

LÉXICO ORTOGRÁFICO Y COMPETENCIA LECTORA

ORTHOGRAPHIC LEXICON AND READING ABILITY

Javier Marín

Universidad de Murcia (España)

Ariel Cuadro

Universidad Católica del Uruguay

Ascensión Pagán

Universidad de Murcia (España)

Resumen: Se presenta un estudio en el que se analizan las relaciones entre la lectura eficiente y la disponibilidad de representaciones ortográficas léxicas en dos muestras de estudiantes universitarios españoles y uruguayos. Mediante el uso de una metodología de análisis de regresión tratamos de explicar los resultados de una prueba de lectura eficiente (TECLE) a partir de una prueba de conocimiento ortográfico (TECOR). Los resultados permiten confirmar el importante papel que juega el léxico ortográfico en la lectura eficiente en español. Esta relación se revela más intensa en la muestra uruguaya que en la española. El patrón de resultados se interpreta en el marco de los efectos de la transparencia ortográfica en el aprendizaje de la lectura y el procesamiento lector.

Palabras clave: Ortografía, Lectura, Transparencia ortográfica, aprendizaje de la lectura, Representaciones ortográficas

Abstract: Relation between efficient reading and availability of lexical orthographic representations is analyzed in Uruguayan and Spanish college students. Using regression analysis methodology we try to explain an assessment of efficient reading (TECLE) with orthographic knowledge scores (TECOR). It is demonstrated the important role of orthographic knowledge in efficient reading. This relation is greater in the Uruguayan sample than in the Spanish. This data is interpreted within the frame of studies about orthographic depth in literacy learning and reading processing.

Keywords: Orthography, Reading, Orthographic depth, literacy learning, orthographic representations

INTRODUCCIÓN

La habilidad para acceder de forma eficiente al contenido de los textos es una capacidad que resulta ineludible para los ciudadanos de la sociedad moderna. La capacidad para leer de forma eficiente muestra su importancia por primera vez en la escuela, donde los alumnos que carecen de unos mecanismos lectores eficaces se ven abocados frecuentemente al fracaso escolar. Unos mecanismos de procesamiento lector eficientes proveen al estudiante de un acceso a los contenidos de los textos sin que ello importe un costo cognitivo que suponga una merma importante para la acción de los meca-

nismos de procesamiento semántico (comprensión) o de almacenamiento. Los alumnos que carecen de esta habilidad, aunque ocasionalmente son capaces de leer oralmente un texto con cierta corrección, no muestran un nivel adecuado de comprensión del mismo, posiblemente debido a que el coste cognitivo de la descodificación ha limitado gravemente los recursos cognitivos disponibles para la acción de los mecanismos semánticos (Vellutino, Fletcher, Snowling y Scanlon, 2004; Vellutino, 1979).

La mayor parte de los modelos actuales de reconocimiento de palabras (véase Jacobs y Grainger, 1994 para una revisión) incluyen un mecanismo de acceso directo a la información

ortográfica léxica a partir del input visual. Uno de los modelos más extendidos, el conocido como modelo de doble ruta (Coltheart, Curtis, Atkins y Haller, 1993; Coltheart, Rastle, Perry, Langdon y Ziegler, 2001), ha establecido la denominación de este tipo de acceso como "acceso directo" o "ruta directa". El acceso directo se diferencia de su alternativa, el "acceso fonológico", en que en éste incluye la conversión de la ortografía en fonología (por medio de las reglas de conversión propias de cada lengua) como paso previo y necesario al acceso al significado o pronunciación de la palabra. En lenguas con ortografía transparente, como el español, la lectura por ruta fonológica puede garantizar la lectura (pronunciación) correcta de cualquier palabra escrita. Sin embargo, el uso de esta ruta implica un alto coste cognitivo, por lo que en lectores que hacen uso intensivo de este procedimiento (en detrimento del procedimiento directo) muestran importantes déficit en la comprensión de los textos que leen. El uso de una u otra ruta depende de la disponibilidad de conocimientos ortográficos en el léxico individual. Estos conocimientos, en condiciones normales, se incorporan al léxico ya desde el inicio mismo del aprendizaje de la lectura y la consiguiente adquisición del principio alfabético (Ehri y Wilce, 1985).

David Share ha formulado la hipótesis del autoaprendizaje (Jorm y Share, 1983; Share, 1995) acerca del proceso de adquisición de la habilidad lectora. En el trabajo del año 1995 presenta una revisión de la evidencia empírica generada en las últimas 4 décadas, donde se apoya la idea de que el tipo de lectura eficaz, que caracteriza a los lectores avanzados, utiliza intensivamente el acceso a la información ortográfica almacenada en el léxico ortográfico. La disponibilidad de esta información permite un acceso rápido y con muy bajo costo cognitivo, tanto al significado de las palabras escritas, como a su fonología. El postulado central de esta hipótesis afirma que, "*cada descodificación exitosa de una palabra no familiar ofrece una oportunidad para adquirir la información ortográfica específica de la palabra que está en la base del reconocimiento de palabras por lectores expertos*" (Share, 1995, p. 155). Según esta formulación la adquisición de nuevas representaciones ortográficas depende de dos factores: el número de exposiciones a un determinado ítem y el éxito de la descodificación. El primer factor predice que aquellas palabras que aparecen

más a menudo en los textos generarán con mayor probabilidad representaciones lexicales en el almacén ortográfico del sujeto. El segundo factor predice que aquellos sujetos que por alguna razón (p.ej., padecer un trastorno disléxico) resulten menos exitosos en la descodificación de las palabras dispondrán de un menor conocimiento ortográfico comparados con aquellos que presenten menos problemas a ese respecto (lectores normales).

