación y análisis de corpus cacográficos, descripción de procesos neológicos, aplicaciones metodológicas, entre otros (García 2011; Paredes 2012a, 2018; Paredes y Gallego 2019; Trigo Ibáñez *et al.* 2020; Perdomo y Alfonso 2021; Reyes *et al.* 2021).

Según el arqueo bibliográfico, puede observarse la aplicación de variables sociales y psicolingüísticas recurrentes, como sexo, tipo de institución, zona de residencia, edad de adquisición de una palabra, tipicidad, familiaridad, estilos de aprendizaje, etc. (Hernández Muñoz e Izura 2010; Hernández Muñoz *et al.* 2014; Hernández Muñoz y Tomé 2017; Domínguez 2019; Ferreira *et al.* 2019). Sin embargo, son pocos los trabajos que han tomado en consideración la incidencia de los consumos culturales, como la lectura, en el léxico disponible de universitarios en formación. En este sentido, resulta importante, por un lado, aportar datos acerca de la posible asociación entre el léxico disponible y los hábitos lectores de un grupo de estudiantes. Por otro lado, conocer el léxico disponible según el nivel de escolaridad de los sujetos. En este estudio, nos interesa conocer el LD de un grupo de alumnos chilenos que se forman para ser maestros.

En este orden de ideas, una investigación que combine factores como año de escolaridad y frecuencia de lectura podría ayudar—desde una perspectiva léxico-estadística— a conocer los alcances de los programas universitarios. Además, ofrecería un panorama sobre el papel de la lectura en el desarrollo académico y profesional del futuro profesorado. Por último, permitiría evaluar la motivación por la lectura y la evolución o no del vocabulario de los sujetos del corpus.

En virtud de lo anterior, considerando centros de interés vinculados a la docencia: *La lectura, La educación y La escuela: muebles y materiales*, nos preguntamos ¿puede observarse alguna incidencia de los años de escolaridad en el léxico disponible de un grupo de estudiantes de la carrera de educación, menciones de Castellano y Filosofía, Ciencias biológicas y Parvularia? ¿Las prácticas lectoras—específicamente, cantidad de horas de lectura optativa—motivan el aumento o no del caudal léxico del grupo bajo análisis?

En este contexto, los objetivos que guían esta investigación son: 1) describir el caudal léxico de un grupo de estudiantes de pedagogía de una

universidad estatal perteneciente al CRUCh, respecto de los centros de interés: *La lectura, La educación y La escuela: muebles y materiales*; y 2) describir la incidencia de los años de escolaridad y cantidad de horas de lectura optativa en el léxico disponible del grupo de encuestados. Por su parte, los objetivos específicos son:

- Conocer los índices generales y específicos de DL en los centros de interés: *La lectura, La educación y La escuela: muebles y materiales*.
- Comparar los índices generales de DL entre los alumnos de primer año y los de cuarto año del grupo bajo análisis.
- Describir la incidencia de la cantidad de horas de lectura optativa en el caudal léxico de los informantes.
- Conocer los vocablos más disponibles en los actualizadores: *La lectura, La educación y La escuela: muebles y materiales*.

En consideración a lo anterior, este es un trabajo descriptivo, cuantitativo, basado en fuentes primarias, con un enfoque sociolingüístico. En los siguientes epígrafes, presentamos los aspectos generales de los estudios de disponibilidad léxica, junto con algunos antecedentes en los que se analizaron variables relacionadas con la lectura; describimos la metodología llevada a cabo; mostramos los resultados generales y específicos; exponemos la discusión de los resultados; y, por último, presentamos las conclusiones.

2. ASPECTOS GENERALES DE LA DISPONIBILIDAD LÉXICA

El término léxico disponible fue acuñado por Michéa, cuando, a partir de una serie de experimentos, logró distinguir las palabras atemáticas de las temáticas. Las primeras aparecen en la interacción, sea cual sea el tema tratado. A este grupo pertenecen las preposiciones, conjunciones, pronombres; verbos y sustantivos de significado general. Por ejemplo: *a, en, por, para, ser, estar, hombre, persona*, etc. Es decir, las palabras atemáticas son, en su mayoría, funcionales. Por su parte, las segundas (temáticas) son las lexías que aparecen en la interacción solo si son requeridas por el contexto enunciativo, ya que están condicionadas por el tópico de la interacción. Así pues, en este conjunto se encuentran los sustantivos y verbos de alta carga léxico-semántica (*mariposa, banana, espejo, constelación, cantar, bailar, correr, manejar, chocar*, etc.). En virtud de esto, el lexicógrafo francés

planteó la diferenciación entre el “léxico frecuente”, que está compuesto por las palabras atemáticas, y el “léxico disponible”, constituido por las palabras temáticas (Gougenheim *et al.* 1954: 144; López Morales 1999: 10-11; Gómez Devís 2019: 12). En consecuencia, en los inicios de los estudios de disponibilidad léxica, Michéa (1953: 340) afirmó que:

Un mot disponible est un mot qui, sans être particulièrement fréquent, est cependant toujours prêt à être employé et se présente immédiatement et naturellement à l'esprit au moment où l'on en a besoin.

De esta manera, el LD se refiere al caudal léxico que más rápido viene a la mente de un individuo cuando se activa una parcela léxico-semántica en un contexto comunicativo específico. En los estudios de DL, este tipo de vocabulario se activa mediante centros de interés⁴ (CI), entendidos como el estímulo que mueve cognitivamente la producción de palabras en la mente de los encuestados (López Morales 1995-1996: 245; Hernández Muñoz 2009: 224; Lara 2006: 173; Paredes 2012b: 54; Sánchez-Saus 2019).

En este orden de ideas, hay que resaltar también que la disponibilidad léxica, como parte de la línea de investigación léxico-estadística, opera sobre el *vocablo*, entendido como la unidad mínima del léxico, y no sobre la *palabra* o unidad mínima textual (Castillo Fadić y Sologuren 2020: 71). Entonces, un vocablo es la unidad léxica que agrupa todas las variantes de una raíz derivada y/o flexionada (Lara 2006: 138; Ávila Muñoz y Lasarte 2010: 93).

3. LÉXICO DISPONIBLE DE UNIVERSITARIOS Y SU RELACIÓN CON LA LECTURA

Las investigaciones enmarcadas en el PPHLD se han centrado en el caudal léxico de estudiantes de secundaria y han tenido como finalidad “Elaborar diccionarios de Disponibilidad Léxica para las diversas zonas del mundo hispánico” (Dispolex 2003-2021)⁵. Sin embargo, en las últimas décadas ha habido un interés por abordar otras realidades lingüísticas. Algunos

⁴ Este término suele recibir otros nombres como: *eje temático*, *actualizador*, *área temática* o *área nocional*, que son utilizados como sinónimos en este texto.

⁵ Para más detalles, visitar: <http://www.dispolex.com/info/el-proyecto-panhispanico>

investigadores, por ejemplo, se han dedicado a analizar el caudal léxico de áreas de especialidad, con el objetivo de conocer y describir el vocabulario con fines comunicativos específicos, como los desarrollados por Guerra y Gómez (2003), Rojas *et al.* (2017), Herranz (2018, 2020), Valenzuela *et al.* (2018), Santos Díaz (2020), Zambrano (2021), entre otros.

Respecto al trabajos de Herranz (2018), la investigadora analizó el léxico disponible de futuros profesores españoles, quienes cursaban primer y cuarto año de las carreras de Educación Infantil y Educación Primaria de una universidad madrileña. El objetivo general era determinar si el léxico de este grupo de estudiantes universitarios españoles aumentaba, disminuía o se mantenía al final de la carrera con respecto al punto de inicio (Herranz 2018: 146). En este mismo orden de idea, su objetivo específico fue comparar los índices generales de disponibilidad léxica de los encuestados de primer año con los de los de cuarto año. La prueba constaba de 18 centros de interés, los 16 tradicionales⁶ y 2 nuevos (*Nuevas tecnologías: TIC y Educación*). Los resultados mostraron que los estudiantes de cuarto año tenían mayores índices de DL que los de primer año de la carrera, lo que supuso que hay un incremento del caudal léxico de los universitarios a medida que avanzan en su formación.

