

## COMPETENCIA LÉXICA, COMPRENSIÓN LECTORA Y RENDIMIENTO ACADÉMICO EN ESTUDIANTES DE ENSEÑANZA MEDIA\*

Bernardo Riffo Ocares\*\*  
Fernando Reyes Reyes\*\*\*  
Abraham Novoa Lagos\*\*\*\*  
Mónica Véliz de Vos\*\*\*\*\*  
Ginette Castro Yáñez\*\*\*\*\*

### Resumen

Este artículo presenta los resultados de un estudio correlacional en el que se midieron las variables comprensión lectora, competencia léxica y rendimiento académico. Se postula que la calidad del conocimiento léxico de los lectores está relacionada con su capacidad para comprender textos y, por ende, afecta su rendimiento académico. Se realizó un estudio con 32 estudiantes del último año de enseñanza media. Los sujetos rindieron una prueba de comprensión lectora, un test de vocabulario pasivo (TEVI-R) y una prueba de decisión léxica con técnica cronométrica. Los resultados muestran correlación entre la precisión de la respuesta en la tarea léxica y la comprensión lectora. El vocabulario pasivo, a su vez, se correlaciona con el rendimiento académico general.

**Palabras clave:** Comprensión lectora, competencia léxica, rendimiento académico.

## LEXICAL COMPETENCE, READING COMPREHENSION AND ACADEMIC PERFORMANCE IN HIGH SCHOOL STUDENTS

### Abstract

This article presents the results of a correlation study in which reading comprehension, lexical competence and academic performance were measured. We postulate, that the quality of readers' lexical knowledge is related to their ability to understand texts and, therefore, it affects their academic performance. A study was conducted with 32 high school last grade individuals. The subjects yielded a reading comprehension test, a passive vocabulary test (TEVI-R) and a lexical decision test with chronometric technique. The results indicate that the quality of the response in lexical decision test correlates with performance on reading comprehension, whereas passive vocabulary correlates with general academic performance.

**Keywords:** Reading comprehension, lexical competence, school achievement.

Recibido: 29-11-2013

Aceptado: 01-06-2014

---

\* La presente investigación fue desarrollada en el marco del proyecto DINV 209.062.043-1FI, de la Vicerrectoría de Investigación y Desarrollo de la Universidad de Concepción y con el apoyo de Conicyt (Gobierno de Chile) a través del proyecto Fondef D0811179.

\*\* Dr. en Lingüística, Universidad de Concepción, Concepción, Chile. [bernardo.riffo@udec.cl](mailto:bernardo.riffo@udec.cl)

\*\*\* Dr. (c) en Psicología, Universidad del Desarrollo, Concepción, Chile. [freyes@udd.cl](mailto:freyes@udd.cl)

\*\*\*\* Mg. en Lingüística, Universidad de Concepción, Concepción, Chile. [abraham.novoa8@gmail.com](mailto:abraham.novoa8@gmail.com)

\*\*\*\*\* Mg. en Lingüística, Universidad de Concepción, Concepción, Chile. [mveliz@udec.cl](mailto:mveliz@udec.cl)

\*\*\*\*\*Dr. en Lingüística, Universidad Católica de Temuco, Temuco, Chile. [gcastro@uct.cl](mailto:gcastro@uct.cl)

## Introducción

La lectura comprensiva constituye un proceso complejo que involucra, junto con el sistema lingüístico, el despliegue de otros códigos de representación. En efecto, desde la percepción de las letras hasta la interpretación de los sentidos del texto, tiene lugar un conjunto de operaciones que implican el uso del conocimiento de la lengua, del ámbito de referencia del discurso y de la situación comunicativa en la que se enmarcan el texto y el lector. Dado su carácter multidimensional, el fenómeno concita el interés de diversas disciplinas y se convierte en objeto para los estudios del discurso, la lingüística y la psicología, entre otras. Desde la perspectiva psicolingüística, en particular, la comprensión lectora ha devenido un punto central de la investigación, tanto así que para Kintsch (1998) la comprensión constituye un paradigma para explicar la cognición humana.