Algunos autores han planteado que las diferencias en transparencia ortográfica existentes entre las diferentes lenguas podrían generar diferencias en el aparato cognitivo que desarrollan los aprendices de la lectura (Frost y Katz, 1992; Katz y Frost, 1992). En particular han propuesto que el desarrollo del léxico ortográfico (y la lectura por ruta directa) estará más avanzado en los países donde exista un sistema ortográfico más opaco en comparación con aquellos en los que el sistema sea más transparente. La razón principal de esta propuesta es que en los sistemas transparentes el uso de las reglas ortográficas es suficiente para acceder a la pronunciación (y al significado) de las palabras, mientras que en el caso opaco, debido a la presencia de un gran número de excepciones a las reglas ortográficas, se hace necesario el acceso a información específica relacionada con cada ítem. Los resultados de la aplicación de esta propuesta, lejos de ser confirmados, en algunos casos han arrojado resultados contrarios a los predichos. Por ejemplo, en Marín, Alegría y Carrillo (remitido) hemos mostrado que el desarrollo del léxico de los escolares españoles en sus primeros años de aprendizaje de la lectoescritura es mucho más acelerado que el de los escolares francófonos. La explicación que sugerimos para este fenómeno es precisamente la sencillez del código ortográfico del español frente al francés, que resulta más complejo, opaco y con frecuentes excepciones. Esta sencillez facilita la descodificación exitosa de las palabras y, por aplicación de la propuesta de Share (1995, 2004), conlleva la generación de una mayor cantidad de conocimiento ortográfico en períodos similares.

Con respecto al efecto de la descodificación fonológica y su efecto en la adquisición de conocimiento ortográfico, hemos realizado algunos trabajos en los que hemos mostrado que existen factores ortográficos específicos, relacionados con la naturaleza de las reglas de conversión de la ortografía en fonología, que pue-

den interferir con la generación de dicho conocimiento (Marín, Aveledo, Pagán y Cuadro, 2007; Marín, Pagán, Cuadro y Aveledo, 2007). En particular, hemos mostrado que la presencia de ciertos grafemas consonánticos que pueden referir más de un fonema interfiere con el almacenamiento ortográfico. Por ejemplo, el grafema 'g' presenta cierta ambigüedad ya que puede representar el fonema /G/ o el fonema /÷/ dependiendo del contexto subsiguiente. Para el caso latinoamericano, existe otro ejemplo de ambigüedad que se debe a la desaparición del fonema /è/ y su sustitución por el fonema /S/ que ha sucedido en estos países. Esto hace que las palabras "siento" y "ciento" sean homófonas en Latinoamérica y que el grafema "c", en la ortografía del español hablado en estos países, presente un caso de ambigüedad similar al descrito para "g". Hemos mostrado que las palabras que contienen el grafema "g" (España y Latinoamérica) y el grafema "c" (Latinoamérica) disponen de representaciones ortográficas más débiles que sus controles, escritos con los grafemas "b" y "v" (Marín et al., 2007; Marín, Aveledo, Vera y Alegría, 2004b). Katz y Frost (2001) han encontrado resultados similares en lengua inglesa, mostrando que se trata de un fenómeno extensible a diferentes ortografías.

Philip Seymour y sus colegas (2003; Seymour, 2005) han publicado los resultados de un amplio proyecto de investigación (COST) en el que se ha comparado el aprendizaje de la lectoescritura en las diferentes lenguas europeas. Uno de los resultados más impactantes muestra que existen grandes diferencias entre diferentes lenguas respecto a la velocidad de adquisición de la capacidad lectora, siendo las lenguas transparentes como el finlandés o el español las que mejores resultados obtienen. En Marín, Aveledo et al., (2007; Marín, Pagán et al., 2007) hemos sugerido que la presencia de diferentes densidades de grafemas ambiguos en las diferentes ortografías podría explicar parcialmente las diferencias observadas en el proceso de adquisición de la lectura, habida cuenta de que éste consiste en gran medida en el desarrollo del léxico ortográfico (Ehri y Wilce, 1985; Share, 1995).