Sobre las variables referidas a consumos culturales como la lectura, estas han sido poco analizadas en los trabajos de DL. No obstante, en el área de la pedagogía, ha habido un mayor interés en estudiar la relación entre la comprensión lectora y el componente léxico. En este sentido, pueden mencionarse las investigaciones de Cepeda *et al.* (2014) y Cepeda *et al.* (2016). Ambos estudios analizaron cómo el caudal léxico incidía en la comprensión de textos académicos de niños en edad escolar. Los investigadores observaron que los sujetos con un mayor caudal léxico obtienen un mayor puntaje en las pruebas de comprensión lectora y, también, en las notas generales del curso. Igualmente, la importancia del léxico en la comprensión lectora se evidencia a través de los resultados de estudios de adquisición léxica en L2. En algunos de estos, se ha determinado que si un estudiante logra alcanzar el umbral léxico adecuado tendrá una

⁶ Los dieciséis ejes temáticos tradicionales y utilizados en el PPHLD son: 1) Las partes del cuerpo, 2) La ropa, 3) La casa, 4) Los muebles de la casa, 5) Los alimentos y las bebidas en la comida, 6) Los objetos colocados sobre la mesa que se utilizan para todas las comidas del día, 7) La cocina, sus muebles y los utensilios que en ella se encuentran, 8) La escuela, sus muebles y su material escolar, 9) La calefacción y la iluminación, 10) La ciudad, 11) El pueblo, 12) Los medios de transporte, 13) Los trabajos del pueblo y del jardín, 14) Los animales, 15) Los juegos y las distracciones, 16) Los oficios.

mayor comprensión lectora en L2 (Llinares *et al.* 2008; Santos Palmou 2017). Al respecto, Santos Palmou (2017: 166), basándose en estudios que correlacionan la variable léxica con la variable comprensión lectora, afirma que: “el éxito de la comprensión lectora depende directamente del conocimiento del léxico”.

En el campo de la estadística léxica, podemos mencionar los trabajos de Ávila Muñoz (2007) y Santos Díaz (2017a), que relacionaron el léxico con la frecuencia de lectura. Ávila Muñoz (2007) buscaba conocer cuáles eran las variables sociales y los hábitos culturales que incidían en la ortografía de estudiantes de la carrera de Traducción e Interpretación de la Universidad de Málaga, basándose en datos de léxico disponible. Entre los denominados hábitos culturales, integró la variable *Cantidad de libros leídos al año*. Los resultados mostraron que las faltas ortográficas aumentaban a medida que disminuían los hábitos lectores de los encuestados, por lo que el autor afirmó que la cantidad de libros leídos podía incidir significativamente en el léxico. Si bien, este trabajo centró su atención más en la producción de errores ortográficos en un corpus de DL, muestra la importancia de considerar factores culturales y cognitivos (como la actividad lectora) para describir y explicar la producción de palabras. Por su parte, Santos Díaz (2017a) estudió la incidencia de la frecuencia de lectura (cantidad de libros leídos al año) en el LD y en la identificación de tecnicismos. Los datos utilizados fueron producidos por estudiantes españoles del posgrado de profesorado en tres lenguas: español, inglés y francés. El análisis de los datos reveló que, por un lado, en los casos del español y el francés no hay una relación estadística significativa entre el caudal léxico y la frecuencia de lectura; mientras que, en la prueba en inglés, los datos mostraron lo contrario: hay una relación significativa entre el LD y la variable de lectura. No obstante, con base en un análisis más detallado, la autora concluyó que “la lectura facilita la actualización de un mayor número de palabras y de identificación de tecnicismos” (Santos Díaz 2017a: 85).

4. METODOLOGÍA

4.1. HIPÓTESIS

Las dos hipótesis que revisamos en este trabajo han sido las siguientes:

- Los estudiantes de cuarto año de la carrera de pedagogía tendrán mayores índices generales de disponibilidad léxica que los estudiantes del primer año.
- Los estudiantes que dedican más horas semanales a la lectura optativa tendrán mayores índices de disponibilidad léxica que aquellos que dedican poco o ningún tiempo a dicha práctica.

4.2. INFORMANTES

Los datos analizados en este estudio son de fuentes primarias, provenientes de las listas de palabras elaboradas por 55 estudiantes universitarios de la carrera de pedagogía, de las menciones: Castellano y Filosofía, Ciencias biológicas y Parvularia, de una universidad estatal del norte de Chile del grupo del CRUCh. 32 informantes cursaban primer año y 23, cuarto año. Cada estudiante contestó un test de disponibilidad léxica de nueve CI. Sin embargo, en este texto, presentamos únicamente los resultados de los siguientes tres ejes temáticos: 1) *La lectura*, 2) *La educación* y 3) *La escuela: muebles y materiales*. Así pues, el corpus de este trabajo son 165 listas, con un total de 2209 palabras y 974 vocablos.

4.3. INSTRUMENTOS Y RECOLECCIÓN DE LOS DATOS

Para la recolección de los datos, aplicamos una prueba compuesta por tres instrumentos, más un consentimiento informado. El primero, *Encuesta sociológica*, recogía la información social de los encuestados. El segundo instrumento era el test de DL, que consistía en hojas en blanco, una por cada eje temático, con las iniciales de cada CI. A manera de ejemplo, en la página correspondiente al área nocional *La escuela: muebles y materiales*, en la parte superior de la hoja estaba rotulada la sigla LEMYM. El último instrumento, *Encuesta de hábitos lectores*, fue una batería sobre prácticas lectoras. Este cuestionario fue el resultado de la adaptación de la encuesta de hábitos lectores propuesta por el proyecto Fondecyt Regular n° 1170779.

La prueba, que contemplaba los tres instrumentos más el consentimiento informado⁷, fue aplicada por el mismo investigador en todos los grupos, quien además explicó en qué consistía el estudio y cómo sería el procedimiento para responder los cuestionarios. El test de DL siguió las pautas metodológicas planteadas por el PPHLD: cada sujeto elaboró manualmente sus listas de palabras, para lo cual contaban con dos minutos por cada CI. El encuestador indicaba el tiempo para empezar y terminar de escribir las listas y el momento para pasar a la siguiente área nocional. Una vez que los sujetos terminaban de escribir la lista de un eje temático, pasaban al siguiente sin posibilidad de regresar al anterior.

4.4. EDICIÓN DE LOS MATERIALES

Las listas de palabras fueron editadas siguiendo los lineamientos propuestos por Samper Padilla (1998) y asumidos por el PPHLD. A grandes rasgos, estos criterios son los siguientes: primero, se mantuvo el orden de aparición de las palabras en las listas; segundo, se eliminaron las palabras repetidas en una misma lista, las palabras incomprensibles y las pseudopalabras⁸, es decir, las unidades léxicas que no estaban registradas en los diccionarios ni formaban parte de los tecnicismos del área; tercero, los sustantivos se lematizaron en singular y género masculino; cuarto, los verbos conjugados se lematizaron en infinitivo; quinto, la unificación ortográfica siguió las pautas de la RAE-ASALE (2005; 2014). Por último, se utilizó el paréntesis en los casos en que una pieza léxica pudiera expresarse con la presencia o ausencia de algún elemento, ejemplo: *aprendizaje (autónomo)*.

4.5. PROCESAMIENTO DE LOS DATOS

Una vez que se transcribieron y editaron todas las listas de palabras en hojas de Excel©, se codificaron en formato Salamanca de acuerdo

⁷ Debe señalarse que la participación de los estudiantes fue voluntaria, por lo que los cuestionarios analizados corresponden solo al de los sujetos que dieron su consentimiento. Los datos proporcionados por los encuestados están codificados de manera que se resguarda el anonimato. Agradecemos a todos los estudiantes por su valiosa colaboración.

⁸ Debemos acotar que, si bien estas unidades no se analizaron en este trabajo, no fueron eliminadas de la base de datos. Por el contrario, están archivada para ser analizadas cualitativamente en futuros trabajos.

con los requerimientos de Dispogen (Echeverría *et al.* 2005), programa computacional creado específicamente para los estudios de DL. Este calcula los índices generales: Total de palabras (TP), Total de vocablos (TV), Promedio de palabras (PP) e Índice de cohesión (IC); y los Índices de Disponibilidad Léxica (IDL) de cada centro de interés, según la fórmula de López Chávez y Strassburger (1987), revisada por los autores en 2001.

El índice de disponibilidad indica el grado de disponibilidad que tiene una palabra para un grupo de sujetos en un tópico específico. Este índice se calcula al ponderar estadísticamente la frecuencia de aparición de un vocablo y la posición de ese vocablo en las listas (Ávila Muñoz y Sánchez 2011: 52; Ávila Muñoz 2016: 3). En tanto que el índice de cohesión muestra el grado de relación o coincidencia de los vocablos producidos por todos los sujetos en cada eje temático (Echeverría *et al.* 1987: 68; Casanova 2017: 146).

Hemos procesado los datos cuantitativamente según las siguientes dos variables: *Año de la carrera* y *Cantidad de horas semanales de lectura optativa*. Para cada una realizamos los análisis descriptivos con base en los valores arrojados por Dispogen. Además, aplicamos pruebas estadísticas paramétricas (t de Student y Anova) con el fin de determinar la significancia de la distribución de los datos por cada una de las variables.

5. RESULTADOS

5.1. ÍNDICES GENERALES DE DISPONIBILIDAD LÉXICA

En el Cuadro 1, puede observarse que en el corpus se contabilizaron 2209 palabras totales en los tres actualizadores, que pueden traducirse en una media de 736,33 unidades léxicas por cada área temática y 40,16 por sujeto. El centro de interés *La escuela: muebles y materiales* fue el que tuvo el mayor número de palabras: 797 (36 %) y, por tanto, el mayor rango (1); seguido por *La educación*, con 766 (35 %), rango 2; y *La lectura*, con 646 (29 %), rango 3. Puede verse que no existe gran oscilación porcentual entre los tres ejes temáticos, siendo de apenas 1 punto porcentual de diferencia entre los

CI *La escuela*⁹ y *La educación*; y 7 puntos de separación entre el CI con menos palabras (*La lectura*) y el de mayor número de palabras (*La escuela*).