De acuerdo con el *modelo de construcción-integración* (C-I), la comprensión de textos implica el procesamiento de información en diversos niveles (Kintsch, 1998). En primer lugar, en el nivel lingüístico se llevan a cabo las operaciones vinculadas con el reconocimiento de palabras y la comprensión de oraciones presentes en la superficie textual. El lector debe dar cuenta de los símbolos gráficos de la página (independiente de su soporte físico, como el papel o la pantalla de un computador), lo que supone procesos perceptivos en los que participa la memoria sensorial (Smith & Kosslyn, 2008). En segundo lugar, en el nivel semántico se efectúa el análisis orientado a establecer el significado de las expresiones de superficie del nivel anterior. El significado de las unidades elementales (palabras) se integra en estructuras de significado llamadas *proposiciones* (entendidas aquí como *unidades de ideas*). Estas proposiciones, a su vez, se incorporan en una red más compleja llamada *base de texto*, formada por secuencias de proposiciones. En términos psicológicos, estas proposiciones se construyen a partir de los significados de las unidades léxicas y su papel en las estructuras oracionales implicadas así como de las relaciones semánticas que entre ellas se establecen, aunque mayoritariamente dichos vínculos no se explicitan. Fundamental para establecer los vínculos entre proposiciones y alcanzar así una representación coherente es la capacidad del lector de realizar inferencias que conectan las distintas partes del texto integrándolas en unidades de mayor alcance. Las *microestructuras* se integran luego en

representaciones más amplias llamadas *macroestructuras*. En conjunto, *microestructura* y *macroestructura* conforman la *base de texto*. Esta, no obstante, aunque representa el significado expresado efectivamente en el texto, no es suficiente para alcanzar la comprensión cabal. Se requiere, en tercer lugar, una representación más profunda vinculada con los hechos denotados por el texto, con la situación referida. A este nivel se le conoce como *modelo de situación*, equivalente en gran medida a lo que Johnson-Laird (1983) llama *modelo mental*. Dicho modelo integra información disponible en el texto con otra proveniente de distintos dominios como la imaginación, las emociones y el conocimiento general del mundo del lector (Kintsch & Rawson, 2007).

La realización de inferencias y las estrategias empleadas por el lector han concentrado buena parte de la atención en los estudios de la comprensión del discurso (Magliano, Wiemer-Hastings, Millis, Muñoz & McNamara, 2002; McNamara & Kintsch, 1996). El procesamiento del léxico, en cambio, se ha considerado principalmente como un nivel relativamente autónomo, en función de lo cual han surgido numerosas propuestas teóricas y un gran volumen de estudios empíricos específicos para dar cuenta de este aspecto en forma separada de su lugar en el discurso. Contrasta con esto, sin embargo, la idea de la palabra como unidad básica de procesamiento del discurso, postulada tempranamente por Just & Carpenter (1980). En esta línea, más tarde, Perfetti desarrolló una interesante propuesta llamada *teoría de la eficiencia verbal* (Perfetti, 1985), y después *hipótesis de la calidad léxica* (Perfetti, 2010; Perfetti, Landi & Oakhill, 2007; Perfetti & Hart, 2001).

De acuerdo con Perfetti (2010), existe una estrecha relación de interdependencia entre la decodificación, el vocabulario y la comprensión. La decodificación, entendida como el conjunto de operaciones implicadas en la percepción de la señal gráfica para convertirla en una representación de orden fonológico, afecta al vocabulario. Si la decodificación tiene éxito, se fortalecen, por una parte, las conexiones forma-significado y, por otra, se establecen vínculos de dependencia contextual para las palabras no familiares o desconocidas. El vocabulario, que implica un conocimiento del significado de las palabras, a su vez, afecta la decodificación, toda vez que, al decodificar una palabra de significado conocido, se fortalece la relación entre significado y forma ortográfica. Por su parte, la relación vocabulario-comprensión se basa en que esta última depende del conocimiento que el lector tenga del significado de las palabras que

están siendo leídas en un determinado momento. Al mismo tiempo, si se logra la comprensión de un fragmento del texto en el que aparecen palabras desconocidas, se produce como consecuencia un aprendizaje del significado de dichas unidades léxicas. Finalmente, decodificación y comprensión no se relacionan en forma directa, sino mediada por el vocabulario: la comprensión del texto provee suficiente significado como para asociar las formas léxicas con algún contenido semántico en forma más eficiente, especialmente para aquellos lectores más desfavorecidos (Perfetti, 1985; 2010; Perfetti et al., 2007). Al lograrse una decodificación más fluida y automatizada, se liberan recursos de memoria para el procesamiento y la comprensión.

