

Complejidad textual, lecturabilidad y rendimiento lector en una prueba de comprensión en escolares adolescentes*

Textual Complexity, Readability and Reading Performance on a Comprehension Test in Adolescents School

Recibido: abril 25 de 2013 | Revisado: septiembre 6 de 2013 | Aceptado: enero 11 de 2014

DANIELA CAMPOS SAAVEDRA **

PAULA CONTRERAS CARMONA

BERNARDO RIFFO OCARES

MÓNICA VÉLIZ

Universidad de Concepción, Chile

ALEJANDRO REYES REYES ***

Universidad Santo Tomas, Chile

RESUMEN

En este artículo se estudia el problema de la *lecturabilidad* –la dificultad intrínseca de los textos para ser comprendidos– desde una perspectiva psicolingüística. El objetivo de este estudio fue establecer en qué medida la lecturabilidad de un texto depende no solo de sus propiedades estructurales (complejidad lingüística), sino también de las habilidades del lector y de la naturaleza de las tareas de comprensión. Se realizó una investigación de tipo cuasiexperimental en la que se construyeron dos pruebas de comprensión de lectura, manipulando la complejidad de los textos. Las pruebas incluyeron ítems para medir diversos aspectos de la comprensión según un modelo de evaluación basado en un enfoque teórico psicolingüístico. Las pruebas fueron aplicadas a 208 escolares de tres tipos de establecimientos educativos distintos. Los resultados de los análisis estadísticos de los datos permiten confirmar la hipótesis de que la lecturabilidad de los textos no depende solo de las propiedades textuales, sino de la relación de estas con las características de los lectores y las tareas de comprensión definidas en las pruebas.

Palabras clave

comprensión lectora; lecturabilidad; índices de complejidad textual; rendimiento lector

doi:10.11144/Javeriana.UPSY13-3.cltlr

Para citar este artículo: Campos, D., Contreras, P., Riffo, B., Véliz, M. & Reyes, A. (2014). Complejidad textual, lecturabilidad y rendimiento lector en una prueba de comprensión en escolares adolescentes. *Universitas Psychologica*, 13(3). <http://dx.doi.org/10.11144/Javeriana.UPSY13-3.cltlr>

* Este estudio fue desarrollado en el marco del proyecto de investigación Fondef D08i1179 "Desarrollo de una prueba validada y normada para el diagnóstico de la comprensión lectora en el sistema escolar chileno", financiado por la Comisión Nacional de Ciencia y Tecnología (Conicyt) de Chile.

** Correos electrónicos: ancampos@udec.cl, paulcontreras@udec.cl, bernardo@udec.cl, mveliz@udec.cl

*** Correo electrónico: areyesr@santotomas.cl

ABSTRACT

In this paper, the problem of readability is studied from a psycholinguistic perspective. Readability is the intrinsic difficulty of texts to be understood. The aim of this study was to determine to what extent the readability of a text depends not only on its structural properties (linguistic complexity), but also the skills of the reader and the nature of comprehension tasks. An experimental research was carried out, in which two reading comprehension tests were developed by manipulating the complexity of texts. The tests included items to measure various aspects of comprehension according to an evaluation model based on a theoretical psycholinguistic approach. The tests were applied to 208 participants in three different types of schools. The results confirmed the hypothesis that the readability of texts depends not only on the textual properties, but also on the relationship of these with the readers' characteristics and the comprehension tasks defined in the tests.

Keywords

reading comprehension; readability; textual complexity indices; reading performance.

Introducción

La facilidad/dificultad con la que un texto puede ser leído y comprendido se conoce como *lecturabilidad*. Si bien se trata de un fenómeno vinculado al ámbito educativo, sus alcances cognitivos no son menores, tanto que ha devenido un problema de investigación en psicolingüística, perspectiva desde la cual se desarrolla el presente estudio. En principio, el interés por el tema tuvo una motivación pedagógica en la búsqueda de criterios para la selección de textos de enseñanza que resultaran adecuados para los diversos niveles de la escuela. Desde los inicios, tanto en Inglaterra como en Rusia a fines del siglo XIX, se consideró que ciertos factores lingüísticos tenían incidencia en la complejidad (y dificultad) de los textos, entre ellos, el vocabulario y la longitud de las oraciones (una exhaustiva revisión del desarrollo del concepto de lecturabilidad se encuentra en Du Bay, 2004 y en Zakaluk & Samuels, 1988).

Más tarde, con el propósito de determinar en forma precisa el grado de dificultad de los textos, se desarrollaron las llamadas *fórmulas de lecturabilidad*, entre las que destacan Reading Ease Score (RES) de Flesch, Dale-Chall, el SMOG, la Flesch-Kincaid, el Gunning FOG test o el Gráfico de Fry (Du Bay, 2004; una adaptación al español del Gráfico de Fry se encuentra en Parodi, 1986). Aun con diferencias específicas en las relaciones de cálculo, todos estos procedimientos tienen en común el considerar que ciertas propiedades verbales de los textos son medibles y, sobre esta base, se puede obtener un valor que da cuenta de la complejidad intrínseca del material textual. Dicho valor constituye, desde esta perspectiva, un predictor de la dificultad potencial que el texto representa para su comprensión y, por ende, del rendimiento de los lectores (Kintsch & Miller, 1981; Parodi, 1986). Así, entre los factores que se incluyen en las fórmulas, se encuentran la longitud de las oraciones, el número de sílabas de las palabras, el que el léxico empleado sea conocido o no por los lectores, etc.