Uno de los problemas que enfrenta el proyecto COST (Seymour, 2004) es la gran heterogeneidad de los sistemas ortográficos que comparan. Aunque es indudable el interés de sus resultados, resulta difícil interpretar cuáles son las causas de las diferencias observadas

entre lenguas, ya que son muy numerosos los factores que las distinguen. Una atractiva posibilidad para estudiar el efecto de la transparencia ortográfica sobre el procesamiento lector en un contexto más simple consiste en estudiar las posibles diferencias entre el español peninsular y el latinoamericano. En este caso, aunque ciertamente se pueden señalar diferencias en el aspecto morfológico, léxico, pragmático, fonológico y sintáctico; estas diferencias no son tan importantes como las que existen entre lenguas más parecidas (p.ej., español e italiano o holandés y alemán) del estudio COST. El mayor grado de similitud que puede obtenerse al permanecer dentro de una misma lengua reduce enormemente la complejidad de las comparaciones y facilita la interpretación de los posibles resultados.

En el presente trabajo mostramos los resultados de un estudio en el que pretendemos estimar la contribución del desarrollo ortográfico a la lectura eficiente en muestras de estudiantes universitarios. Comparamos los resultados obtenidos en una muestra española y uruguaya con el objetivo de estudiar las posibles diferencias entre grupos y asignarles una interpretación en el marco del estudio de los efectos de las propiedades de la ortografía sobre el perfil cognitivo de la capacidad lectora.

EXPERIMENTO

Uno de los aspectos que establecen el grado de transparencia/opacidad de un sistema ortográfico es la complejidad del sistema de reglas de conversión de la ortografía en fonología. La regla más simple es la que establece una correspondencia bidireccional y única entre un grafema simple y un fonema. Este sería el caso de la gran mayoría de las reglas ortográficas del español. Un ejemplo lo tenemos con la regla /t/ \leftrightarrow "t", por la que se expresa que el fonema /t/ siempre se transcribe con el grafema "t" y que el grafema "t" siempre se convierte en el fonema /t/ durante la descodificación lectora. Cualquier desviación de este principio unívoco y simple supone un incremento en la complejidad del sistema. Por ejemplo, tenemos reglas del tipo /b/ \leftrightarrow {"b", "v"}, por la que un mismo fonema puede transcribirse de dos formas diferentes, aunque éstas siempre se leen del mismo modo. Un caso aún más complejo es /x/ (e, i) \leftrightarrow {"g", "j"} (e, i); en el que en función de unos parámetros contextuales (la pre-

sencia de las vocales “e” o “i”) la transcripción del fonema /x/ puede efectuarse alternativamente con los grafemas “g” o “j”. Esta complejidad se ve incrementada cuando tenemos en cuenta la regla /G/ ↔ “g”(a, o, u, consonante); que, en el caso de que no tengamos en cuenta el contexto, podría llevarnos a un caso de ambigüedad en la conversión fonológica: {/G/, /x/} ↔ “g”. Como ya hemos relatado en la introducción, la presencia de estas complejidades introduce un componente de dificultad que afecta a la eficiencia del almacenamiento ortográfico. Este fenómeno se extiende al caso de la regla /S/ (e, i) ↔ {“c”, “s”}(e, i), operativa en la ortografía del español hablado en los países latinoamericanos y que resulta funcionalmente idéntica al caso del grafema “g” (Marín, Aveledo et al., 2007; Marín, Pagán et al., 2007). Debido a la alta frecuencia de palabras que contienen la secuencia /se-/si/, entendemos que existe una diferencia en el grado de transparencia entre el español latinoamericano y el peninsular.

En la Tabla 1 mostramos un recuento realizado sobre el corpus LEXESP (Sebastián, Martí, Carreiras y Cuetos, 2000) de la frecuencia type y token de los segmentos fonológico/ortográficos que hemos referido como ambiguos. Como se puede ver en la tabla, existe un alto porcentaje de palabras que contienen los segmentos /se-/si/, lo que indica que los escritores latinoamericanos afrontan permanentemente el problema de la transcripción de segmentos ambiguos. Casi una cuarta parte (23.23%) de las palabras

registradas en LEXESP contiene una ambigüedad en la transcripción del fonema /S/. Esta circunstancia no opera para el caso peninsular, en el que esas palabras no revisten una dificultad especial para su transcripción. Nuestro objetivo es probar hasta qué punto esta diferencia puede introducir cambios en el perfil cognitivo de los lectores, en un sentido similar al observado cuando se comparan sujetos de diferentes lenguas.

MÉTODO

1.1. Materiales

Prueba de lectura TECLE

Para medir la competencia lectora se ha utilizado el test TECLE (Marín y Carrillo, 1999). Esta prueba ha sido desarrollada por nosotros y ha demostrado su eficacia a la hora de evaluar el nivel lector tanto en muestras españolas (Marín, Alegría y Carrillo, remitido; Marín, Aveledo, Vera y Alegría, 2004b) como uruguayas (Cuadro, 2006). Está compuesta de un total de 64 ítems que contienen una frase encabezado y cuatro opciones de respuesta. A la frase encabezado le falta la última palabra, que debe ser seleccionada entre las cuatro opciones de respuesta. Las tres opciones falsas son dos pseudopalabras que constituyen pares mínimos ortográfico y fonológico de la respuesta correcta, y por último, una palabra que resulta inadecuada desde el punto de vista semántico

TABLA 1.
Frecuencias type y token (frecuencia por millón) de las palabras que contienen diferentes segmentos subléxicos. Los valores absolutos se indican primero, a continuación y entre paréntesis los valores relativos al total del corpus.