Centro de Interés	Rango	TP	TV	PP	IC
La escuela	1	797	283	14,49	0,051
La educación	2	766	405	13,93	0,034
La lectura	3	646	286	11,75	0,041

Cuadro 1. Índices generales de DL por centro de interés, ordenados jerárquicamente según el rango

Respecto al número de vocablos, los tres actualizadores sumaron 974 lemas, lo que dio un promedio de unidades léxicas diferentes de 324,66 por área nocional, y 17,70 por informante. El centro de interés *La educación* fue el que tuvo mayor número de unidades de cita: 405, seguido por *La lectura*, con 286, y, por último, *La escuela*, con 283. El CI *La educación* fue el que tuvo el mayor porcentaje de vocablos (42 %), mientras que los otros dos actualizadores obtuvieron el mismo porcentaje (29 %). En este caso, la diferencia porcentual entre los dos ejes temáticos con menos vocablos (*La lectura* y *La escuela*) y el de mayor número (*La educación*) fue de 13 puntos.

En consideración a los promedios de palabras, puede observarse que el área nocional con mayor rango (*La escuela*) tuvo el mayor promedio de palabras (14,49), seguido por *La educación* (13,93), rango 2; y, por último, *La lectura* (11,75), rango 3. La media de palabras por centro de interés fue 13,39, así los actualizadores *La escuela* y *La educación* tuvieron un promedio de palabras por encima de la media, mientras que *La lectura* estuvo por debajo de este índice.

Por último, los resultados indicaron que el centro de interés con mayor índice de cohesión fue *La escuela*, con 0,051; seguido por *La lectura*, con 0,041, y *La educación*, con 0,034. Estos valores señalaron que el actualizador *La escuela* fue el más compacto de los tres y estuvo un poco por encima del límite del percentil referencial en este tipo de estudios (0,05) (cf. Sefarti 2016: 113), mientras que los otros dos CI estuvieron por debajo de dicho percentil, siendo el CI *La educación* el más difuso o abierto de los campos nocionales aquí estudiados. Debemos acotar que los tres CI mostraron IC altos, por lo

⁹ El nombre del centro de interés es *La escuela: muebles y materiales*. Sin embargo, por razones de espacio, en el texto, los cuadros y gráficos, se simplificará tan solo como *La escuela*.

que puede afirmarse que los tres son compactos, según la clasificación de Mateus Ferro *et al.* (2018: 137-138).

5.2. ÍNDICES DE DISPONIBILIDAD LÉXICA POR AÑO DE LA CARRERA

Una variable social recurrente en los estudios de disponibilidad léxica ha sido el nivel educativo de los encuestados. Esta ha resultado importante en numerosos estudios (cf. Echeverría *et al.* 1987; Salcedo y Leo 2013). En el Cuadro 2, presentamos los resultados de este factor social en este trabajo.

Centro de interés \ Año de la carrera	Primer Año				Cuarto Año			
	TP	TV	PP	IC	TP	TV	PP	IC
La lectura	356	184	11,13	0,060	290	175	12,61	0,072
La educación	380	237	11,88	0,050	386	248	16,78	0,068
La escuela	445	165	13,91	0,084	352	186	15,30	0,082

Cuadro 2. Distribución de los valores de los índices generales de cada CI según la variable Año de la carrera

Como puede apreciarse en el Cuadro 2, el total de palabras reportadas por los estudiantes de primer año en los tres ejes temáticos fue 1181, que da un promedio de 393,66 palabras por área nocional y 36,91 por sujeto. El mayor número de palabras de este grupo se registró en el CI *La escuela* (445/38 %), seguido por *La educación* (380/32 %) y *La lectura* (356/30 %). Por su parte, los alumnos de cuarto año mencionaron 1028 palabras, traducidas en un promedio de 342,66 por centro de interés y 44,70 por sujeto. Este grupo reportó más palabras en *La educación* (386/38 %), seguido por *La escuela* (352/34 %), y *La lectura* (290/28 %). Este índice no puede verse aisladamente, pues el total de palabras depende en gran medida del tamaño muestral (Serfati 2016: 108). Sin embargo, podemos observar que la distribución del total de palabras por individuo es mayor entre los informantes de cuarto año frente a los de primer año –44,70 y 36,91 respectivamente–.

A pesar de las limitaciones del índice de Total de palabras, existe una diferencia significativa ($< 0,001$) entre el grupo de primer año y el de cuarto año, según la prueba t de Student.

En referencia al número de vocablos, los encuestados del primer año escribieron 586 palabras diferentes, un promedio de 135,33 unidades léxicas por CI y 18,31 por sujeto. El eje temático con mayor número de vocablos entre los estudiantes de primer año fue *La educación* (237/40 %), seguido por *La lectura*, con 184 (31 %) lemas, y *La escuela*, con 165 (28 %). Por su parte, los informantes de cuarto año reportaron 609 lemas, lo que significa 203 piezas léxicas por área nocional y 26,48 por encuestado. En este caso, el área nocional con mayor número de vocablos fue *La educación*, con 248 lemas (41 %); seguido por *La escuela*, con 186 (31 %), y *La lectura*, con 175 (29 %). En el Gráfico 1, a continuación, ilustramos los resultados del Total de palabras (TP) y Total de vocablos (TV).

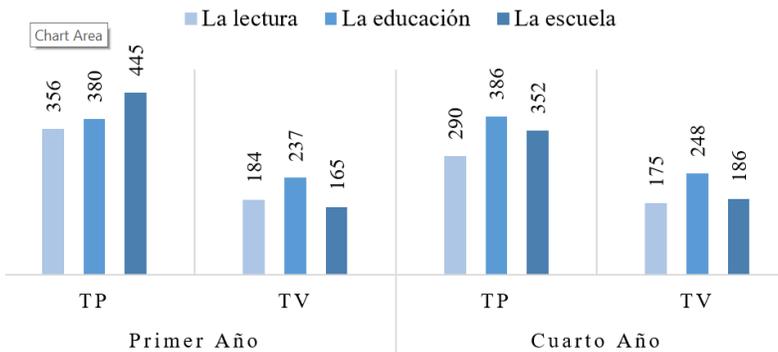


Gráfico 1. Total de palabras y Total de vocablos por CI, según la variable Año de la carrera

Respecto al promedio de palabras, entre los sujetos del primer año, el eje temático con mayor PP fue *La escuela*, con 13,91; en segundo lugar, *La educación*, con 11,88; y, por último, *La lectura*, con 11,13. En cuanto a los universitarios de cuarto año, el mayor PP lo obtuvo el actualizador *La educación* (16,78), seguido por *La escuela* (15,30) y *La lectura* (12,61). Es importante resaltar que en los tres ejes temáticos hubo un aumento del PP en los estudiantes de cuarto años en comparación con los del primer año, siendo justamente el CI *La educación* el que tuvo el mayor incremento, de 11,88 a 16,78, es decir 4,9 palabras de diferencia entre ambos grupos. En el Gráfico 2, exponemos estos resultados.

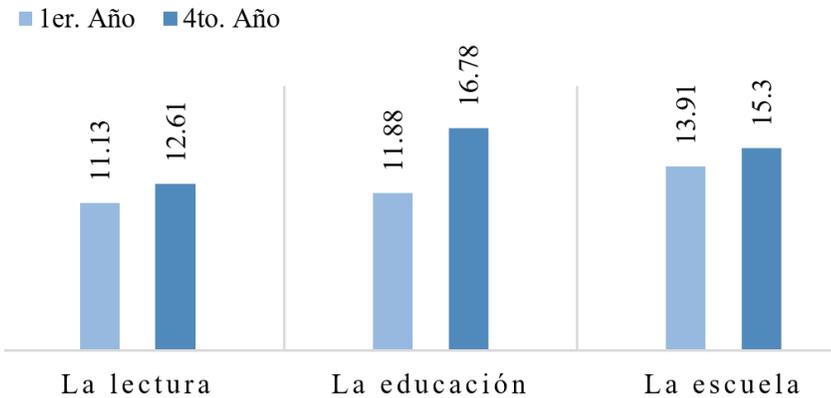


Gráfico 2. Promedio de palabras por CI según la variable Año de la carrera

En relación con el índice de cohesión entre los universitarios de primer año, el eje temático con mayor IC fue *La escuela* (0,084), seguido por *La lectura* (0,060), y *La educación* (0,050). Por su parte, entre los informantes de cuarto año, el eje temático con el mayor IC fue *La escuela*, 0,082; seguido por *La lectura* (0,072) y *La educación* (0,068). Puede observarse que ambos grupos coincidieron en tener los IC más altos en el CI *La escuela*. Estos valores son \geq al percentil (0,05), por lo que puede decirse que los tres actualizadores son compactos. En el Gráfico 3, presentamos estos resultados.