Desde el punto de vista de la psicolingüística actual, sin embargo, la comprensión lectora constituye un proceso cognitivo de alta complejidad en el que, además de los factores textuales, están involucrados

los conocimientos previos del lector, sus habilidades cognitivas y las operaciones implicadas en las tareas que este debe realizar para conseguir su objetivo de lectura (Kintsch, 1998; Kintsch & Rawson, 2007; Snow, 2002; Zwaan & Singer, 2003). La dificultad o facilidad con la que un determinado texto puede ser comprendido depende, entonces, de varios factores. Este enfoque psicolingüístico multifactorial concibe la lecturabilidad como el resultado de las propiedades del texto en relación con las características del lector y la naturaleza y complejidad de las tareas de lectura (Kintsch & Miller, 1981; McNamara, Louwerse & Graesser, 2002; McNamara & Shapiro, 2005; Snow, 2002).

La comprensión lectora

Una de las teorías más influyentes en la actualidad, conocida como *modelo de construcción e integración* (Kintsch, 1998), postula que la comprensión de un texto escrito implica dos fases. Durante la primera (fase de construcción) tienen lugar los llamados procesos de bajo nivel, como lo son aquellos responsables de la decodificación de la señal gráfica (percepción de las letras), el reconocimiento de palabras y el procesamiento sintáctico. Los contenidos semánticos producto de estas operaciones forman representaciones proposicionales, las que se integran luego en una red llamada *base de texto*.

En un segundo momento (fase de integración), algunos de los elementos activados de las diversas proposiciones se mantienen en esta condición para permitir su vinculación en representaciones de la memoria de largo plazo. Junto con la red de proposiciones, el lector construye otra representación, una de carácter análogo, conocida como *modelo de situación* y que corresponde a la situación referida por el texto. Debido a que la memoria no puede operar sino con un número limitado de unidades, este complejo proceso se lleva a cabo por partes (palabra a palabra, frase a frase). Así, el procesamiento de cada unidad gramatical involucrará todas las demás operaciones de bajo nivel ya mencionadas y su integración en una representación mayor (Kintsch, 1998).

Propiedades textuales y lecturabilidad

Existe hoy abundante evidencia empírica para sostener que determinados factores lingüísticos inciden en el desempeño de los lectores. De los hallazgos experimentales se concluye que en los diversos niveles de organización del lenguaje es posible identificar propiedades relacionadas con el rendimiento lector, entre las que destacan el léxico, la gramática y el significado.

Complejidad léxica

La dificultad intrínseca del vocabulario de un texto se entenderá en este estudio como complejidad léxica y, aunque existen numerosos aspectos de las palabras que podrían tener relevancia psicolingüística, incluimos tres de ellos, a saber, *densidad léxica*, *diversidad léxica* y *frecuencia léxica*. La densidad léxica (DeL) se define como el porcentaje de palabras nocionales en oposición a los ítems gramaticales en un texto dado (Baker, 1995). Esta medida permite conocer en forma indirecta la complejidad y la carga informativa de un texto, lo que influye en la comprensión lectora; un texto con una alta proporción de palabras de contenido contiene más información que un texto con una alta proporción de palabras funcionales (preposiciones, intersecciones, pronombres, conjunciones y palabras contables) (Johansson, 2008). Evidencia empírica ha demostrado que existe una correlación entre una baja densidad léxica y la facilidad con que un texto es comprendido (Perfetti, Landi & Oakhill, 2005). Por el contrario, una elevada densidad léxica es un rasgo que diferencia la escritura de la oralidad, ya que esta última tiende a un léxico más repetitivo y a una menor complejidad sintáctica (Halliday, 1985).

La diversidad léxica (DiL), a su vez, representa la proporción entre palabras diferentes del texto respecto del número total de palabras (Johansson, 2008). Mientras mayor sea la diversidad léxica de un texto, más palabras no repetidas contendrá este, con lo que se eleva la carga semántica al introducir un mayor volumen de significados distintos. Una baja diversidad léxica, en cambio, implica un mayor número de palabras repetidas y, en consecuencia,

una carga semántica más liviana. Como es de esperar, estos valores inciden en el rendimiento en tareas de lectura: mientras más palabras diferentes, más tiempo de lectura y mayor consumo de recursos (Graesser, McNamara, Lowerse & Cai, 2004).

La frecuencia léxica, por su parte, refleja el uso efectivo de las palabras en la comunidad. El valor se obtiene del registro de grandes volúmenes de texto (varios cientos de millones de palabras) y, aunque el corpus puede presentar algún sesgo (por ejemplo, hacia el texto escrito más que hacia el texto oral), es la variable que mejor predice la velocidad de reconocimiento de palabras en diversas pruebas experimentales. Mientras más frecuente es una palabra, es decir, en tanto más veces aparece en diferentes textos que circulan en una comunidad, menos recursos cognitivos exigirán su percepción, reconocimiento e integración en el procesamiento del texto. Esto tiene, sin duda, importantes efectos para la comprensión; de hecho, en la medida que las palabras de un texto son menos frecuentes, más “pesada” se hará la lectura, más lento el proceso y mayor la tasa de error en las tareas que se le asignen al lector (Álvarez, Alameda & Domínguez, 1999; Graesser et al., 2004).