Fonema	Grafema	Silaba	Frecuencia Type	Frecuencia Token
/S/			30 742 (23.23%)	143 512 (14.35%)
	“c”		16 160 (12.21%)	69 723 (6.97%)
		“CE”	4 831 (3.65%)	19 836 (1.98%)
		“CI”	11 329 (8.56%)	49 887 (4.99%)
	“s”		14 582 (11.02%)	73 789 (7.38%)
		“SE”	8 696 (6.57%)	43 059 (4.31%)
“SI”		5 886 (4.45%)	30 731 (3.07%)	
/X/			5 711 (4.32%)	16 777 (1.68%)
	“g”		4 105 (3.10%)	11 052 (1.11%)
		“GE”	1 827 (1.38%)	5 413 (0.54%)
		“GI”	2 278 (1.72%)	5 638 (0.56%)
	“j”		1 606 (1.21%)	5 725 (0.57%)
		“JE”	1 229 (0.93%)	5 218 (0.52%)
“JI”		377 (0.28%)	507 (0.05%)	

y/o sintáctico para completar la frase. Por tanto, el evaluado necesita ser capaz de discriminar entre sutiles cambios fonológicos, ortográficos y semántico/sintácticos para poder contestar adecuadamente. La prueba está cronometrada, por lo que la selección de alternativas debe producirse con el menor consumo de tiempo disponible. Para muestras de lectores avanzados utilizamos un tiempo total de 3 minutos, que normalmente impide que el 95% de los evaluados no alcance el final de la prueba.

La puntuación en la prueba es el número de respuestas correctas emitidas.

Prueba de ortografía TECOR

Esta prueba ha sido desarrollada específicamente para este trabajo y consiste en un conjunto de pares de palabra/pseudopalabra que resultan homófonos. Para ello utilizamos palabras que contienen los grafemas “g”/“j” (seguidos de “e” ó “i”), “v”/“b” y, para la versión latinoamericana, palabras escritas con “c”/“s” seguidas de “e” ó “i”.

Cada grafema estaba representado un total de 24 veces, de las cuales 12 se correspondían con palabras de baja frecuencia y 12 con palabras de alta frecuencia. Se seleccionaron los ítems de tal forma que los promedios de frecuencia para cada uno de los subgrupos grafema x frecuencia no resultaran estadísticamente diferentes para niveles iguales de frecuencia (alta, baja) y alcanzaran la significación estadística para niveles diferentes. Esta estructura nos permite obtener una medida del efecto de la frecuencia independiente para cada grafema y estudiar las posibles variaciones idiosincráticas asociadas a cada uno de ellos.

La versión española contenía los ítems con “b”, “v”, “g” y “j”, lo que supone un total de $24 \times 4 = 96$ ítems. La versión latinoamericana incorporó los ítems para “s” y “c”, lo que arroja una cantidad total de 144 ítems en esta versión. Se añadieron ítems de control formados por parejas de pseudopalabras (4 a la versión española, 6 a la latinoamericana) lo que convierte a la versión española en una prueba con 100 ítems y a la latinoamericana en una prueba con 150 ítems. Los ítems de control informan de si el sujeto sigue correctamente las instrucciones de la prueba —y consiguientemente marca como “desconocido” estos ítems, o está desarrollando un protocolo de respuestas contaminado por la adivinación al azar.

La tarea del sujeto consiste en identificar la secuencia que está mal escrita (la pseudopalabra) para aquellas palabras cuyo significado conozca de forma aproximada (ésto excluye a los ítems de control). La puntuación en la prueba es el porcentaje de aciertos obtenido para cada uno de los subgrupos grafema (b, v, g, j, c, s) x nivel de frecuencia (alto, bajo).

1.2. Sujetos

En el estudio español se emplearon un grupo de 150 estudiantes de 4º curso de la licenciatura de psicología de la universidad de Murcia. En el estudio latinoamericano se utilizaron 60 estudiantes uruguayos de primer curso de la licenciatura de Psicología en la universidad católica del Uruguay. Todos los participantes afirmaron ser nativos de los respectivos países y haber seguido en su totalidad el proceso de escolarización en las instituciones del país.

1.3. Procedimiento

Las pruebas TECLE y TECOR fueron aplicadas en el aula. Primero se administró la prueba TECLE con un tiempo de 3 minutos. A continuación se aplicó la prueba TECOR, con un tiempo de 10 minutos para la versión española y de 15 para la latinoamericana. Esta prueba constaba de dos formas paralelas en las que los ítems habían sido ordenados de forma diferente para evitar que algunos participantes pudieran copiar sus respuestas de otros compañeros.