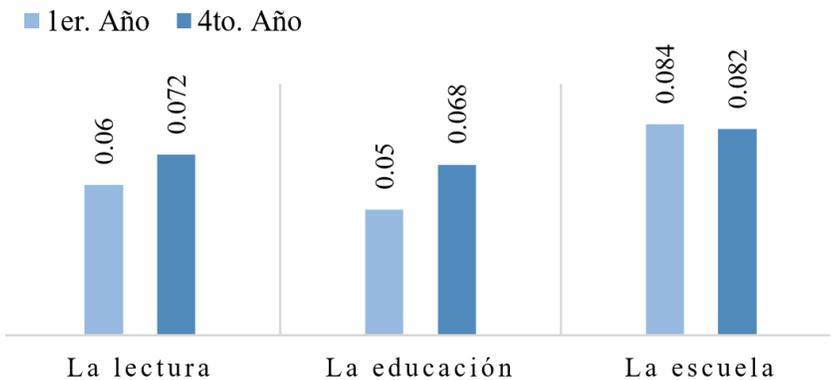


Gráfico 3. Índice de cohesión por CI según la variable Año de la carrera

5.3. ÍNDICES DE DISPONIBILIDAD LÉXICA Y HORAS SEMANALES DE LECTURA OPTATIVA

¿Horas semanales? Centro de interés	Ninguna				1 a 5 horas				6 a 10 horas				Más de 10 horas			
	TP	TV	PP	IC	TP	TV	PP	IC	TP	TV	IC	TP	TP	TV	PP	IC
La lectura	73	54	10,43	0,193	485	239	11,54	0,048	54	47	13,5	0,287	34	32	17	0,531
La educación	101	91	14,42	0,158	550	312	13,09	0,041	73	63	18,25	0,289	42	41	21	0,512
La escuela	104	64	14,85	0,232	643	261	15,31	0,058	68	47	17	0,361	48	39	24	0,615

Cuadro 3. Índices generales de disponibilidad léxica según la variable Cantidad de horas semanales de lectura optativa

La variable *Cantidad de horas semanales de lectura optativa* está compuesta por cuatro variantes: 1) *ninguna hora semanal de lectura optativa*, 2) *de 1 a 5 horas semanales de lectura optativa*, 3) *de 6 a 10 horas semanales de lectura optativa*, y 4) *más de diez horas de lectura optativa*. Como puede observarse en el Cuadro 3, el mayor número de palabras se encuentra entre los informantes que indicaron que leen por placer entre 1 y 5 horas semanales, sumando entre los tres CI 1678 palabras; en segundo lugar, los encuestados que señalaron que no leen por placer sumaron 278 palabras; seguidos por los que afirmaron leer entre 6 y 10 horas semanales, 195 palabras; por último, quienes indicaron que leen más de 10 horas semanales reportaron 124 palabras. No obstante, debemos recalcar que este índice depende del tamaño de la muestra; en este caso la mayoría de los encuestados se decantó por afirmar que realizaba lectura optativa entre 1 a 5 horas por semanas, mientras que un bajo número afirmó leer más de 10 horas semanales. Con el objetivo de determinar si existen diferencias significativas entre los grupos de esta variable, hemos aplicado la prueba paramétrica de Anova, que ha arrojado una significancia de 0,101, mayor a 0,05, y valor $F = 2,184$. Es decir, la distribución de las medias entre los grupos no tiene diferencia significativa.

En cuanto al total de vocablos, se aprecia el mismo patrón que en el índice anterior. El mayor número de palabras diferentes se observa entre los encuestados que realizan lectura optativa entre 1 a 5 horas semanales (812 lemas), seguido por quienes indicaron que no leen (209 vocablos), los que leen entre 6 y 10 horas semanales (157 lemas) y los que afirmaron leer más de 10 horas semanales (112 vocablos).

En relación con el PP, puede apreciarse que los promedios de palabras por centros de interés aumentaron proporcionalmente a medida que se pasó

de la variante 1 (*ninguna hora semanal de lectura optativa*) a la variante 4 (*más de 10 horas semanales de lectura optativa*).

En el caso del área nocional *La lectura*, esta tuvo el PP más bajo en comparación con los otros dos centros de interés, pasando de un promedio de palabras de 10,43 en la variante 1, hasta alcanzar un promedio de 17 palabras en la variante cuatro. Es decir, hubo un aumento de casi 7 palabras. Por su parte, en el CI *La educación* los encuestados que indicaron no leer nunca textos optativos tuvieron un PP de 14,42, llegando a alcanzar un valor de 21 palabras por sujeto entre los informantes que leen más de 10 horas semanales; es decir, hubo un incremento de 6,58 palabras. Por último, el CI *La escuela* fue el que tuvo los PP más altos de las tres áreas nocionales; en este, el PP en la variante *ninguna hora semanal de lectura optativa* fue de 14,85 y llegó a un valor de 24 palabras en la cuarta variante (*más de 10 horas semanales de lectura optativa*), es decir hubo un incremento de 9,15 palabras. Estos resultados pueden apreciarse detalladamente en el Gráfico 4.

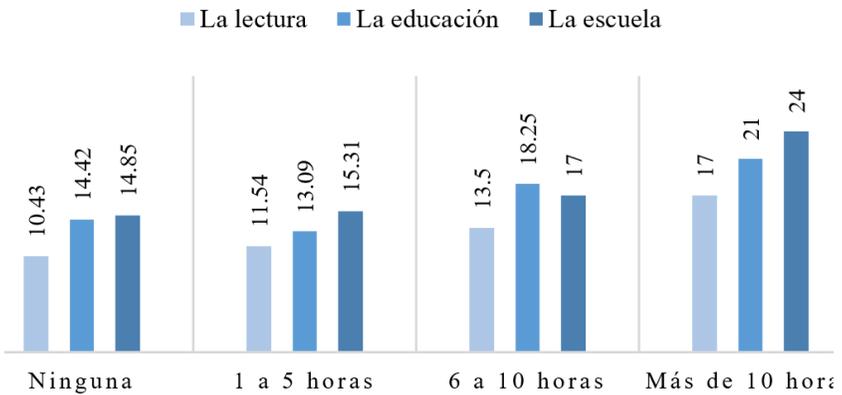


Gráfico 4. Promedio de palabras por CI según la variable Cantidad de horas semanales de lectura optativa

Respecto al índice de cohesión, puede verse que los valores aumentaron a medida que se pasó de la variante 1 (*ninguna hora semanal de lectura optativa*) a la variante 4 (*más de 10 horas semanales de lectura optativa*). Si bien todos los valores de IC de esta variable pueden ubicarse en una escala de compactibilidad alta, según la propuesta de Mateus Ferro *et al* (2018), este nivel de compactibilidad se va acrecentando o fortaleciendo a medida que se pasa de la variante 1 a la variante 4 de lectura optativa. En esta variante, los valores fueron bastante bajos comparativamente con las otras tres. Los

actualizadores *La educación* (0,041) y *La lectura* (0,048) alcanzaron valores menores al percentil de referencia (0,05), que fue superado solo por el CI *La escuela* (0,058). Sin embargo, los IC aumentaron exponencialmente en las variantes 3 y 4, siendo más altos en el último nivel de análisis (*más de 10 horas semanales de lectura opcional*). En este punto, el CI *La escuela* fue el más cohesionado, con un valor de 0,615, seguido por *La lectura*, con 0,531, y *La educación*, con 0,512. En el Gráfico 5, ilustramos estos resultados.

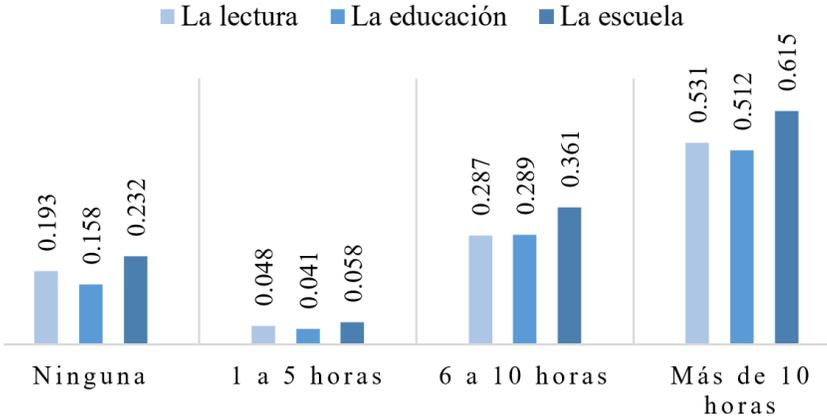


Gráfico 5. Índice de cohesión léxica de cada CI según la cantidad semanal de lectura opcional

5.4. ÍNDICES DE DISPONIBILIDAD LÉXICA POR CENTRO DE INTERÉS

En este tipo de estudio léxico-estadístico, el índice más importante es justamente el de disponibilidad léxica (IDL). Este indica cuáles son las palabras más disponibles en un grupo de encuestados. Como mencionamos en la Introducción, entre los objetivos de este trabajo se encuentra conocer las palabras más disponibles de los siguientes centros de interés: *La lectura*, *La educación* y *La Escuela: muebles y materiales*, todos vinculados al área de formación de los sujetos: la pedagogía. Si bien no hay unanimidad entre los investigadores en definir un criterio de corte en la selección de las palabras más disponibles (Santos Díaz 2017b), hemos considerado idóneo presentar en este trabajo los vocablos con $IDL \geq 0,1$.