Varios estudios, tanto de tipo evolutivo como experimental, proveen abundante evidencia a favor de las propuestas de Perfetti. El autor mismo señala que trabajos como el de Muter, Hulme, Snowling & Stevenson (2004) y el de Oakhill, Cain & Bryant (2003), entre otros, muestran resultados que permiten considerar las habilidades léxicas como un predictor del crecimiento en materia de comprensión lectora en niños de educación básica. Durante el proceso de alfabetización, mediante un mecanismo de auto-aprendizaje, los niños transitan desde una lectura completamente dependiente de la codificación fonológica de la palabra impresa a una en la que el acceso utiliza el código ortográfico, con lo que se hace más rápido y eficiente el procesamiento con un menor consumo de recursos (Share, 1999). En el estudio de Oakhill et al. (2003), en el que participaron escolares de 3° a 6° grado (rango etario: 7-11 años), se utilizaron pruebas para medir comprensión lectora, conocimiento de vocabulario, inteligencia verbal, memoria y conciencia fonológica. Los resultados fueron concluyentes en cuanto a establecer que el vocabulario se correlaciona significativamente con el desempeño en comprensión de textos; más aún, la variable léxica mostró capacidad de predecir el desarrollo futuro de los sujetos en lectura comprensiva. Estos hallazgos coinciden con nuestra experiencia durante el desarrollo de una prueba de comprensión lectora estandarizada (Riffo, Véliz, Castro, Reyes, Figueroa, Salazar & Herrera, 2011), donde utilizamos tareas de repetición de palabras y de discriminación en los niveles preescolar (4-5 años) y primer ciclo básico (6-9 años). El método de discriminación léxica empleado se basó en la tarea de decisión léxica (TDL), técnica ampliamente difundida en la psicolingüística experimental, que consiste básicamente en presentar a los sujetos series de letras, respecto de las cuales deben decidir si es

una palabra (e.g. POCO) o no (e.g. PONO). La modalidad empleada difería de la estándar en que se aplicó en forma impresa y no mediante un computador, por lo que no se midió la latencia de respuesta, sino solo el acierto/error. Los resultados de los análisis mostraron una correlación significativa entre la TDL y el desempeño en comprensión. En cambio, no hubo correlación entre la repetición de palabras y la comprensión.

Considerando los antecedentes expuestos, cabe preguntarse si lo constatado empíricamente en niños se da en forma equivalente en individuos mayores, específicamente en aquellos que se encuentran al final de la etapa escolar. Perfetti ha señalado que la calidad léxica está relacionada con las experiencias de lectura de los individuos y que existirían diferencias individuales asociadas a variables demográficas (Perfetti, 2010). En este sentido, nos proponemos: 1) determinar si la competencia léxica se correlaciona con la comprensión lectora y 2) determinar si la competencia léxica se relaciona con el rendimiento académico. Tales objetivos se apoyan en la hipótesis de que la competencia léxica se correlaciona con la comprensión lectora y el rendimiento académico.

## Método

### Diseño de investigación

El diseño de investigación general es de tipo correlacional; sin embargo, los datos de decisión léxica se obtuvieron mediante un experimento que utilizó un diseño con pre-prueba sin grupo control.

### Participantes

La muestra estuvo constituida por estudiantes de dos cursos de cuarto año de enseñanza media, pertenecientes a un colegio particular-subvencionado de la comuna de Concepción. La muestra total corresponde a 32 estudiantes, de los cuales 17 (53%) eran hombres y 15 (47%) eran mujeres, entre 17 y 18 años de edad ( $M = 17.6$ ,  $DE = .62$ ).

### Instrumentos

#### a) Prueba de Comprensión Lectora (LECTUM)

Instrumento desarrollado por Riffo et al. (2011) que evalúa, mediante ítems de selección múltiple, los aspectos textuales, pragmáticos

y críticos implicados en la comprensión lectora en estudiantes, quienes debían responder un total de 32 ítems correspondientes a cuatro diferentes textos. La prueba es una batería conformada por 7 niveles, cada uno de los cuales tiene una forma A y B, equivalentes en dificultad. Se usaron las dos formas de LECTUM 7, que es la prueba que corresponde al nivel más alto, es decir, para estudiantes de 3° y 4° año de enseñanza media (rango etario: 16-17 años). Cada ítem correcto se calificaba con un punto, y el puntaje global de la prueba es la sumatoria simple de las respuestas correctas.

b) Tarea de decisión léxica (TDL)

Se realizó una medición en condiciones experimentales con la técnica en su modalidad estándar (Perea & Rosa, 1999). En un computador a los participantes se les presentó un estímulo, en este caso una serie de letras, y debían decidir si ello correspondía a una palabra o no (palabras v/s pseudopalabras). Las pseudopalabras tienen similar constitución silábica a las palabras, pero no existen en la lengua. El computador registraba el tiempo de respuesta tanto en los aciertos como en los errores. Como un valor de frecuencia se utilizó el Índice de Disponibilidad Léxica de estudiantes de enseñanza media del registro de léxico disponible de Valencia y Echeverría (1999). Se contrabalanceó la lista de estímulos incluyendo equitativamente palabras de alta y baja disponibilidad (IDL alto, IDL bajo), con lo que se aseguró contar con palabras de alta y baja frecuencia. Toda la medición se hizo en un espacio controlado y la prueba se aplicó de forma individual.