RESULTADOS

Se estableció un criterio de exclusión de los participantes basado en la tasa de aciertos obtenida en los ítems de control y se excluyeron todos aquellos con un porcentaje de aciertos inferior al 20%. Esto supuso la eliminación de 6 sujetos en la muestra uruguaya (10% del total), restando un total de 54 participantes. De igual forma en la muestra española se eliminaron 3 participantes (2%) que no cumplieron con los criterios, quedando la muestra reducida a un total de 147 participantes. Los resultados de la muestra seleccionada en las pruebas TECLE y TECOR constituyen el conjunto de datos que someteremos a análisis. En la Tabla 2 y Tabla 3 se muestran los promedios y desviaciones típicas para cada uno de los subgrupos de grafema x frecuencia y nivel lector.

TABLA 2.
Resultados de las pruebas TECLE y TECOR en la muestra Española.
Los promedios se muestran primero y entre paréntesis las desviaciones típicas.

		España			
		Nivel Lector			
		Bajo	Medio	Alto	
N		53	52	42	Promedio
Lectura (TECLE)		41.40 (3.64)	49.87 (1.83)	57.24 (3.60)	49.50
Grafema	Frecuencia				Total
B	Alta	95.19% (14.88)	90.20% (22.40)	97.62% (10.78)	94.34%
	Baja	76.92% (22.67)	84.84% (17.79)	89.48% (12.35)	83.75%
V	Alta	83.33% (23.85)	91.99% (15.74)	94.44% (10.84)	89.92%
	Baja	70.33% (24.92)	80.22% (14.04)	86.99% (13.99)	79.18%
G	Alta	91.84% (19.87)	94.77% (12.24)	90.08% (20.51)	92.23%
	Baja	83.76% (19.70)	87.26% (19.03)	88.39% (18.61)	86.47%
J	Alta	83.96% (30.66)	93.27% (19.87)	95.24% (18.51)	90.82%
	Baja	68.11% (28.47)	67.79% (25.25)	63.69% (24.34)	66.53%
	Promedio	81.68	86.29	88.24	

TABLA 3.
Resultados de las pruebas TECLE y TECOR en la muestra Uruguaya.
Los promedios se muestran primero y entre paréntesis las desviaciones típicas.

		Nivel Lector			
		Bajo	Medio	Alto	
N		20	18	18	
Lectura (TECLE)		36.50 (05.24)	48.06 (02.53)	56.33 (04.51)	
Grafema	Frecuencia				Total
B	Alta	79.17% (30.05)	92.59% (14.26)	96.30% (10.78)	89.35%
	Baja	75.93% (34.88)	62.96% (36.41)	59.26% (26.95)	66.05%
V	Alta	89.17% (17.33)	94.44% (12.78)	96.30% (10.78)	93.30%
	Baja	71.42% (28.11)	70.65% (24.21)	74.63% (20.06)	72.23%
G	Alta	95.00% (15.39)	94.44% (23.57)	97.22% (11.79)	95.55%
	Baja	84.53% (19.76)	90.36% (11.90)	95.29% (08.36)	90.06%
J	Alta	82.92% (19.62)	95.28% (11.44)	97.22% (08.08)	91.81%
	Baja	53.36% (25.26)	75.62% (23.14)	72.43% (16.97)	67.14%
C	Alta	79.48% (19.97)	89.68% (09.56)	90.48% (10.96)	86.55%
	Baja	76.26% (23.91)	84.11% (10.87)	83.23% (14.46)	81.20%
S	Alta	85.60% (23.18)	91.27% (17.07)	98.41% (06.73)	91.76%
	Baja	62.81% (20.52)	70.01% (23.79)	77.70% (26.26)	70.17%
Total		77.97%	84.28%	86.54%	

En primer lugar sometimos a un análisis de regresión de las puntuaciones de lectura (TE-CLE) tomando como criterio las estimaciones de conocimiento ortográfico segmentadas por grafema y frecuencia. Los resultados muestran una capacidad para explicar el 42.4% de varianza ($r^2 = 0.554$, error típico de estimación = 6.81) en el caso uruguayo mientras que en el caso español solo se explica el 15.7% ($r^2 = 0.21$, error estándar de estimación = 6.45). Es posible que la diferencia en poder explicativo se deba a que en el caso uruguayo, hemos utilizado cuatro variables predictoras más (las correspondientes a los grafemas “c” y “s” con alta y baja frecuencia) que en el caso español. Para estudiar esta posibilidad reanalizamos los datos utilizando las mismas variables que en el caso español. Los resultados de este nuevo análisis muestran que el modelo conserva la capacidad explicativa del 40.5% ($r^2 = 0.495$, error típico de estimación = 6.92) por lo que no podemos entender que la diferente capacidad explicativa de la or-

tografía sobre el nivel lector que observamos se deba a un artefacto metodológico.