Complejidad sintáctica

La complejidad sintáctica se define en términos de la estructura gramatical de las oraciones que componen el texto. Desde las primeras etapas del desarrollo, los niños incorporan paulatinamente estructuras gramaticales más complejas al repertorio de oraciones que pueden producir y comprender, proceso que se extiende más allá de la primera infancia a través de toda la etapa escolar (Echeverría, 1978; Autor, 1988). De acuerdo con Hunt (1970), la capacidad de combinar más oraciones formando estructuras de mayor complejidad es un signo de “madurez sintáctica”. Desde esta perspectiva, podemos considerar que un texto alcanzará un mayor grado de complejidad si las oraciones que lo conforman están formadas, a su vez, por un mayor número de elementos constitutivos, siendo estos elementos no solo palabras, sino cláusulas, algunas incrustadas dentro de otras (subordinación).

Se sabe, además, que los textos varían considerablemente en su complejidad gramatical según el género (tipo de texto) y la trama estructural (narrativa, argumentativa o expositiva) (Autor, 1999). Las oraciones, su longitud y constitución gramatical tienen también un importante papel en la comprensión del discurso, pues este complejo proceso se hace efectivo palabra a palabra, frase a frase, debido a las limitaciones de la memoria de trabajo (Baddeley, 2000; Baddeley & Hitch, 1974). A ello debe agregarse que la naturaleza secuencial de las expresiones verbales, resultado de la combinación de unidades mínimas carentes de significado (fonemas) para producir unidades mayores dotadas de significado (morfemas, palabras, oraciones), determina que su procesamiento sea también secuencial, distribuido en el tiempo (Just & Carpenter, 1980).

Bajo tales condiciones, la longitud y complejidad de las oraciones que conforman el texto tiene importantes consecuencias para el desempeño lector. En efecto, oraciones breves y simples –constituidas por una sola cláusula, esto es, por un solo verbo conjugado (finito) y los elementos que lo determinan– pueden incluso ser procesadas sin que en el límite final se produzca la integración, sino hasta concluir la siguiente oración. Por el contrario, oraciones muy extensas y complejas –que incluyen cláusulas subordinadas incrustadas o que presentan ambigüedad estructural– demandarán mayores recursos cognitivos con la consiguiente recarga de la memoria (Graesser et al., 2004; Kintsch, 1998; McNamara & Shapiro, 2005).

Complejidad semántica

En el marco de este estudio, determinar si un texto presenta una mayor o menor complejidad semántica supone definir una forma de representar este aspecto del lenguaje desde una perspectiva cognitiva. Las proposiciones se han incorporado en la psicología como una alternativa muy productiva para estos efectos (de Vega, 1985; de Vega, Carreiras, Gutiérrez-Calvo & Alonso-Quecuty, 1990; Kintsch, 1998). Estas unidades básicas de significado se encuentran compuestas por *palabras-concepto* (Kintsch, Kozminsky, Streby, McKoon &

Keenan, 1975), las que funcionan como entradas en el lexicon de la memoria semántica y pueden ser representadas por una o más palabras en la estructura superficial del texto –*tokens* para el concepto abstracto– (Turner & Greene, 1978).

Las proposiciones se componen de un predicado o concepto proposicional y uno o más argumentos, cada uno cumple un papel semántico único. Los predicados pueden ser identificados como verbos, adjetivos, adverbios y frases conectivas y se refieren a propiedades o relaciones. Los argumentos, por su parte, pueden ser conceptos u otras proposiciones insertadas y cumplen diferentes funciones semánticas, tales como agente, objeto o meta y se refieren a individuos, sean cosas o personas (de Vega et al., 1990; Kintsch, 1998). Como unidades de procesamiento del lenguaje, las proposiciones constituyen piezas fundamentales para la elaboración de la representación semántica (base de texto). Abundante evidencia de la realidad psicológica de estas unidades se encuentra en la investigación experimental de las últimas décadas. Entre otros hallazgos, se puede mencionar que las proposiciones tienen un papel relevante en la memoria de textos: las palabras se recuerdan mejor cuando estas pertenecen a una determinada proposición que cuando se encuentran fuera de ella (Goetz, Anderson & Schallert, 1981). Otro efecto importante es del *priming*: una palabra permite activar con mayor fuerza el reconocimiento de otra cuando ambas pertenecen a la misma unidad proposicional (Ratcliff & McKoon, 1978). La densidad proposicional de un texto (DeP), aspecto de interés en esta investigación, ha demostrado ser un potente predictor del consumo de recursos cognitivos (de Vega et al., 1990; Kintsch & Keenan, 1973).

Las tareas de comprensión

Comprender lo que se lee depende no solo de las características del texto, sino también de los objetivos de la lectura y las operaciones cognitivas que en función de estos propósitos se efectúen. En el marco de un procedimiento evaluativo, estas operaciones pueden llamarse *tareas de comprensión*, las que implican diversos niveles de dificultad según

los recursos que ellas demandan del lector. Así, leer para identificar cierta información textual explícita en el texto (el nombre de un personaje, una cifra, el color de un objeto, etc.) tendrá un costo más bajo que resumir el contenido del texto o vincular causalmente dos hechos cuando no se expresa ningún vínculo entre ellos, sino que este debe ser inferido.

Con el propósito de desarrollar procedimientos de evaluación de la comprensión lectora, el equipo de investigadores ha elaborado un modelo psicolingüísticamente fundamentado, cuya primera versión se publicó en 1992 (Autor & Autor, 1992, 1993). Recientemente, los autores han incorporado nuevos elementos al modelo, el que distingue tres dimensiones de la comprensión, a saber, *textual*, *pragmática* y *crítica* (Autor & Autor, 2011). La comprensión textual se centra en las operaciones orientadas al procesamiento de la información lingüística contenida en el texto en los niveles léxico, proposicional, microestructural y macroestructural (Kintsch, 1998; van Dijk & Kintsch, 1983); se trata de la comprensión de palabras (identificar o generar el significado de un vocablo utilizando las claves que proporciona el texto), de oraciones (identificar los componentes del esquema proposicional), de secuencias de dos o tres oraciones (vincular los hechos referidos en las oraciones mediante relaciones condicionales o funcionales) o de fragmentos mayores del texto (reconocer el significado global del texto, establecer relaciones de causalidad entre elementos que abarcan dimensiones mayores o bien el texto en su conjunto, entre otras).