5.4.1. La lectura

Vocablo	IDL	Frecuencia	% Aparición	Frec. Acumulada
libro	0,49110377	32	58,1818163	0,0495356
letra	0,33719492	23	41,8181807	0,08513932
comprensión	0,21643822	17	30,9090912	0,11145511
texto (científico)	0,2027128	15	27,2727281	0,13467492
leer	0,18694994	14	25,454545	0,15634675
cuento	0,14001346	13	23,6363634	0,17647059
palabra	0,13274349	10	18,1818187	0,19195046

Cuadro 4. Vocablos con $IDL \geq 0,1$ en el CI La lectura

Como puede observarse en el Cuadro 4, en el CI *La lectura* solo 7 vocablos tuvieron un $IDL \geq 0,1$, siendo el lema *libro* el que tuvo el valor más alto (0,49), seguido por *letra*, *comprensión*, *texto (científico)*, *leer*, *cuento* y *palabra*. Además, puede apreciarse que, entre los tres primeros vocablos más disponibles, los IDL tienen una mayor oscilación en cuanto a la frecuencia y porcentaje. Respecto al porcentaje de aparición, puede verse que solo el lexema más disponible supera el 50 % en el corpus, con una diferencia de 16,36 puntos porcentuales de separación entre este y el vocablo que le sigue (*letra*).

De los siete vocablos más disponibles de este eje temático, seis fueron sustantivos y uno fue un verbo. Según sus propiedades semánticas, pueden agruparse en dos tipos: por un lado, lexemas que refieren a ‘elementos básicos de la lectura’ (*libro*, *letra*, *texto (científico)*, *cuento* y *palabra*); por otro lado, unidades léxicas que aluden a ‘procesos cognitivos vinculados a la lectura’ (*comprensión* y *leer*).

5.4.2. La educación

Vocablo	IDL	Frecuencia	% Aparición	Frec. Acumulada
aprendizaje (significativo)	0,31156072	22	40,0000006	0,02872063
profesor	0,26719567	20	36,3636374	0,05483029
conocimiento	0,21032353	16	29,0909082	0,07571802
colegio	0,20048188	18	32,7272713	0,09921671
enseñanza	0,15833077	11	20,0000003	0,11357702
oportunidad	0,12051546	9	16,3636357	0,12532637
derecho	0,11479283	7	12,7272725	0,13446475
estudiante	0,10574218	9	16,3636357	0,1462141

Cuadro 5. Vocablos con $IDL \geq 0,1$ en el CI *La educación*

En el centro de interés *La educación* se registraron 8 vocablos con $IDL \geq 0,1$, siendo *aprendizaje (significativo)* el que obtuvo el valor más alto (0,31), seguido de *profesor*, *conocimiento*, *colegio*, *enseñanza*, *oportunidad*, *derecho* y *estudiante*. No se observaron grandes diferencias entre la frecuencia y los porcentajes de este grupo de vocablos. Sin embargo, llamó la atención que ninguno de los vocablos de este conjunto de lemas haya alcanzado el 50 %, por el contrario, la pieza léxica más disponible obtuvo solo 40 % de aparición.

Según se aprecia en el Cuadro 5, seis lexemas fueron sustantivos y dos, adjetivos, los cuales pueden agruparse en cuatro grupos semánticos: 1) lexemas referidos a ‘actores de la educación’ (*profesor* y *estudiante*); 2) palabras concernientes a ‘lugar vinculado con la educación’ (*colegio*); 3) piezas léxicas alusivas a ‘procesos cognitivos de la educación’ (*aprendizaje (significativo)*, *enseñanza* y *conocimiento*); y 4) unidades léxicas referidas a ‘cualidades de la educación’ (*oportunidad* y *derecho*).

5.4.3. *La escuela: muebles y materiales*

Vocablo	IDL	Frecuencia	% Aparición	Frec. Acumulada
mesa	0,63033897	40	72,7272749	0,05018821
silla	0,61436641	41	74,545455	0,10163112
pizarra	0,46905699	37	67,2727287	0,14805521
plumón	0,34943011	33	60,0000024	0,18946048
lápiz (de color)	0,28968629	28	50,90909	0,22459222
cuaderno	0,2307273	23	41,8181807	0,25345044
libro	0,22321709	23	41,8181807	0,28230866
computador	0,12774949	15	27,2727281	0,30112923
estante	0,12572038	13	23,6363634	0,3174404
estuche	0,12306412	13	23,6363634	0,33375157
data (proyector)	0,11897144	15	27,2727281	0,35257215
escritorio	0,11690986	10	18,1818187	0,3651192

Cuadro 6. Vocablos con IDL \geq 0,1 en el CI *La escuela: muebles y materiales*

El centro de interés que tuvo más vocablos con IDL \geq 0,1 fue *La escuela: muebles y materiales*, con 12 casos, siendo el más disponible, *mesa* (0,63) seguido por *silla*, *pizarra*, *plumón*, *lápiz (de color)*, *cuaderno*, *libro*, *computador*, *estante*, *estuche*, *data (proyector)* y *escritorio*. Entre los dos primeros vocablos (*mesa* y *silla*) no hubo gran oscilación de frecuencia y porcentajes, pero sí puede verse una variación en la oscilación entre el resto de los lemas más disponibles. Respecto a los porcentajes de aparición, en este actualizador, los cinco vocablos más disponibles tuvieron valores por encima de 50 %, lo que contrastó con los otros dos ejes temáticos; el lema más disponible, *mesa*, tuvo un porcentaje de 72,72 %, el segundo, *silla*, 74,54 %, el tercero, *pizarra*, 67,27 %, el cuarto, *plumón* (60 %) y, el quinto, *lápiz (de color)*, 50,9 %.

Todos los vocablos con IDL \geq 0,1 de este actualizador fueron sustantivos, que pueden organizarse semánticamente en dos grupos, justamente en los nombres del área nocional: lexemas que refieren a ‘muebles de la escuela’: *mesa*, *silla*, *pizarra*, *estante* y *escritorio*; y unidades léxica que aluden a ‘materiales de la escuela’: *plumón*, *lápiz (de color)*, *cuaderno*, *libro*, *computador*, *estuche* y *data (proyector)*.

6. DISCUSIÓN

En esta investigación nos hemos centrado en describir el léxico disponible de un grupo de futuros docentes chilenos, en tres centros de interés vinculados al área de la pedagogía. Del análisis, puede destacarse que las áreas nocionales que tuvieron mayores índices generales de disponibilidad léxica (TP, TV, PP¹⁰ e IC) fueron *La escuela* y *La educación*. Es decir, los actualizadores que se relacionan más con la formación de los encuestados.

Al igual que en Herranz (2018), quien analizó ejes temáticos similares en estudiantes universitarios españoles de carreras de docencia, observamos que en ambos trabajos el área nocional *La escuela* tiene mayor número de palabras, mientras que el actualizador *La educación* tiene el mayor número de vocablos. A grandes rasgos, podemos señalar que el hecho de que los encuestados de los dos estudios hayan producido mayor número de palabras diferentes en el centro de interés *La educación* puede deberse quizá al contacto y uso de términos del área de la docencia que tienen los informantes mediante sus lecturas, clases y actividades curriculares universitarias. En el Cuadro 7, a continuación, ilustramos estos datos.

	Este trabajo		Herranz (2018)	
	TP	TV	TP	TV
La educación	766	405	10858	1675
La escuela	797	283	12913	839

Cuadro 7. Comparación de los índices TP y TV de este estudio con el de Herranz 2018

Respecto a la variable año de escolaridad en el presente estudio, los resultados parecen indicar que sí existen diferencias cuantitativas entre el léxico disponible de los informantes de primer año y cuarto año, siendo los del último año los que tienen una mayor riqueza léxica *versus* los del primer año. Además, la prueba paramétrica t de Student arrojó que hay un contraste significativo entre las igualdades de medias de ambos grupos.

¹⁰ El promedio de palabras es representado por el signo \bar{x} .

En relación con el promedio de palabras del grupo chilenos, los estudiantes de cuarto año tuvieron mayores valores que los del primer año –44,70 y 36,91 respectivamente–, una diferencia de 7,79 palabras entre ambos grupos. Sobre el X de producción de palabras diferentes, el grupo de cuarto año tuvo 26,48 vocablos por sujeto, mientras que los del primer año, 18,31; una diferencia de 8,17 lemas. Al detallar los vocablos, los encuestados de cuarto año produjeron más unidades de cita que los del primer año, especialmente en el actualizador *La educación*. Además, debemos destacar que, en los tres ejes temáticos, este índice aumentó en el grupo de cuarto año, siendo, justamente, el CI *La educación* el que tuvo el incremento más alto: de 11,88 (primer año) a 16,78 (cuarto año).