Para comprender mejor cuál es la estructura explicativa de los modelos realizamos nuevos análisis utilizando la metodología del análisis de regresión por pasos. Los resultados para la muestra uruguayo se muestran en la Tabla 4. El aspecto a destacar es que las tres variables con mayor capacidad explicativa están constituidas por medidas de ítems con alta frecuencia (los grafemas B, J y V). Solo en el último modelo entra una variable de baja frecuencia (el grafema B). Un segundo aspecto a destacar es que ninguno de los grafemas que hemos rotulado como ambiguos entran en el modelo.

En la Tabla 5 se muestran los resultados del análisis de regresión por pasos en la muestra española. En este caso es importante resaltar que las variables que obtienen mayor poder predictivo son todas de baja frecuencia y nuevamente quedan aquellas que dependen de los grafemas ambiguos.

TABLA 4.
Modelos de regresión de nivel lector sobre diferentes agrupamientos de variables ortográficas en la muestra uruguayo.

Modelo	R	R2	R2 Ajustado	Error típico de estim.
1 (B_F+)	,445(a)	,198	,182	8,110
2 (B_F+, J_F+)	,564(b)	,318	,291	7,552
3 (B_F+, J_F+, V_F+)	,630(c)	,397	,361	7,172
4 (B_F+, J_F+, V_F+, B_F-)	,672(d)	,451	,406	6,911

TABLA 5.
Modelos de regresión de nivel lector sobre diferentes agrupamientos de variables ortográficas en la muestra española

Modelo	R	R2	R2 Ajustado	Error típico de estima.
1 (V_F-)	,332(a)	,111	,104	6,647
2 (V_F-, B_F-)	,400(b)	,160	,147	6,484

DISCUSIÓN

El patrón de resultados obtenido resulta ciertamente complejo. Como primer aspecto a analizar tenemos que en las dos muestras estudiadas se manifiesta una relación positiva y significativa entre la disponibilidad de representaciones ortográficas y la lectura rápida y eficiente. Aquellos estudiantes que demuestran poseer mayor número de conocimientos ortográficos también resultan más rápidos y precisos que aquellos que carecen del mismo nivel de conocimiento ortográfico. En sendos análisis discriminantes efectuados para predecir el nivel lector de los sujetos (bajo, medio, alto) encontramos que en el caso español somos capaces de clasificar correctamente más del 51% de los sujetos y en el caso uruguayo alcanzamos el 63% de clasificaciones correctas. Esto nos indica que la lectura eficiente en español tiene un componente muy importante que está basado en el acceso al léxico ortográfico y posiblemente que la carencia de un número suficiente de representaciones en el almacén puede interferir con el desarrollo de una lectura rápida y eficaz.

Un aspecto interesante de los resultados es el que atribuye mayor capacidad predictiva al conocimiento léxico en la muestra uruguaya que en la española. Este resultado puede interpretarse en el marco de las diferencias de transparencia ortográfica descritas entre el español peninsular y el latinoamericano. En línea con la propuesta de la "hipótesis de la profundidad ortográfica" (Frost y Katz, 1992; Katz y Frost, 1992) las lenguas con un código ortográfico más opaco generarán un perfil cognitivo del lector con mayor uso de la información ortográfica.

Sin embargo, este resultado debe contemplarse con precaución. Las diferencias observadas entre las muestras podrían estar causadas parcialmente por ciertos problemas metodológicos. En primer lugar, tenemos que la prueba de ortografía ha sido elaborada basándonos en un corpus peninsular. Esta circunstancia puede haber incrementado el grado de dificultad objetiva de la prueba para la muestra uruguaya ya que, debido a las diferencias léxicas entre el español peninsular y uruguayo, es posible que algunas palabras que

consideramos como de alta frecuencia según los criterios peninsulares no lo sean realmente para la muestra de estudiantes uruguayos. El caso inverso es menos probable, ya que el número de palabras infrecuentes es muy alto y la probabilidad de seleccionar por azar una palabra de baja frecuencia peninsular y alta frecuencia uruguaya es muy reducida. El segundo factor crítico está conectado con el anterior: la prueba TECOR ha resultado demasiado sencilla para una muestra de estudiantes universitarios. Su capacidad discriminativa ha quedado muy mermada por la aparición de posibles efectos techo. Esto es especialmente cierto en el caso español, donde más del 34% de los participantes obtuvieron una proporción de aciertos igual o superior al 90%, mientras que en la muestra uruguaya solo alcanzaron estas puntuaciones el 24% de la muestra.