La comprensión pragmática, por su parte, considera las relaciones entre el texto y el contexto; se evalúa aquí la capacidad del lector de relacionar el texto con el contexto de situación (el entorno inmediato, el marco temporal-espacial en el que el texto se produce, transmite y recibe) y con el contexto cultural al cual pertenece. Lo anterior supone: 1) reconocer o identificar los participantes en el circuito comunicativo del texto, 2) situar y vincular el texto con el contexto inmediato y en el contexto cultural, 3) determinar los propósitos del texto y 4) establecer los sentidos del texto. Por último, en la comprensión crítica se evalúa la capacidad del lector para realizar un análisis crítico del texto, lo que puede expresarse

en: 1) utilizar la información que se obtiene de un texto, transfiriéndola a un contexto diferente para resolver una tarea (e. g. resolver problemas y otras actividades que impliquen una acción reflexiva), 2) emitir un juicio que dé cuenta de que el lector aplica un sistema de valores para evaluar diversas situaciones representadas en el texto y 3) comprender significados no convencionales y todas las operaciones retóricas del discurso.

Teniendo en consideración los factores descritos aquí, el presente estudio se propone determinar empíricamente en qué medida la lecturabilidad de un texto depende no solo de sus propiedades estructurales (complejidad lingüística), sino también de las habilidades del lector y de la naturaleza de las tareas de comprensión. Así, el objetivo de la investigación está orientado por la pregunta de investigación que sigue: ¿existen otros factores relevantes, además de la complejidad textual, que determinan la lecturabilidad de los textos en el contexto de la evaluación de la comprensión lectora? Como respuesta a esta interrogante, se propone la siguiente hipótesis: debido a la complejidad del proceso de comprensión lectora, la lecturabilidad depende no solo de las propiedades textuales, sino también de las características del lector y de las tareas de comprensión.

Método

Diseño

Se realizó un estudio de tipo cuasiexperimental con un diseño bifactorial 2x3 de efectos fijos, el que incluyó como variables independientes el factor Complejidad del Texto y el factor Tipo de Establecimiento Educativo. Se conformaron así seis tratamientos resultantes de la combinación de los dos niveles vinculados al grado de complejidad textual (alta dificultad/baja dificultad) con los tres niveles correspondientes al tipo de establecimiento educativo (municipal/particular subvencionado/particular pagado)¹. Como variable dependiente

1 Existen tres formas de administración en el sistema escolar de Chile: las escuelas municipalizadas (públicas), las particulares subvencionadas (dependientes de personas o corporaciones

se utilizó la comprensión lectora medida por dos pruebas construidas en función de los textos (alta dificultad/baja dificultad). Ambas pruebas consideraron un puntaje total del rendimiento lector y tres puntajes parciales, resultantes de la descomposición del puntaje total, que dan cuenta del rendimiento en las dimensiones comprensión textual, pragmática y crítica.

Participantes

La muestra estuvo compuesta por un total de 208 participantes, 109 hombres (52.4 %) y 99 mujeres (47.6 %), que cursaban octavo año de enseñanza básica², con un rango de edad entre 13 y 16 años y una media de edad de 13.59 años ($DE = 0.646$). Los sujetos pertenecían a tres tipos de establecimientos distintos desde la perspectiva de su dependencia administrativa, como se muestra en la Tabla 1.

Instrumentos

Para evaluar la comprensión lectora se diseñaron dos pruebas, cada una de ellas consistente en un texto expositivo sobre el ciclo del carbono, seguido de 14 ítems de selección múltiple con alternativa única de respuesta correcta. Estos ítems apuntaban a tres dimensiones de la comprensión, a saber: textual, pragmática y crítica, construidos a partir del modelo de evaluación de Autor y Autor (2011). Los dos textos utilizados fueron intervenidos, mo-

dificando sus propiedades textuales con el de fin de obtener dos versiones con índices distintos en su complejidad léxica (densidad, diversidad y frecuencia léxica), sintáctica (longitud de la cláusula y la oración e índice de subordinación) y semántica (densidad proposicional). Luego, para confirmar que ambos textos diferían en complejidad textual, estos fueron analizados mediante un sistema automático (Trunajod)³, basado en un etiquetador que permite obtener cantidades (palabras, oraciones, vocablos, cláusulas, palabras nocionales y proposiciones) e índices de complejidad textual (longitud de la cláusula y de la oración, índice de subordinación, densidad léxica, diversidad léxica, densidad proposicional, frecuencia promedio y frecuencia promedio logarítmica). Este análisis computacional corroboró las diferencias de ambos textos en los indicadores: densidad, diversidad y frecuencia léxica. Del mismo modo, el material lingüístico difería en el promedio de longitud de las oraciones y cláusulas, así como en el índice de subordinación. En cuanto a la complejidad semántica, también se alcanzó una diferencia en la densidad proposicional (Tabla 2). Además, con el propósito de confirmar que los textos eran efectivamente más complejos y más simples, respectivamente, se aplicó una encuesta a 162 escolares de 8.º básico y a 25 docentes. Los resultados del análisis estadístico basados en la prueba *t* de Student mostraron diferencias significativas para las respuestas de valoración subjetiva con respecto a la lecturabilidad de cada texto, realizada tanto por estudiantes como por docentes ($p = 0.0001$ en ambos casos), lo que confirma que estas diferencias

privadas, pero con financiamiento del Estado) y las particulares pagadas (privadas, sin financiamiento estatal).