La tendencia antes mencionada es comparable con la de la investigación de Herranz (2018), especialmente sobre el eje temático *La educación*. Pues, los estudiantes españoles de cuarto año tuvieron mayor número de vocablos que los del primer año, 843 y 832, respectivamente. Igualmente, se observó un aumento en el promedio de palabras y vocablos entre ambos grupos: de 17,7 palabras y 10,5 vocablos (primer año) a 19,1 palabras y 11,1 vocablos (cuarto año). Respecto al CI *La escuela: muebles y materiales escolares*, también pudo apreciarse un aumento del número de vocablos, mas no de palabras: de 410 lemas (primer año) a 429 lemas (cuarto año). Con base en estos datos, Herranz (2018: 156) afirma que “sí existe un cambio en el vocabulario” de los estudiantes, sobre todo en campos semánticos vinculados con su formación académica.

Acerca del índice de cohesión, los resultados del presente trabajo mostraron que a mayor nivel educativo mayor el índice de cohesión, lo que significa que los ejes temáticos fueron más compactos entre los estudiantes de cuarto año, siendo el actualizador *La educación* el que tuvo el mayor aumento de este índice, pues pasó de 0,05 (primer año) a 0,068 (cuarto año). Esto sugiere que los alumnos del último año de la carrera utilizan un vocabulario más acotado y específico cuando se refieren a tópicos de su carrera, al contrario de los universitarios de reciente ingreso, quienes aún podrían tener una concepción general acerca de “la educación”, lo que podría incidir en la utilización de un vocabulario más difuso. En líneas generales, estos resultados parecen apoyar nuestra hipótesis de que a medida que los estudiantes avanzan en sus estudios, aumenta también su componente léxico, especialmente, en aquellos tópicos vinculados con sus áreas de formación.

Sobre la incidencia de la cantidad de lectura optativa en el caudal léxico de un grupo de futuros profesores chilenos en tópicos de pedagogía, los índices generales de DL han mostrado algunas diferencias entre los estudiantes que afirmaron leer más y los que señalaron que leen menos. Si bien la prueba estadística paramétrica Anova arrojó que no existen diferencias significativas

entre las categorías analizadas, a grandes rasgos, la tendencia ha mostrado que los alumnos del corpus que dedican más horas semanales a realizar lecturas optativas tienen una mayor riqueza léxica que los que leen menos o nada.

Entre los resultados pudo verse que el promedio de palabras en cada centro de interés se incrementaba a medida que aumentaba la cantidad de horas de lectura. El mismo patrón se apreció en los índices de cohesión: los estudiantes que indicaron leer más de 10 horas semanales textos no obligatorios tuvieron índices más altos que aquellos sujetos que afirmaron no dedicar ninguna hora semanal a realizar lecturas optativas. Los resultados respecto al \bar{X} e IC permiten apoyar nuestra hipótesis de que los estudiantes que dedican más tiempo semanal a leer textos no obligatorios tendrán mayores índices de disponibilidad léxica que aquellos que dedican poco o ningún tiempo semanal a la lectura opcional.

En relación con las palabras más disponibles, en primer lugar, pudimos observar que el actualizador *La escuela* fue el que tuvo más vocablos con $IDL \geq 0,1$, a saber: 12 lexemas, todos ellos sustantivos organizados en dos grupos semánticos referidos justamente al nombre del CI ‘muebles’ y ‘materiales’ de la escuela. El hecho de que este eje temático haya tenido más vocablos puede deberse a que es el más concreto de los tres analizados en este trabajo. Además, alude a una realidad a la que los informantes han estado vinculados desde niños. Las tres palabras más disponibles de este centro de interés fueron: *mesa*, *silla* y *pizarra*, que son también las tres más disponibles del trabajo de Herranz (2018).

El eje temático *La educación* fue el segundo con más vocablos con $IDL \geq 0,1$ (8 lemas). De este grupo de lexemas debemos resaltar el más disponible: *aprendizaje* (*significativo*), ya que se distingue entre *aprendizaje*, como palabra de uso general utilizada por un hablante de cualquier nivel educativo; y *aprendizaje significativo*, que es una unidad léxica especializada en la pedagogía y, por lo tanto, perteneciente a la jerga de los educadores. Esto nos hace inferir que esta unidad léxica se adquiere, comprende y usa por quienes han tenido acceso a textos científicos especializados del área de la docencia. En este sentido, puede decirse que la presencia de esta pieza léxica como la más disponible en este CI es quizá un indicador de la posible adquisición e incorporación al léxico de términos especializados de la carrera de los encuestados. En comparación con el trabajo de Herranz (2018), en ambos estudios los vocablos *profesor* y *colegio* aparecen entre los más disponibles. Igualmente, comparten términos afines, en el presente se reporta *estudiante*, mientras que entre los madrileños aparece *alumnado*, lo cual podría deberse a razones de divergencia dialectal.

Sobre el actualizador *La lectura*, este tuvo apenas 7 vocablos con $IDL \geq 0,1$, siendo *libro* el más disponible. De este grupo de lexemas, hay que resaltar el vocablo *texto (científico)*, puesto que, por una parte, se hace la distinción entre *texto* a secas, que es el hiperónimo en el que pueden agruparse todos los textos, sea cual sea su género literario y, por lo tanto, el término utilizado por las personas independientemente de su formación académica para identificar cualquier manuscrito. Por otra parte, se diferencia *texto científico*, refiriéndose ya a los textos especializados. Este término alude a un tipo específico de género literario (artículos, capítulos de libros, tesis, ensayos técnicos especializados) que es familiar entre los encuestados, quienes tienen que recurrir a la lectura de este tipo de documento en sus asignaturas, por lo que podría sugerirse que esta pieza léxica evidencia el nivel y el contexto educativo de los sujetos.

Por último, debemos señalar que el área nocional *La lectura* fue la que resultó con los índices generales y específicos más bajo. Si bien se espera que un docente sea un lector nato, que motive prácticas lectoras entre sus estudiantes (Larrañaga *et al.* 2008; Parrado *et al.* 2018; Juárez 2019), estos resultados parecen reflejar los pocos hábitos lectores de los futuros profesores, al menos en aspectos estadísticos del léxico. No obstante, habría que ahondar más en el análisis de este campo temático y las prácticas lectoras de los universitarios.

7. CONCLUSIÓN

En este trabajo descriptivo, cuantitativo, enmarcado en los estudios de disponibilidad léxica, planteamos dos objetivos generales: por un lado, describir las características estadísticas del LD de un grupo de estudiantes universitarios de la carrera de pedagogía de una universidad estatal chilena del CRUCH, y, por otro lado, describir la incidencia de los años de escolaridad y frecuencia de lectura (cantidad de horas semanales de lectura optativa) en el léxico disponible de los encuestados. Para esto hemos analizado un corpus compuesto por las listas de palabras de 55 sujetos, sobre tres centros de interés: *La lectura*, *La educación* y *La escuela: muebles y materiales*.

Los resultados generales han mostrado que de los tres CI el de mayor número de palabras y, por ende, el de rango 1 fue *La escuela* (797), seguido por *La educación* (766) y, por último, *La lectura* (646). Sin embargo, el CI *La educación* (405) tuvo el mayor número de vocablos, muy por encima del promedio de palabras diferentes (324 , 66).

En relación con el contraste de los índices de disponibilidad léxica según el nivel educativo de los informantes, los resultados muestran que hubo un incremento en el caudal léxico de los estudiantes de último año en comparación a los de primer año. Esto nos permite afirmar que, respecto al corpus aquí analizado, existen diferencias cuantitativas entre el caudal léxico de los estudiantes de primer año y los de cuarto año.

Igualmente, pudo determinarse que la cantidad de horas semanales que los sujetos del corpus analizado dedican a realizar lecturas no obligatorias parece incidir en el caudal léxico, pues, a mayor cantidad de horas de lectura, mayor es el promedio de palabras e índice de cohesión en cada tópico tratado.

Respecto a las palabras más disponibles, aquellas que tuvieron un $IDL \geq 0,1$, pudo apreciarse que hubo tanto palabras de conocimiento general (*libro, leer, colegio, profesor, mesa, silla*) como un par de léximas con rasgos léxico-semánticos especializados de la vida universitaria y la docencia, a saber: *texto (científico)*, en el CI *La lectura*; y *aprendizaje (significativo)*, en el CI *La educación*.

Con este estudio hemos querido aportar datos acerca del caudal léxico de un grupo de estudiantes chilenos que se forman para ser maestros, y conocer la relación del léxico disponible con los años de escolaridad y la cantidad de lectura realizada. Con base en los análisis, recomendamos desarrollar una investigación similar en un corpus más amplio, con métodos no solo cuantitativos sino también cualitativos. En la que, además, se tomen en cuenta otros factores vinculados con las prácticas lectoras. Por último, se recomienda desarrollar un escrutinio cualitativo detallado sobre las piezas léxicas más disponibles, en el que se consideren también análisis de clúster y redes semánticas.