Por último tenemos que las subpuntuaciones TECOR relacionadas con grafemas ambiguos ("g" y "c") no demuestran una capacidad predictiva comparable a la que muestran los grafemas no ambiguos ("b", "v", "j"). En un análisis de correlación entre las puntuaciones TECLÉ y las subpuntuaciones TECOR encontramos que ninguna de las puntuaciones relacionadas con un grafema ambiguo alcanzó niveles significativos en la muestra española. En la muestra uruguaya encontramos un patrón similar con la excepción de la puntuación del grafema C de alta frecuencia que alcanzó un coeficiente de correlación igual a 0.40, $p < 0.05$. Resultados anteriores de nuestro grupo han demostrado que los individuos almacenan con mayor dificultad la información ortográfica de los ítems que contienen grafemas ambiguos. Este fenómeno alcanza el extremo de que en algunas ocasiones hemos detectado un nivel de conocimiento ortográfico para estos ítems prácticamente nulo¹. Se puede interpretar que las subpuntuaciones TECOR relacionadas con grafemas ambiguos no indican conocimiento genuinamente ortográfico, sino la acción de principios estratégicos que pueden operar cuando el sujeto afronta la prueba y carece de conocimientos específicos para resolver este tipo de ítems. No obstante, las implicaciones últimas de este patrón de resultados caen fuera del propósito de este trabajo. Remitimos al lector a tra-

¹ La alta tasa de aciertos que algunos sujetos obtienen en estos ítems se debe más que a la disponibilidad de representaciones ortográficas específicas, al uso de estrategias por defecto, del tipo de las descritas en Alegria y Mousty (1996; Alegria et al., 2003).

bajos donde exponemos la “hipótesis de la ambigüedad fonológica” (Marín, Aveledo, Pagán y Pérez, 2006; Marín, Aveledo, Vera y Alegría, 2004, 2004b; Marín, Aveledo, et al., 2007; Marín, Pagán et al., 2007) para conocer con mayor profundidad algunas de las implicaciones teóricas de este resultado.

CONCLUSIONES

El principal hallazgo de este trabajo consiste en mostrar que existen diferencias en el perfil cognitivo de los lectores españoles y uruguayos. En efecto, hemos mostrado que la relación entre el conocimiento ortográfico léxico y la eficiencia de los mecanismos de la lectura es mayor en la muestra uruguaya que en la española. Esta variación en el perfil cognitivo de los lectores de dos países con una lengua con un gran número de elementos en común es un fenómeno cuya explicación última no puede abordarse a partir de los datos correlacionales disponibles en este trabajo. Sin embargo si es posible formular un marco teórico adecuado donde encuadrar este fenómeno.

En primer lugar, es preciso reconocer la posibilidad de que las diferencias socioculturales existentes entre las dos poblaciones de referencia podrían estar en la base del fenómeno observado. Estas diferencias afectan a las características del sistema escolar y de enseñanza reglada: la organización del currículo, los métodos y materiales utilizados en la enseñanza de la lectoescritura; o las variables económicas, su distribución en la población y el efecto sobre los aspectos educativos (Bruck, Treiman y Caravolas, 1995; Goswami y Ziegler, 2006). Sin embargo, cabe suponer que el efecto de tales diferencias, en su caso, se mostrará principalmente en el período de adquisición inicial de la habilidad lectora, cuando ésta depende en mayor medida de factores educativos, y tenderá a disminuir su influencia en muestras de lectores expertos y adultos, como los analizados en nuestro estudio.

Desde nuestro punto de vista, las diferencias encontradas en el procesamiento lector entre universitarios españoles y uruguayos están directamente relacionadas con las diferencias en consistencia de la ortografía del español hablado en Uruguay y el hablado en España. En particular, hemos sugerido que el factor específico que podría explicar las diferencias es el hecho de que en el caso uruguayo (y por ex-

presión latinoamericano) la desaparición del fonema /è/ y su sustitución por el fonema /S/ genera un nuevo conjunto de inconsistencias que logra ejercer un efecto en los mecanismos de procesamiento (y posiblemente de aprendizaje) debido a su alta frecuencia. En efecto, mientras que para un español la escritura de la palabra “cine” no reviste dificultad alguna ya que existe una regla del tipo /è/(e,i)à'c' [nota al pie: Las reglas de la ortografía española incluyen la posibilidad de utilizar la regla /è/à'z', que produciría una representación fonológicamente aceptable. Sin embargo, este tipo de error es muy poco frecuente (Valle Arroyo, 1990), posiblemente debido a la extremada baja frecuencia del grafema “z”, que impide que se constituya un competidor importante del grafema “c” cuando se trata de transcribir las sílabas /èE/ o /èl/], para el español latinoamericano, donde la palabra se pronuncia /SINE/, se puede transcribir de una forma fonológicamente válida como “cine” ó “sine”. Este tipo de ambigüedades se incluyen dentro de los elementos que aumentan el grado de opacidad de una ortografía (Perry, Ziegler y Coltheart, 2002; Ziegler, Jacobs y Stone, 1996, 1997).

El incremento en opacidad ha sido relacionado con una divergencia en el curso de adquisición de la lectoescritura y los procesos cognitivos implicados (Frost, Feldman y Katz, 1990; Katz y Feldman, 1983, 1992; Ziegler y Goswami, 2006). Por ejemplo Katz y Frost (1992), en su formulación de la hipótesis de la profundidad ortográfica, sugiere que en los sistemas más transparentes habrá un uso mínimo (o nulo) de los mecanismos léxicos en el proceso de reconocimiento visual de palabras mientras que en los sistemas opacos habrá un uso intensivo de la información ortográfica léxica en el desempeño de la función lectora. Aunque algunos estudios han matizado y moderado los postulados extremos de dicha hipótesis (Perfetti y Bell, 1991; Van Orden, 1987) disponemos de evidencia empírica que muestra la existencia de diferencias cognitivas entre los lectores de lenguas opacas y transparentes (Katz y Frost, 2001; Seymour, Aro y Erskine, 2003).