2 En el sistema escolar chileno, la enseñanza básica incluye 8 grados (1.º a 8.º). La edad oficial para el inicio del ciclo es de 6 años, se espera que en 8.º básico los alumnos tengan una edad no inferior a 13 años.

3 Software diseñado por Mónica Véliz y Bruno Karelovic en el marco del proyecto Fondef D08i1179.

TABLA 1

Distribución de estudiantes de enseñanza general básica según tipo de establecimiento educacional y sexo

Tipo de establecimiento	Hombres	Mujeres	N	%
Municipal	40	33	73	35.1
Part. Subvencionado	33	41	74	35.6
Part. Pagado	36	25	61	29.3
Total	109	99	208	100

Fuente: elaboración propia.

en los índices de complejidad concuerdan con la percepción por parte de escolares y docentes de que un texto es más “difícil” de comprender que el otro.

Los ítems de las pruebas de comprensión lectora se conceptualizaron como elementos de una prueba de aprovechamiento, estableciéndose la validez de contenido mediante la opinión de seis jueces, los que determinaron la respuesta correcta al ítem y la dimensión a la cual pertenecía (textual, pragmática y crítica), dado lo cual, como plantean Nunnally y Bernstein (1995), no se requiere correlacionar la medida con un criterio externo para ser validada. La confiabilidad de la prueba, obtenida con la muestra objeto de estudio, fue establecida mediante el coeficiente Kuder-Richardson (KR20) y el método de división por mitades con corrección de Spearman Brown ($\rho_{xx'}$). Para el texto de baja complejidad textual se obtiene un KR20 = 0.56 y un coeficiente $\rho_{xx'} = 0.58$ y para el texto de alta complejidad la confiabilidad fue de un KR20 = 0.45 y un $\rho_{xx'} = 0.49$. Estos coeficientes indican una baja consistencia interna del instrumento, lo que se explica tanto por la baja extensión de la prueba como por su heterogeneidad (Aiken, 2003; Anastasi & Urbina, 1998).

Procedimientos

Los participantes fueron asignados en forma aleatoria a las pruebas con textos de distinto nivel de complejidad léxica, sintáctica y semántica. La prueba fue administrada de manera colectiva en un único momento en el aula habitual de clases de cada

grupo. El instrumento utilizado se entregó impreso y las instrucciones estandarizadas impartidas por los investigadores. La aplicación de la prueba tardó entre 30 y 40 minutos y se realizó en tres establecimientos educacionales que fueron seleccionados aleatoriamente del universo de establecimientos de la provincia de Concepción, Chile. Previo a la aplicación de los instrumentos, se consiguió el consentimiento de los directores de los distintos establecimientos, habiéndoseles proporcionado la información necesaria para decidir respecto de su participación en la investigación.

Análisis de datos

Con el objeto de establecer el efecto de los factores Complejidad del Texto y Tipo de Establecimiento Educacional sobre el rendimiento global y las dimensiones de la comprensión lectora evaluados, se realizaron análisis de varianza factorial (ANOVA factorial) y análisis multivariado de la varianza (MANOVA). Para estudiar las diferencias entre los promedio de los grupos se utilizaron las pruebas de comparaciones múltiples *post hoc* DHS de Tukey, Scheffe o la prueba de Bonferroni. Se efectuaron también cálculos de estadísticas descriptivas para poder caracterizar a la muestra. Los grupos presentan una adecuada equivalencia, la cual fue verificada mediante la prueba chi cuadrado de homogeneidad ($\chi^2 = 0.51, p = 0.775$). Para todos los análisis estadísticos se empleó el programa IBM SPSS Statistics versión 20 (Statistical Package for the Social Sciences).

TABLA 2
Índices de complejidad textual

Complejidad del texto	Índices de complejidad								
	LO	LC	IS	DeP	DeL	DiL	DFN	FP	FPL
Texto baja complejidad	16.5	10.5	1.6	35	50	57	3.7	1260	124.5
Texto alta complejidad	30.9	12.6	2.5	38	51	62	3.5	944	94.4

Nota. LO = Longitud promedio de la oración; LC = Longitud promedio de la cláusula; IS = Índice de subordinación; DeP = Densidad proposicional; DeL = Densidad léxica; DiL = Diversidad léxica; DFN = Densidad de la frase nominal; FP = Frecuencia promedio; FPL = Frecuencia promedio logarítmica.
Fuente: elaboración propia.

Resultados

Del análisis factorial de la varianza se desprende que existen diferencias entre los puntajes medios de comprensión lectora global de los participantes en función del tipo de establecimiento educacional y la complejidad del texto ($F_{(3,204)} = 8.102$; $p = 0.0001$). Habiéndose descartado la existencia de interacción entre los factores ($F_{(4,207)} = 1.046$; $p = 0.353$), se conservaron en el modelo solo los efectos principales (Tabla 3).

Los datos permiten afirmar que la complejidad del texto genera un efecto en la comprensión lectora global de los escolares ($F_{(1,204)} = 15.82$; $p = 0.0001$). Estas diferencias en el rendimiento lector global dependen de esta variable con independencia del tipo

de establecimiento educacional, sin que se observe una interacción entre ambos factores. Específicamente, se obtiene como resultado que el texto de alta complejidad fue más difícil de comprender para los estudiantes que el texto de baja complejidad textual, en el cual logran un mejor rendimiento global en la prueba de comprensión ($t = 4.014$; $p = 0.001$) como se muestra en la Tabla 4.