REFERENCIAS BIBLIOGRÁFICAS

- AABIDI, LAHOUSINE. 2020. La disponibilidad léxica en español de alumnos marroquíes de enseñanza media: resultados generales. *Philologica Canariensia* 26: 1-19.
- ÁVILA MUÑOZ, ANTONIO. 2007. Léxico disponible y ortografía. condicionantes sociales y hábitos culturales de influencia. En Juan Moya y Marcin Sosinsky (eds.). *Las hablas andaluzas y la enseñanza de la lengua. Actas de las XII Jornadas sobre la Enseñanza de la Lengua Española*, pp. 25-46. Granada: Universidad de Granada.
- _____. 2016. El léxico disponible y la enseñanza del español. Propuesta de selección léxica basada en la teoría de los conjuntos difusos. *Journal of Spanish Language Teaching* 3(1): 31-43.

- ÁVILA MUÑOZ, ANTONIO Y MARÍA LASARTE. 2010. Estudio del conocimiento social del léxico disponible. Objetivos, metodología y criterios de edición del diccionario. En Antonio Ávila Muñoz y Juan Villena (eds.), pp. 83-112.
- ÁVILA MUÑOZ, ANTONIO Y JOSÉ SÁNCHEZ. 2011. La posición de los vocablos en el cálculo del índice de disponibilidad léxica: Proceso de reentrada en las listas del léxico disponible en la ciudad de Málaga. *ELUA* 25: 45-74.
- ÁVILA MUÑOZ, ANTONIO, INMACULADA SANTOS DÍAZ Y ESTER TRIGO IBÁÑEZ. 2020. Análisis léxico-cognitivo de la influencia de los medios de comunicación en las percepciones de universitarios españoles ante la COVID-19. *Círculo De Lingüística Aplicada a La Comunicación* 84: 85-95.
- ÁVILA MUÑOZ, ANTONIO Y JUAN VILLENA (eds.). 2010. *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis*. Málaga: Editorial Sarriá.
- BLANCO, OSCAR, PEDRO SALCEDO Y GABRIELA KOTZ. 2020. Análisis del léxico de las emociones: una aproximación desde la disponibilidad léxica y la teoría de los grafos léxicos. *Lingüística y Literatura* 78: 56-84.
- CASANOVA, MANUELA. 2017. *Léxico disponible de Castellón. Estudio y diccionarios*. Tesis para optar al grado de Doctora, Universitat Jaume I.
- CASTILLO FADIĆ, MARÍA NATALIA Y ENRIQUE SOLOGUREN. 2020. Léxico frecuente, riqueza léxica y estereotipos sobre la lectura de profesores en formación. *Logos* 30(1): 69-85.
- CASTILLO FADIĆ, MARÍA NATAL Y JOSUÉ PINO CASTILLO. 2020. Hacia la construcción de un instrumento para evaluar la familiaridad de pacientes crónicos con unidades léxicas relevantes para el automanejo de su condición de salud. *Nueva Revista del Pacífico* 72: 86-115.
- CASTILLO FADIĆ, MARÍA NATAL E INMACULADA SANTOS DÍAZ. 2021. Dos miradas sobre el cuerpo humano en el marco de la literacidad en salud: profesionales de la salud y personas con diabetes mellitus 2. En Maribel Serrano Zapata y M.^a Ángeles Carrero Fernández (eds.), *Aplicaciones de la disponibilidad léxica: 277-293*. Universitat de Lleida, Lleida, España.
- CEPEDA, MILKO, MARIBEL GRANADA Y MARÍA POMÉS. 2014. Disponibilidad léxica en estudiantes de primero básico. *Literatura y Lingüística* 30: 166-181.
- CEPEDA, MILKO, ÁNGELA Cárdenas, MACARENA CARRASCO, NICOLE CASTILLO, JOSELYNE FLORES, CONSTANZA GONZÁLEZ Y MELANIE ORÓSTICA. 2016. Relación entre disponibilidad léxica y comprensión lectora, en un contexto de educación técnico profesional rural. *Sophia Austral* 18: 81-93.
- CERDA, GAMAL, CARLOS PÉREZ Y EUGENIO CHANDÍA. 2021. Disponibilidad léxica en centros de interés asociados a ejes curriculares de matemáticas en estudiantes de contextos de alta vulnerabilidad social. *Revista de Estudios y Experiencias en Educación* 20(42): 33-51. [En línea] disponible en <http://www.rexe.cl/ojournal/index.php/rexe/article/view/1012/641>
- DISPOLEX. 2003-2021. ¿Qué es el proyecto panhispánico? [En línea] disponible en <http://www.dispox.com/info/el-proyecto-panhispanico> [Consulta 26/09/2020].
- DOMÍNGUEZ PELEGRÍN, JUAN. 2019. La incidencia de los estilos de aprendizaje en el índice de disponibilidad léxica individual de estudiantes extranjeros de español. *Verba Hispanica* XXVII: 93-113.
- ECHEVERRÍA, MAX, MARÍA HERRERA, PATRICIO MORENO Y FRANCISCO PRADENAS. 1987. Disponibilidad léxica en Educación Media: resultados cuantitativos. *Revista de Lingüística Teórica y Aplicada* 25: 55-115.
- ECHEVERRÍA, MAX, PAULA URZÚA e ISRAEL FIGUEROA. 2005. *Dispogen II. Programa computacional para el análisis de la disponibilidad léxica*. Concepción, Chile: Universidad de Concepción.
- FERREIRA, ROBERTO, JAIME GARRIDO Y ALEXIA GUERRA. 2019. Predictors of lexical availability in English as a second language. *Onomázein* 46: 18-34.

- GARCÍA, FRANCISCO. 2011. La creatividad léxica a través de los recursos morfológicos del léxico disponible de español de Galicia. En Belén López Meirama (ed.) *Estudios sobre disponibilidad léxica en el español de Galicia*, pp. 217-266. Santiago de Compostela: Universidad de Santiago de Compostela.
- GÓMEZ DEVÍS, M.^a BEGOÑA. 2019. A propósito de las redes semánticas en el léxico disponible de escolares de primero de educación primaria. *Ogigia* 25: 165-183.
- GÓMEZ MOLINA, JOSÉ RAMÓN y M.^a BEGOÑA GÓMEZ DEVÍS. 2004. *La disponibilidad léxica de los estudiantes preuniversitarios valencianos. Estudio de estratificación sociolingüística*. Valencia: Universitat de València.
- GOUGENHEIM, GEORGES, RENÉ MICHÉA, PAUL RIVENC y AURÉLIEN SAUVAGEOT. 1954. *L'élaboration du français élémentaire*. Paris: Didier.
- HERNÁNDEZ MUÑOZ, NATIVIDAD. 2009. Aspectos sociolectales del léxico dialectal. *Spanish in Context* 6(2): 224-248.
- HERNÁNDEZ MUÑOZ, NATIVIDAD y CRISTINA IZURA. 2010. ¿Perro o musarafa? La recuperación léxica en las categorías semánticas. *Ciencias cognitivas* 4(1): 1-3. [En línea] Disponible en <http://www.cienciacognitiva.org/?p=84> [Consulta: 26/09/2020]
- HERNÁNDEZ MUÑOZ, NATIVIDAD y CARMELA TOMÉ. 2017. Léxico disponible en primera y segunda lengua: bases cognitivas. En Florencio del Barro (ed.), pp. 99-122.
- HERNÁNDEZ MUÑOZ, NATIVIDAD, CRISTINA IZURA y CARMELA TOMÉ. 2014. Cognitive Factors of Lexical Availability in a Second Language. En Rosa Jiménez Catalán (ed.) *Lexical Availability in English and Spanish as a Second Language*, pp. 169-186. Netherlands: Springer
- HERRANZ, CRISTINA. 2018. Disponibilidad léxica de los futuros profesores de educación infantil y primaria. *Revista Electrónica Interuniversitaria de Formación del Profesorado* 21(1): 143-159. [En línea] Disponible en <https://revistas.um.es/reifop/article/view/295271> [Consulta 26/09/2020]
- _____. 2020. *Palabra de maestro. Análisis del léxico disponible de futuro docentes*. Berlín: Peter Lang.
- HERRANZ, CRISTINA y MIGUEL MARCOS. 2019. Análisis del léxico disponible español de extranjeros que estudian los grados de educación. *OGIGIA* 26: 5-30
- JIMÉNEZ, FELIPE. 2019. *Estudio sociolingüístico del léxico disponible de escolares navarros*. Pamplona: EUNSA.
- JUÁREZ, MARÍA. 2019. Influencia de la formación inicial del profesorado en los hábitos lectores y en el concepto de educación literaria. *Investigaciones Sobre Lectura* 12: 99-115.
- LARA, LUIS FERNANDO. 2006. *Curso de lexicología*. México: El Colegio de México.
- LARRAÑAGA, ELISA, SANTIAGO YUBERO y PEDRO CERRILLO. 2008. *Estudio sobre los hábitos lectores de los universitarios españoles*. Cuenca: CEPLI/UCLM.
- LIN, JIN. 2012. El estudio de disponibilidad léxica de los estudiantes chinos de español como lengua extranjera. *MarcoELE* 14 [En línea] Disponible en https://marcoele.com/descargas/14/lin-disponibilidad_lexica.pdf [Consulta 26/09/2020]
- LLINARES, GENOVEBA, BERTHA LEIVA, NOELA CARTAYA y RUELA ST. LUIS. 2008. Acquisition of L2 Vocabulary for Effective Reading: Testing Teachers' Classroom Practice. *The Reading Matrix* 8(2): 55-69 [En línea] Disponible en <http://www.readingmatrix.com/articles/llinares-leiva-cartaya-stlouis/article.pdf> [Consulta 26/09/2020]
- LÓPEZ CHAVES, JUAN y CARLOS STRASSBURGER. 1987. Otro cálculo del índice de disponibilidad léxica. En *Presente, y perspectiva de la investigación computacional en México. Actas del IV Simposio de la Asociación Mexicana de Lingüística Aplicada*, pp. 20-31. México: UNAM.
- _____. 2001. El diseño de una fórmula matemática para obtener un índice de disponibilidad léxica confiable. *Anuario de Letras* 38: 227-251.