c) Test de Vocabulario en Imágenes (TEVI-R)

El test de vocabulario en imágenes en su versión revisada, TEVI-R (Echeverría, Herrera & Sègure, 2002), permite determinar el nivel de comprensión del vocabulario pasivo que posee un sujeto entre 2, 6 y 17 años. El procedimiento de aplicación consiste en que el examinador entrega una palabra estímulo y el sujeto debe indicar en una lámina de cuatro imágenes la que corresponda a la palabra estímulo. El test puede aplicarse de manera grupal. Los puntajes se corrigieron según las normas disponibles para la población escolar.

## Procedimiento

A todos los participantes se les aplicó primero la prueba de comprensión lectora correspondiente al nivel del curso (LECTUM 7) de forma grupal en la sala de clases. En segundo lugar, se ejecutó la prueba experimental de procesamiento léxico (TDL), lo que se realizó en forma individual en un recinto distinto de la sala habitual de clases. Finalmente, se aplicó la prueba de vocabulario TEVI-R en forma individual en el mismo lugar que se llevó a efecto la prueba TDL. Para la TDL se utilizaron los procedimientos habituales: en la pantalla del computador aparecían palabras con valores de IDL variados (palabras de alto IDL y de bajo IDL), además de pseudopalabras. Los estímulos se presentaron de manera aleatoria, considerando las 20 primeras palabras como ensayo. Los alumnos debían responder lo más rápidamente posible, presionando una tecla de color verde situada a la derecha del teclado, en caso de que el estímulo correspondiera a una palabra, o una tecla roja ubicada a la izquierda en caso de que el estímulo no correspondiera a una palabra del español.

## Análisis de datos

En función de los objetivos, junto con el análisis a nivel descriptivo, se realizó análisis bivariado usando el coeficiente de correlación producto-momento de Pearson para establecer la relación entre las variables de interés. Para conocer la relación entre determinadas variables dependientes y su capacidad de explicar la variable comprensión lectora se estimaron diversos modelos de regresión lineal múltiple, a la vez que se evaluó mediante regresión logística binaria la capacidad del modelo de poder clasificar entre buenos y malos lectores. El análisis se realizó usando el software estadístico SPSS 19 (IBM, Chicago, IL).

## Resultados

Se presentan a continuación los estadísticos descriptivos de las variables medidas y las relaciones que se establecen en tanto simple asociación y capacidad predictiva de las variables independientes sobre la variable criterio.

Tabla 1

**Media y desviación estándar para las variables analizadas**

<b>Variable</b>	<b>M</b>	<b>DE</b>	<b>Min</b>	<b>Max</b>
Acierto TDL	148.22	9.83	101	158
Errores TDL	11.94	9.78	2	59
Tiempo IDL alto	727.59	134.66	490	1026
Tiempo correctas IDL alto	739.80	130.21	491	1025
Tiempo IDL bajo	830.94	131.55	537	1099
Tiempo correcta IDL bajo	846.82	140.76	551	1149
Promedio notas Lenguaje	5.97	0.53	4.9	6.9
Promedio general de notas	5.39	0.49	5.1	6.8
TEVI-R	45.06	9.44	25	65
LECTUM	12.66	5.27	5	25

Nota: TDL=Tarea de decisión léxica; IDL= Índice de disponibilidad léxica;  
Tiempo= Promedio de tiempo.

Si bien en general no se encuentran diferencias estadísticamente significativas en las variables TDL y tiempo de procesamiento de palabras con alto y bajo IDL en relación con el género y los buenos y malos lectores, es posible apreciar que el tiempo empleado en las palabras de IDL bajo siempre es mayor que el requerido para las de IDL alto, lo que corresponde a lo esperado<sup>1</sup>.

Se presentan a continuación los estadísticos bivariados, tanto a nivel de las variables como de los grupos contrastados por género y nivel de competencia lectora.

1 Una premisa en procesamiento léxico es que las palabras de alta frecuencia se leen más rápido y provocan menos respuestas erróneas en una TDL, lo contrario ocurre con las palabras de baja frecuencia. Como se dijo antes, en este estudio, el IDL se consideró como un valor equivalente al de la frecuencia.