Homólogamente, el factor tipo de establecimiento educacional genera un efecto sobre la comprensión global del estudiante ($F_{(2,204)} = 3.87$; $p = 0.022$). Al utilizar el test de Levene para examinar la homogeneidad de la varianza de la variable comprensión global, se constata que es igual a lo largo de todos los grupos ($F_{(5,202)} = 0.44$; $p = 0.82$), lo que permitió realizar la prueba *post hoc* de Bonferroni,

TABLA 3

ANOVA de los efectos del tipo de colegio y dificultad del texto sobre la comprensión lectora global de los estudiantes

Origen	SC tipo III	gl	MC	F	p
Modelo corregido	141.984 ^a	3	47.328	8.102	0.0001
Intersección	11808.583	1	11808.583	2021.426	0.0001
Complejidad del texto	92.428	1	92.428	15.822	0.0001
Tipo de establecimiento	45.215	2	22.608	3.87	0.022
Error	1191.709	204	5.842		
Total	13154.000	208			
Total corregida	1333.692	207			

R cuadrado = 0.106 (R cuadrado corregida = 0.093).

Fuente: elaboración propia.

TABLA 4

Estadísticos descriptivos de la comprensión lectora total según dificultad del texto

Complejidad del texto	N	M	DE	Error típ. de la media
Texto de baja complejidad	106	8.21	2.521	0.245
Texto de alta complejidad	102	6.84	2.375	0.235

Fuente: elaboración propia.

TABLA 5

Estadísticos descriptivos de la comprensión lectora global según tipo de establecimiento

Tipo de establecimiento	N	M	DE	Error típ. de la media
Municipal	73	7.36	2.568	0.301
Part. Subvencionado	74	7.11	2.447	0.284
Part. Pagado	61	8.28	2.491	0.319

Fuente: elaboración propia.

cuyo resultado llevó a encontrar diferencias en la media de comprensión lectora global entre los establecimientos particular subvencionado y particular pagado ($p = 0.017$). Este resultado es consistente con los encontrados con la prueba de Scheffe y DHS de Tukey. No se advierten diferencias entre los demás pares de establecimientos, como se aprecia en la Tabla 5.

Con el propósito de estudiar el efecto de los factores complejidad del texto y tipo de establecimiento en la comprensión textual, comprensión pragmática y comprensión crítica, se realizó un diseño multivariado de la varianza (MANOVA), cuyo resultado confirmó la ausencia de interacción entre estos factores (Lamdda de Wilks = 1.112; $p = 0.354$), lo que permitió construir un modelo considerando solo los efectos principales. Los datos

muestran que tanto la comprensión textual ($F_{(3,204)} = 7.302$; $p = 0.0001$) como la comprensión pragmática ($F_{(3,204)} = 5.521$; $p = 0.001$) presentan variaciones atribuibles principalmente a la complejidad del texto al cual fueron expuestos los participantes, con cierto efecto del tipo de establecimiento sobre la comprensión textual ($F_{(2,204)} = 3.052$; $p = 0.049$). No obstante, no se observa un efecto sobre la comprensión crítica atribuible a la complejidad del texto o al tipo de establecimiento educacional ($F_{(3,204)} = 0.769$; $p = 0.513$), como se aprecia en la Tabla 6.

Con claridad, los datos evidencian que la alta complejidad del texto produce una disminución de la comprensión textual y la comprensión pragmática; ningún efecto se aprecia, en cambio, sobre la comprensión crítica, donde se obtienen rendimientos similares y muy bajos, pese a la diferencia

TABLA 6
MANOVA de los efectos del tipo de establecimiento y complejidad del texto sobre las tres dimensiones de la comprensión lectora de los estudiantes

Origen	Variable dependiente	Suma de cuadrados tipo III	gl	Media cuadrática	F	p
Modelo corregido	Comp. Textual	98.111 ^a	3	32.704	7.302	0
	Comp. Pragmática	3.714 ^b	3	1.238	5.521	0.001
	Comp. Crítica	0.811 ^c	3	0.27	0.769	0.513
Intersección	Comp. Textual	7911.131	1	7911.131	1766.384	0
	Comp. Pragmática	77.556	1	77.556	345.88	0
	Comp. Crítica	119.162	1	119.162	339.112	0
Complejidad texto	Comp. Textual	68.154	1	68.154	15.217	0
	Comp. Pragmática	2.745	1	2.745	12.243	0.001
	Comp. Crítica	0.089	1	0.089	0.254	0.615
Tipo de establecimiento	Comp. Textual	27.337	2	13.669	3.052	0.049
	Comp. Pragmática	0.813	2	0.407	1.814	0.166
	Comp. Crítica	0.733	2	0.366	1.043	0.354
Error	Comp. Textual	913.658	204	4.479		
	Comp. Pragmática	45.743	204	0.224		
	Comp. Crítica	71.685	204	0.351		
Total	Comp. Textual	8938.000	208			
	Comp. Pragmática	127.000	208			
	Comp. Crítica	191.000	208			
Total corregida	Comp. Textual	1011.769	207			
	Comp. Pragmática	49.457	207			
	Comp. Crítica	72.495	207			

a. R cuadrado = 0.097 (R cuadrado corregida = 0.084)

b. R cuadrado = 0.075 (R cuadrado corregida = 0.061)

c. R cuadrado = 0.011 (R cuadrado corregida = -0.003)

Fuente: elaboración propia.

de complejidad del texto, como se aprecia en la Tabla 7.