- LÓPEZ MORALES, HUMBERTO. 1995-1996. Los estudios de disponibilidad léxica: pasado y presente. *Boletín de Filología* XXXV: 245-259.
- _____. 1999. *Léxico disponible de Puerto Rico*. Barcelona: Arco/Libro.
- LÓPEZ MORALES, HUMBERTO y ESTER TRIGO Ibáñez. 2019. La disponibilidad léxica: tendencias actuales y perspectiva de futuro. Editorial. *Ogigia* 25: 7-10.
- MAHECHA, VIVIANA y GERAL MATEUS FERRO. 2017. El léxico disponible y sus mecanismos de asociación: un análisis con grafos. En Florencio del Barro (ed.), pp. 123-141.
- MATEUS FERRO, GERAL, LAURA CASTIBLANCO y PEDRO ÁLVAREZ-BERMÚDEZ. 2018. Mecanismos lógicos y analógicos en la producción del léxico disponible. *FOLIOS* 47: 133-152.
- MENDOZA PUERTAS, JORGE. 2018. El léxico disponible de 82 estudiantes coreanos de español como lengua extranjera. *MarcoEle* 26 [En línea] Disponible en https://marcoele.com/descargas/26/mendoza-lexico_disponible-coreanos.pdf [Consulta 26/09/2020]
- MESA CANALES, ROSA. 1989. *Disponibilidad léxica en preescolares*. México, D.F.: Universidad Autónoma de México.
- MICHÉA, RENÉ. 1953. Mots fréquents et mots disponibles. un aspect nouveau de la statistique du langage. *Les Langues Modernes* 47: 338-344.
- PARRADO, MILAGROSA, MANUEL ROMERO y ESTER TRIGO IBÁÑEZ. 2018. La experiencia literaria en la formación de futuros docentes: el viaje iniciático de nuestras biografías lectoras en 10 hashtag. En Víctor Amar. (coord.) *Miradas y voces de futuros maestros*, pp. 57-84. Barcelona: Octaedro.
- PAREDES, FLORENTINO. 2012a. Desarrollos teóricos y metodológicos recientes de los estudios de disponibilidad léxica. *Revista Nebrija de Lingüística Aplicada* 6/4 [En línea] Disponible en <https://revistas.nebrija.com/revista-linguistica/article/view/177/147> [Consulta 27/09/2020]
- _____. 2012b. A vueltas con la selección de ‘centros de interés’ en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE. *Revista Nebrija de Lingüística Aplicada* 16 [En línea] Disponible en <https://www.nebrija.com/revista-linguistica/a-vueltas-con-la-seleccion-de-centros-de-interes-en-los-estudios-de-disponibilidad-lexica-para-una-propuesta-renovadora-a-proposito-de-la-disponibilidad-lexica-en-ele.html> [Consulta 26/09/2020]
- _____. 2018. Neologismos y creaciones léxicas en las pruebas de disponibilidad léxica de la comunidad de Madrid. En Lynn Williams (coord.) *Estudios de lengua y lingüística española. Homenaje a Orlando Alba*, pp. 125-150. Berlín: Peter Lang.
- PAREDES, FLORENTINO y DIEGO GALLEGU. 2019. Procedimientos neológicos en el léxico disponible de español como lengua materna y como lengua extranjera. *OGIGIA* 25: 109-138.
- PERDOMO, NORAI DA y ROSA ALFONSO. 2021. La resemantización de vocablos en estudiantes de bachillerato: una mirada desde los estudios de disponibilidad léxica. *Amauta* 19(37) [En línea] disponible en <http://investigaciones.uniatlantico.edu.co/revistas/index.php/Amauta/article/view/2885>
- RAE-ASALE = REAL ACADEMIA ESPAÑOLA y ASOCIACIÓN DE ACADEMIAS DE LA LENGUA. 2005. *Diccionario Panhispánico de Dudas*. Madrid: Espasa.
- _____. 2014. *Diccionario de la Lengua Española*. Madrid: Espasa.
- REYES, DALIA, JOSÉ REYES, FLORES TREVIÑO MARÍA y RINA OJEDA. 2021. Una propuesta de herramientas informáticas para el tratamiento estadístico del índice de disponibilidad léxica en estudios correlacionales de educación y movilidad social. *Forma y función* 34(1) [En línea] disponible en <https://revistas.unal.edu.co/index.php/formayfuncion/article/view/80581>
- ROJAS, DARÍO, CAROLINA ZAMBRANO y PEDRO SALCEDO. 2017. Metodología de análisis de disponibilidad léxica en alumnos de pedagogía a través de la comparación jerárquica de lexicones. *Formación Universitaria* 10(4): 3-14.

- SALCEDO, PEDRO y MARÍA DEL VALLE LEO. 2013. Disponibilidad léxica matemática en estudiantes de enseñanza media de Concepción, Chile. *Revista Científico Pedagógica Atenas* 4, 21: 1-16.
- SAMPER PADILLA, JOSÉ ANTONIO. 1998. Criterios de edición del léxico disponible: sugerencias. *Lingüística* 10: 311-333.
- SÁNCHEZ-SAUS, MARTA. 2019. *Centros de interés y capacidad asociativa de las palabras*. Sevilla: EUS.
- SANTOS DÍAZ, INMACULADA. 2017a. Incidencia de la lectura en el vocabulario en lengua materna y extranjera. *Ocnos Revista de Estudios sobre Lectura* 16(1): 79-88.
- _____. 2017b. Selección del léxico disponible: propuesta metodológica con fines didácticos. *Porta Linguarum* 27: 122-139.
- _____. 2020. *El léxico bilingüe del futuro profesorado. Análisis y pautas para estudios de disponibilidad léxica*. Berlín: Peter Lang.
- SANTOS DÍAZ, INMACULADA, ESTER TRIGO y MANUEL ROMERO. 2020. La activación del léxico disponible y su aplicación a la enseñanza de lenguas. *Porta Linguarum* 33: 75-93.
- SANTOS PALMOU, XANDRA. 2017. Métodos de selección léxica aplicados a la enseñanza. El vocabulario fundamental del español. Tesis para optar al grado de Doctora, Universidad de Salamanca.
- SEFARTI, MOHAMED. 2016. La disponibilidad léxica de estudiantes marroquíes de nivel universitarios: resultados generales. *Philologica Canariensia* 22: 105-116.
- _____. 2021. La evolución del léxico disponible en español como lengua extranjera en aprendices marroquíes de enseñanza secundaria y universitaria. *Lingüística y Literatura* 79: 164-180.
- TRIGO IBÁÑEZ, ESTER. 2011. *Dialectología y cultura. El léxico disponible de los preuniversitarios sevillanos*. Valencia: Advana vieja.
- TRIGO IBÁÑEZ, ESTER, MANUEL ROMERO OLIVA e INMACULADA SANTOS DÍAZ. 2019. Aproximación al léxico gastronómico dialectal andaluz desde los repertorios de disponibilidad léxica para una propuesta didáctica. *Verba Hispánica* XXVII: 115-130
- TRIGO IBÁÑEZ, ESTER, INMACULADA SANTOS DÍAZ y MANUEL ROMERO OLIVA. 2020. Disponibilidad léxica y dominio de la ortografía: un estudio empírico basado en la influencia de los factores sociales. *Onomázein* 47: 27-47.
- URZÚA, PAULA. 2018. *Disponibilidad léxica en adultas mayores de la ciudad de Concepción: Estudio descriptivo*. Concepción: Universidad de Concepción.
- VALENZUELA, MARCO, VICTORIA PÉREZ, CLAUDIO BUSTO y PEDRO SALCEDO. 2018. Cambios en el concepto aprendizaje de estudiantes de pedagogía: análisis de disponibilidad léxica y grafos. *Estudios Filológicos* 61: 143-173.
- ZAMBRANO, CAROLINA. 2021. Un estudio de la disponibilidad léxica en el ámbito de la autorregulación del aprendizaje en la formación inicial docente. *Lingüística y Literatura* 42(79): 11-33.