Tabla 2

**Coefficiente de correlación producto-momento de Pearson para las variables medidas**

Variables	1	2	3	4	5	6	7	8	9	10
1. Acierto TDL	--									
2. Errores TDL	-.99**	--								
3. Tiempo IDL alto	.08	-.07	--							
4. Tiempo correctas IDL alto	-.04	.05	.98**	--						
5. Tiempo IDL bajo	.11	-.10	.91**	.90**	--					
6. Tiempo correcta IDL bajo	.04	-.03	.94**	.94**	.97**	--				
7. Promedio notas lenguaje	.01	-.03	-.16	-.13	-.32	-.23	--			
8. Promedio general de notas	.09	-.10	-.13	-.12	-.29	-.20	.79**	--		
9. TEVI-R	.02	-.02	-.17	-.15	-.23	-.14	.31	.44*	--	
10. LECTUM	.37*	-.36*	-.26	-.30	-.32	-.31	.17	.43*	.20	--

Nota:

TDL=Tarea de decisión léxica; IDL= Índice de disponibilidad léxica; Tiempo= Promedio de tiempo.

\* $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ 

Como se aprecia en la tabla 2, considerados en su conjunto, los coeficientes de correlación son bajos, exceptuando el que se da entre los tiempos globales con palabras de alto y bajo IDL y su respectivo porcentaje de respuestas correctas. Esto se debe a que la cantidad de aciertos, que es siempre mayor a los errores, está incluida en el tiempo total empleado en la tarea. Se decidió inicialmente considerar ambos indicadores esperando que hubiera alguna diferencia en los resultados; sin embargo, como se verá, la variable tiempo no permite diferenciar los grupos.

En relación con el rendimiento académico, se usó como indicador el promedio de notas de la asignatura de Lenguaje y el promedio general, los que correlacionan de forma inversa con el tiempo empleado en el reconocimiento de palabras con alto y bajo IDL. Aunque estos valores no son estadísticamente significativos, se puede interpretar, de forma global, que a menor nota, mayor tiempo empleado en la tarea. El rendimiento académico, a su vez, presenta una correlación directa con vocabulario pasivo, medido con el TEVI-R ( $r = .44$ ,  $p = .013$ ), y con comprensión lectora,

medida con LECTUM 7 ( $r=.43$ ,  $p=.014$ ). En el caso del promedio general de notas, la relación es moderada y estadísticamente significativa; en el caso de la nota de Lenguaje, está más asociada con vocabulario pasivo que con la comprensión lectora evaluada.

Junto al promedio general de notas, otras dos variables correlacionan de forma significativa con comprensión lectora, estas son los aciertos y los errores en la tarea de decisión léxica. Si bien los coeficientes de correlación no son altos ( $r=.37$ ,  $p=.038$  y  $r=-.36$ ,  $p=.042$  para aciertos y errores respectivamente), los datos se encuentran en la dirección esperada, es decir, a mayor rendimiento lector, mayor cantidad de aciertos y menor cantidad de errores en la tarea de decisión léxica. Esto permite sostener la idea de que un lector competente tendrá un buen desempeño en ambos procesos, tanto en lo que concierne a la comprensión lectora como al reconocimiento léxico, lo cual sugiere un grado de asociación entre el proceso de lectura y el reconocimiento de palabras. La frecuencia de aciertos en la TDL tiene relación con la calidad de las representaciones que un sujeto pueda elaborar sobre las palabras, incidiendo en los procesos de construcción del significado global del texto.

Debido a que la muestra estaba constituida por hombres y mujeres, y siendo conocidas las diferencias entre grupos halladas en resultados de otras mediciones que consideran, como parte del proceso, la comprensión lectora (SIMCE, PSU, PISA)<sup>2</sup>, se utilizó la prueba *t de Student* para muestras independientes, a fin de determinar si había diferencias de género verificando previamente que la variable contrastada se distribuyera de forma normal. En términos generales, existe diferencia en los tiempos empleados entre hombres y mujeres en todos los indicadores medidos, pero sin que en ninguno de ellos estas diferencias resultaran estadísticamente significativas. La única variable que presenta una diferencia considerable es vocabulario pasivo ( $t_{(30)}=2.769$ ,  $p=.01$ ), siendo el promedio de los hombres ( $M=49$ ,  $DE=8.28$ ) superior al de las mujeres ( $M=40.6$ ,  $DE=8.87$ ).