Al examinar la diferencias de medias en la comprensión textual según tipo de establecimiento educacional utilizando la prueba DHS de Tukey, se encuentran diferencias significativas ($p = 0.032$) entre los colegios particular subvencionado ($M = 5.81$; $DE = 2.105$) y particular pagado ($M = 6.74$; $DE = 2.097$), resultados que se confirman con otras pruebas *post hoc* como la de Scheffe y Bonferroni.

Discusión

Los hallazgos de este estudio permiten comprobar, en primer lugar, que la complejidad textual incide en el rendimiento lector de los escolares en tareas de comprensión. Esto se sustenta en que los análisis estadísticos realizados proporcionan resultados robustos y consistentes en los que se establece una relación inversamente proporcional entre las propiedades estructurales del material verbal y los resultados, esto es, un texto con menor complejidad textual permite un aumento del rendimiento en tareas de comprensión. Por el contrario, a mayor complejidad textual se obtiene un menor resultado lector en la prueba de comprensión. De estos hallazgos se concluye que las propiedades textuales manipuladas corresponden efectivamente a factores lingüísticos relevantes para la comprensión. En congruencia con los supuestos y antecedentes teóricos considerados, la carga informativa del vocabulario, la complejidad sintáctica y la densidad proposicional inciden en el rendimiento lector constituyendo, por sí mismos y en conjunto, un valor que permite anticipar el comportamiento de

los sujetos. En otras palabras, es posible predecir con cierto alcance la mayor dificultad o facilidad con la que un texto puede ser leído y comprendido observando los componentes verbales antes mencionados (Graesser et al., 2004; McNamara et al., 2002; Snow, 2002).

En segundo lugar, no obstante, el valor predictor de los indicadores de complejidad lingüística para efectos de anticipar el rendimiento lector debe ser relativizado en función de las características del lector y de las tareas de comprensión (Kintsch & Miller, 1981; Snow, 2002). Las evidencias proporcionadas por esta investigación son claras en tanto muestran diferencias significativas asociadas a la variable tipo de establecimiento educacional, hecho que no debe sorprender, toda vez que existen abundantes antecedentes de que los alumnos de establecimientos privados, provenientes de los grupos sociales de mayor ingreso económico, donde los padres tienen en promedio un nivel educacional muy superior y transmiten un mayor capital cultural a sus hijos, obtienen mejores resultados en evaluaciones de diversa índole, entre las que destacan en Chile la prueba de medición de la calidad de la educación (SIMCE) y la prueba PISA de la OCDE (Brunner & Elacqua, 2003; Contreras, Kluttig, Espinoza, Ugarte & Rubio, 2012; Eyzaguirre & Fontaine, 2008; Fontaine, 2002; Ministerio de Educación de Chile, 2012). Del mismo modo, las diferencias atribuidas al efecto de las variables lingüísticas determinantes de una mayor complejidad (y dificultad intrínseca) del texto ven anulado su efecto ante la tarea de comprensión crítica. Todo indica que las operaciones cognitivas implicadas aquí representan una dificultad mayor para los escolares, tanto que muchos fracasan, con independencia de las carac-

TABLA 7

Estadísticos descriptivos de las dimensiones de la comprensión lectora según complejidad del texto

Complejidad del texto	N	Comprensión					
		Textual		Pragmática		Crítica	
		M	DE	M	DE	M	DE
Texto de baja Complejidad	106	6.75	2.169	0.73	0.448	0.74	0.637
Texto de alta Complejidad	102	5.58	2.103	0.49	0.502	0.77	0.543

Fuente: elaboración propia.

terísticas del texto o si se trata de estudiantes de establecimientos públicos o privados.

Un aspecto de los resultados que llama la atención es el hecho de constatar que los ítems correspondientes a la dimensión Comprensión Crítica no arrojan diferencias significativas entre los grupos. La explicación para ello radica en la naturaleza de las operaciones cognitivas requeridas por las tareas de comprensión incluidas en este tipo de reactivos. La exigencia mayor que ello conlleva tiene como consecuencia un bajo rendimiento en un grado tal que las características diferenciadoras de los lectores se anulan, reforzando así la hipótesis que guió la investigación, esto es, que la lecturabilidad depende de varios factores.

En suma, los hallazgos del estudio permiten sostener que la lecturabilidad, observada desde una perspectiva psicolingüística multifactorial, adquiere un carácter más complejo y, por lo mismo, más interesante como objeto de investigación. Los alcances de este trabajo son limitados, en tanto el diseño utilizado no hace posible despejar la incógnita de cuál es el peso específico de cada uno de los factores lingüísticos y cómo ellos se vincularían con aspectos también más específicos de las tareas de comprensión. Ello, no obstante, permanece como un desafío para nuevas investigaciones.