---

2 SIMCE: Sistema de Medición de la Calidad de la Educación, instrumento utilizado por el Ministerio de Educación de Chile para medir el desempeño en diversas áreas de conocimiento, entre otras, matemáticas, lengua materna, ciencias e inglés. PSU: Prueba de Selección Universitaria, instrumento utilizado por las universidades chilenas más antiguas para seleccionar estudiantes. PISA: Prueba de comprensión lectora empleada por la OECD.

Al separar los sujetos de la muestra en altos y bajos puntajes en comprensión lectora (buenos lectores v/s malos lectores), no se observan diferencias relevantes en los indicadores de TDL e IDL; en cambio, como es obvio, los grupos se diferencian en el rendimiento en comprensión lectora ( $t_{(30)}=9.23, p=.001$ ) con un promedio de 14.4 puntos ( $DE= 3.25$ ) los buenos lectores y 8.5 puntos ( $DE=2.18$ ) los malos lectores. En relación con el promedio de notas generales, los malos lectores alcanzan un 5,7 ( $DE=0.44$ ), mientras que los buenos lectores tienen un promedio de 6.1 ( $DE=0.50$ ), lo que también resulta ser una diferencia estadísticamente significativa ( $t_{(30)}=2.12, p=.043$ ).

Para determinar la capacidad de predecir el resultado en comprensión lectora a partir de las variables independientes, se usó un modelo de regresión lineal múltiple, tomando como variable dependiente el puntaje global obtenido en la prueba LECTUM 7 y como variables predictoras el número de aciertos en TDL, y el tiempo promedio de reconocimiento de palabras de alto y de bajo IDL. Para evitar la multicolinealidad, se empleó solo el tiempo de reconocimiento correcto y no el de los errores ni el promedio global de todas las palabras con alto o bajo IDL. Se utilizó como método de regresión el de pasos sucesivos (*stepwise*).

Se evaluó el cumplimiento de los supuestos básicos para ajustar un modelo de regresión. En primer lugar, se evaluó la distribución de las variables y se comprobó que correspondía a una distribución normal, con un leve sesgo positivo, pero que no afectaba la interpretación de los datos. Al mismo tiempo, se verificó que las variables explicativas contaran con un factor de inflación de la varianza (VIF) menor que 3, descartando así multicolinealidad. Los resultados del modelo se presentan en la tabla 3.

Tabla 3

**Regresión lineal múltiple para variable comprensión lectora**

	B	EE	$\beta$	t	sr <sup>2</sup>
Intercepto	39.07***	15.18		-2.57*	
Promedio general de notas	.78***	1.67	.40	8.87*	.16
Acierto TDL	1.51***	.08	.33	2.13*	.11

R=.54\*\*\*, R<sup>2</sup>=.30, R<sup>2</sup>ajustado=.25

\*p<.05, \*\* p< .01, \*\*\* p<.001

De todas las variables empleadas, solo dos resultaron ser estadísticamente significativas como predictor de la comprensión lectora ( $F_{(2,29)}=6.09, p=.006$ ), que en conjunto explican un 30% del cambio en la varianza de la variable explicada: 1) el promedio general de notas de los estudiantes y 2) la tarea de decisión léxica, siendo el primero un predictor levemente más robusto. Todas las demás variables evaluadas no mostraron ninguna contribución al modelo predictivo.

Debido a que la prueba de comprensión lectora empleada no contaba, en el momento de la aplicación para este estudio, con puntajes normativos que permitieran evaluar si el nivel de comprensión lectora era alto o bajo para la población estudiada, se decidió reclasificar a los estudiantes como buenos o malos lectores según el puntaje global, que en la muestra se distribuía entre 5 y 25 puntos de un total de 31 puntos posibles. De esta forma los alumnos que presentaron 12 puntos o menos se etiquetaron como malos lectores y aquellos en el rango entre 13 y 25 puntos, como buenos lectores. Con estos valores, se estimó un modelo de regresión logística binaria, usando como variable dependiente la nueva variable dicotómica y se evaluaron nuevamente todas las variables predictoras usadas anteriormente para, finalmente, determinar un modelo que permitió clasificar correctamente en promedio al 72% de los estudiantes entre buenos y malos lectores a partir de la variable aciertos en TDL, promedio de notas generales y sexo. En la tabla 4 se pueden apreciar los indicadores generales del modelo de regresión logística binaria ajustado.

Tabla 4

**Regresión logística binaria para clasificar a estudiantes con alto y bajo puntaje en comprensión lectora.**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inf.	Sup.
Acierto TDL	.16	.101	2.55	1	.110	1.175	.964	1.431
Promedio general notas	1.95	.975	4.01	1	.045	7.052	1.043	47.667
Sexo (hombre)	-1.67	.927	3.26	1	.071	.187	.030	1.153
Constante	-34.93	17.57	3.95	1	.047	.000		

$\chi^2_{(3)}=10.73, p=.013$ . Prueba de Hosmer y Lemeshow  $\chi^2_{(8)}=2.28, p=.971$ .