Referencias

- Aiken, L. (2003). *Test psicológicos y evaluación* (11.a ed.). México: Pearson Educación.
- Álvarez, C., Alameda, R. & Domínguez, A. (1999). El reconocimiento de las palabras: procesamiento ortográfico y silábico. En M. de Vega & F. Cuetos (Eds.), *Psicolingüística del español* (pp. 231-270). Madrid: Trotta.
- Anastasi, A. & Urbina, S. (1998). *Test psicológicos* (7.a ed.). México: Prentice Hall.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, 4(11), 417-423.
- Baddeley, A. D. & Hitch, G. J. (1974). Working memory. En G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47-89). New York: Academic Press.
- Baker, M. (1995). Corpora in translation studies: An overview and suggestions for future research. *Target*, 7(2), 223-244.
- Brunner, J. & Elacqua, G. (2003). *Capital Humano en Chile* (Informe). Santiago, Chile: Universidad Adolfo Ibáñez, Escuela de Gobierno. Recuperado de http://www.oei.es/etp/informe_capital_humano_chile_brunner.pdf
- Contreras, D., Kluttig, M., Espinoza, P., Ugarte, A. & Rubio, M. (2012). *El capital cultural familiar y su influencia sobre las habilidades lectoras*. Santiago, Chile: Fondo de Investigación y Desarrollo en Educación. Recuperado de http://www.agencia-educacion.cl/wp-content/files_mf/elcapitalculturalfamiliarysuinfluencia55.pdf
- de Vega, M. (1985). *Introducción a la psicología cognitiva*. Madrid: Alianza.
- de Vega, M., Carreiras, M., Gutiérrez-Calvo, M. & Alonso-Quecuty, M. (1990). *Lectura y comprensión: una perspectiva cognitiva*. Madrid: Alianza.
- Du Bay, W. (2004). *The Principles of readability*. Costa Mesa, CA: Impact Information. Recuperado de <http://www.impact-information.com/impactinfo/readability02.pdf>
- Echeverría, M. (1978). *Desarrollo de la comprensión infantil de la sintaxis española*. Concepción, Chile: Universidad de Concepción.
- Eyzaguirre, B. & Fontaine, L. (2008). *Las escuelas que tenemos*. Santiago, Chile: Centro de Estudios Públicos. Recuperado de http://www.cepchile.cl/dms/lang_1/doc_4167.html#.UW9nlhyvFhQ
- Fontaine, A. (2002). *Equidad y calidad de la educación: cinco proposiciones interrelacionadas* (Nº 87). Santiago, Chile: Centro de Estudios Públicos. Recuperado de http://www.cepchile.cl/1_3137/doc/equidad_y_calidad_de_la_educacion_cinco_proposiciones_interrelacionadas.html#.UW9uBRyvFhQ
- Goetz, E. T., Anderson, R. C. & Schallert, D. L. (1981). The representation of sentences in memory. *Journal of Verbal Learning and Verbal Behavior*, 20(4), 369-385.
- Graesser, A., McNamara, D., Lowerse M. & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202.

- Halliday, M. A. K. (1985). *Spoken and written language*. Oxford: Oxford University Press.
- Hunt, K. W. (1970). How little sentences grow into big ones. En M. Lester (Ed.), *Readings in applied transformational grammar* (pp. 193-201). New York: Holt, Rinehart and Winston.
- Johansson, V. (2008). Lexical diversity and lexical density en speech and writing: A developmental perspective. En *Working Papers* (Vol. 53, pp. 61-79). Paradisgatan, Suecia: Lund University.
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329-354.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: University Cambridge Press.
- Kintsch, W. & Keenan, J. (1973). Reading rate as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, 5, 257-274.
- Kintsch, W., Kozminsky, E., Streby, W., McKoon, F. & Keenan, J. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 14, 196-214.
- Kintsch, W. & Miller, J. (1981). Readability: A view from cognitive psychology. En *Teaching: Research reviews* (pp. 220-232). Newark, DE: International Reading Assn.
- Kintsch, W. & Rawson, K. A. (2007). Comprehension. En M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209-226). Malden, MA: Blackwell.
- Ministerio de Educación de Chile. (2012). *Orientaciones e instrumentos de evaluación diagnóstica inicial, intermedia y final en comprensión lectora. 1.º año de educación media*. Santiago, Chile: Autor. Recuperado de <http://www.mineduc.cl/usuarios/media/doc/201210161055470.CLectoralMedio.pdf>
- McNamara, D. S., Louwerse, M. M. & Graesser, A. C. (2002). *Coh-Matrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension* (Technical report). Memphis, TN: University of Memphis, Institute for Intelligent Systems.
- McNamara, D. & Shapiro, A. (2005). Multimedia and hipermedia solutions for promoting metacognitive engagement, coherence, and learning. *Journal Educational Computing Research*, 33(1), 1-29.
- Nunnally, J. C. & Bernstein, I. J. (1999). *Teoría psicométrica* (3.a ed.). México: McGraw-Hill.
- Parodi, G. (1986). Validación y desarrollo del gráfico de Fry para el castellano hasta el nivel universitario. *Signos*, 19(24), 125-132.
- Perfetti, C., Landi, N. & Oakhill, J. (2005). The acquisition of reading comprehension skill. En M. Snowling & C. Hulme (Eds.), *The science of reading. A handbook* (pp. 227-247). Oxford: Blackwell.
- Ratcliff, R. & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior*, 17, 403-417.
- Autor & Autor (2011)
- Snow, C. (2002). *Reading for understanding. Toward and R & D program in reading comprehension*. Santa Monica, CA: RAND.
- Turner, A. & Greene, E. (1978). *Construction and use of a propositional text base* (Informe técnico N° 87-02). Boulder, CO: Institute of Cognitive Science. Recuperado de <http://www.colorado.edu/ics/sites/default/files/attached-files/87-02.pdf>
- van Dijk, T. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Autor (1988)
- Autor (1999)
- Autor & Autor (1992)
- Autor & Autor (1993)
- Zakaluk, B. L. & Samuels, S. J. (Eds.). (1988). *Readability: Its past, present, and future*. Newark, DE: International Reading Association.
- Zwaan, R. & Singer, M. (2003). Text comprehension. En A. Graesser, M. A. Gernsbacher & S. Goldman (Eds.), *Handbook of discourse processes* (pp. 83-121). Mahwah, NJ: Erlbaum.