-2 log de la verosimilitud=.33,46;  $R^2_{Cox y Snell}=.29$ ;  $R^2_{Nagelkerke}=.38$

La variable categórica sexo se ha incluido en el modelo, ya que mejora la clasificación de los casos a los grupos; de esta forma el modelo propuesto clasifica correctamente al 71% de los malos lectores y al 73% de los buenos lectores. Al reemplazarse el promedio general de notas por el promedio de la asignatura de lenguaje mejora a un 77% para malos lectores, pero disminuye a un 60% el de los buenos lectores. Lo anterior implica que el ser un buen estudiante, medido a través del indicador global promedio general de notas, constituye un buen indicador para distinguir entre los buenos y malos lectores. El rendimiento en la asignatura de Lenguaje, por su parte, permite aumentar el acierto solamente en la clasificación de malos lectores. Se prefirió dejar en el modelo solo la variable promedio general de notas, porque mejoraba la clasificación con ambos grupos, lo que no ocurría si se consideraban ambas en el modelo.

## Discusión y conclusiones

Del análisis de los datos se desprende que la comprensión lectora y el rendimiento académico están relacionados con la competencia léxica. En efecto, la frecuencia de aciertos y de errores en la TDL revelan distintos niveles de los sujetos en su competencia léxica y la relación intrínseca entre esta y la competencia lectora. Tal como sostiene la hipótesis de la calidad léxica (Perfetti, 2010; Perfetti et al., 2007; Perfetti & Hart, 2001), la existencia de limitaciones en la eficiencia con que las palabras son reconocidas afecta el desempeño lector. De igual forma, el esfuerzo desplegado en la recuperación de las propiedades fonológicas y semánticas de una palabra se vincula con una mayor demanda de recursos de la memoria, lo que finalmente afecta los procesos de comprensión que dependen de las habilidades léxicas.

Aunque las diferencias que se observan entre los buenos y malos lectores en el tiempo promedio de palabras de alto y bajo IDL, así como en vocabulario pasivo, no son estadísticamente significativas, es posible interpretar los resultados y suponer que en una muestra mayor pudiera acrecentarse la tendencia de las diferencias observadas en este estudio. En la frecuencia de errores y en la frecuencia de aciertos de la TDL, se aprecia que los buenos lectores se equivocan menos, a diferencia de los lectores menos competentes, ya que el primer grupo es capaz de reconocer una mayor cantidad de palabras y, lógicamente, cometen menos errores. A pesar de su bajo valor de significación, estos resultados apoyarían la hipótesis propuesta esta investigación. De igual forma, al

comparar los tiempos de reacción de ambos grupos, es posible observar que los buenos lectores tienen un desempeño más alto que los lectores menos competentes, siendo los tiempos de reacción del primer grupo menores que los del segundo grupo. Estos resultados están directamente relacionados con el IDL de las palabras del experimento de decisión léxica, ya que se observan mayores tiempos de reconocimiento de las palabras de bajo IDL a diferencia de las palabras de alto IDL, tanto para los buenos, como para los malos comprendedores. El léxico disponible, como índice de frecuencia, sería un indicador que permitiría avalar la idea de que la competencia léxica se relaciona con la competencia lectora, facilitando o dificultando la eficacia en ambos casos. Cabe señalar que la relación entre las variables opera de manera lógica: al existir un índice de disponibilidad léxica más alto, habrá un menor tiempo de reconocimiento y un mayor acierto en la TDL. Lo contrario ocurre si analizamos el tiempo promedio de las palabras de bajo IDL, en donde se observa un mayor tiempo de reacción, comparado en el promedio de las palabras de alto IDL y una menor cantidad de aciertos por parte de ambos grupos. Estos resultados permiten determinar que el IDL es una variable significativa para considerar como propiedad de la calidad en las representaciones de las palabras por parte de los sujetos y como variable significativa en un experimento de TDL.

Los hallazgos de este estudio, junto con confirmar la hipótesis de Perfetti en nuestro medio (Perfetti, 2010; Perfetti et al., 2007; Perfetti & Hart, 2001) en cuanto al papel del vocabulario en la comprensión y, con ello, reforzar los principios establecidos en la teoría de Kintsch (1998), contribuyen a entender mejor la problemática de la comprensión lectora en escolares. Todo indica, por una parte, que el bajo desarrollo léxico alcanzado por los estudiantes tiene una incidencia significativa en el rendimiento lector, esto es, alumnos con menor nivel de vocabulario tendrán peores resultados en pruebas de comprensión lectora. Por otra parte, una mejor competencia léxica asegura mayores niveles de rendimiento en lectura comprensiva. Al mismo tiempo, el vocabulario, sobre todo el de mayor especialización y menor frecuencia, se encuentra principalmente en textos científicos y humanistas, así como en los medios de comunicación. Este vocabulario no es accesible fácilmente en la comunicación oral coloquial, sino que se encuentra más vinculado a espacios académicos y formales. Podríamos estar frente a un “callejón sin salida”, ya que se requiere elevar el nivel de vocabulario para mejorar la

lectura, pero ello pasa principalmente por la lectura. Aquí cabe un papel primordial a la pedagogía y, dentro de ella, al docente, quien puede actuar como una suerte de “mediador cultural” a través de actividades que apunten a incorporar vocabulario nuevo mediante ciertas prácticas pedagógicas, entre ellas, la lectura guiada. Nuevas investigaciones en esta dirección podrían arrojar luces sobre la efectividad de las líneas de acción aquí sugeridas.

## Referencias bibliográficas

- Echeverría, M., Herrera, M. O. & Sègure, T. (2002). *TEVI-R, test de vocabulario en imágenes*. Concepción: Universidad de Concepción.
- Johnson-Laird, Ph. (1983). *Toward a Cognitive Science of Language, Inference and Consciousness*. Cambridge, MA: Harvard University Press.
- Just, M. A., Carpenter, P. A. (1980). A theory of reading: from eye fixation to comprehension. *Psychological Review*, 87, 329–354.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. New York: Cambridge University Press.
- \_\_\_\_\_ & Rawson, K. A. (2007). Comprehension. En M.J. Snowling & Ch. Hulme, (Eds.), *The Science of Reading. A Handbook* (pp. 209-226). Malden, MA: Blackwell Publishing.
- Magliano, J., Wiemer-Hastings, K., Millis, K., Muñoz, B. & McNamara, D. (2002). Using latent semantic analysis to assess reader strategies. *Behavior Research Methods, Instruments & Computers*, 34 (2), 181-188.
- McNamara, D. & Kintsch, W. (1996). Learning from texts: effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247- 288.
- Muter, V., Hulme, C., Snowling, M. & Stevenson, J. (2004). Phonemes, rimes, vocabulary and grammatical skills as foundations of early reading development.

- Evidence from a longitudinal study. *Developmental Psychology*, 40 (5), 665-681.
- Oakhill, J. V., Cain, K. E., & Bryant, P. E. (2003). Prediction of comprehension skill in the primary school years. Paper presented at *Biennial meeting of the Society for Research in Child Development*. Tampa, EE.UU.
- Perea, M. & Rosa, E. (1999). Psicología de la lectura y procesamiento léxico visual: Una revisión de técnicas experimentales y procedimientos de análisis. *Psicológica*, 20, 65-90.
- Perfetti, Ch. (1985). *Reading ability*. New York: Oxford University Press.
- \_\_\_\_\_. (2010). Decoding, Vocabulary and Comprehension. The Golden Triangle of Reading Skill. En M. McKeown & L. Kucan (Eds.), *Bringing Reading Research to Life* (pp. 291-303). New York: The Guilford Press.
- Perfetti, Ch. & Hart (2001). The lexical bases of comprehension skill. En D. Gorfiend, (Ed.), *On consequences of meaning selection* (pp. 67-96). Washington, D. C.: APA.
- \_\_\_\_\_, Landi, N. & Oakhill, J. (2007). The Acquisition of Reading Comprehension Skill. En M.J. Snowling & Ch. Hulme (Eds.). *The Science of Reading. A Handbook* (pp. 227-247). Malden, MA: Blackwell Publishing.
- Riffo, B., Véliz, M., Castro, G., Reyes, F., Figueroa, B., Salazar, O. & Herrera, M. O. (2011). LECTUM. *Prueba de comprensión lectora*. Conicyt, Proyecto Fondef D08i1179.
- Share, D. L. (1999). Phonological decoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72, 95-129.
- Smith, E. & Kosslyn, S. (2008). *Procesos cognitivos. Modelos y bases neurales*. Madrid: Pearson Prentice Hall.
- Valencia, A. & Echeverría, M. (1999). *Disponibilidad léxica en estudiantes chilenos*. Santiago: Universidad de Chile y Universidad de Concepción.