Para comprobar la validez de esta predicción sobre nuestros datos hemos sometido a un análisis más detallado la diferencias encontradas. En primer lugar hemos categorizado a todos los participantes en el experimento como altos, medios o bajos según la puntuación obtenida en la prueba TECLE. El procedimiento se-

guido ha consistido en calcular la puntuación típica para cada sujeto, tomando para ello la media y desviación típica de la muestra (española o uruguaya) a la que pertenece, y posteriormente aplicar la regla de clasificación que asigna el valor “bajo” a todos los que presentan un valor z inferior a -0.65, “alto” a los que presentan una puntuación típica superior a 0.65 y “medio” a los que tienen un valor z situado entre esos dos valores. Un procedimiento equivalente hemos seguido para la categorización de los sujetos según su conocimiento ortográfico, con la salvedad de que en el caso uruguayo hemos utilizado la puntuación TECOR que no considera las respuestas a los ítems /SE/-/SI/, para que los resultados en ambas muestras resulten comparables con mayor facilidad. A continuación hemos construido sendas tablas de contingencia, donde calculamos el porcentaje de sujetos que cae en cada una de las 9 categorías resultantes de cruzar las dos variables. En la tabla 6 se muestran los resultados obtenidos.

La diferencia principal entre los datos españoles y uruguayos consiste en los valores que aparecen en las casillas NL-Alto/NO-Bajo y NL-Bajo/NO-Alto, que en ambos casos representan valores en torno al 10% (8.1% y 7.2%, para el caso español y uruguayo, respectivamente). La diferencia entre ambas muestras aparece cuando observamos la diferente composición de este grupo. Mientras que en el caso español tenemos que el 75% de estos participantes son del tipo NL-Bajo/NO-Alto, quedando un 25% en la categoría NL-Alto/NO-Bajo; en el caso uruguayo la distribución en ambos grupos es igual al 50%.

Si se nos permite la licencia de denominar a la categoría NL-Alto/NO-Bajo como “superficial”

y a la categoría NL-Bajo/NO-Alto como “fonológica”, dado que representan patrones que rememoran las características de los principales subtipos de dislexia, podemos decir que en la muestra española hay una mayor incidencia del síndrome fonológico mientras que existe un equilibrio entre ambos síndromes en el caso uruguayo. Estos resultados resultan consistentes con los obtenidos por un lado por Calvo (1999; Carrillo, Alegría y Calvo, 2007) en muestras de niños españoles y por otro por Cuadro (2006) en muestras de niños uruguayos, utilizando en ambos casos materiales y procedimientos muy similares.

Por lo tanto en el presente trabajo hemos intentado describir una iniciativa novedosa para abordar el estudio de los mecanismos cognitivos que subyacen al procesamiento lector. Frente al gran proyecto COST (Seymour, 2004), que ha destinado una gran cantidad de esfuerzo y recursos para estudiar los aspectos diferenciales de la adquisición de la habilidad lectora en diferentes lenguas occidentales, queremos incluir nuestro trabajo en el que abordamos el estudio de aspectos más sutiles pero de más fácil interpretación. Los resultados procedentes de las diferentes iniciativas podrán combinarse en una teoría global sobre el efecto de las particularidades ortográficas en el proceso de adquisición de la lectura.

Asimismo, los resultados del presente estudio necesitan ser contrastados con otros trabajos realizados sobre población más joven, que podría mostrar el componente evolutivo de estos resultados al tiempo que podría circunvenir el problema del efecto techo en la evaluación de los conocimientos ortográficos.

TABLA 6.
Porcentajes de alumnos que obtuvieron una categorización baja, media o alta en las puntuaciones de lectura y conocimiento ortográfico. Los valores están calculados respecto al total de la muestra (española o uruguaya).

	Nivel Lector [NL] (TECLE)	Nivel Ortográfico [NO] (TECOR)		
		Bajo	Medio	Alto
España	Bajo	12.2%	10.2%	6.1%
	Medio	7.5%	26.5%	12.9%
	Alto	2.0%	12.9%	9.5%
Uruguay	Bajo	14.3%	12.5%	3.6%
	Medio	7.1%	25.0%	16.1%
	Alto	3.6%	10.7%	7.1%

REFERENCIAS

- Alegría, J., Marín, J., Carrillo, M. S. y Mousty, P. (2003). Les premières pas dans l'acquisition de l'orthographe en fonction du caractère profond ou superficiel du système alphabétique: comparaison entre le français et l'espagnol. In Romdhane, M.N., Gombert, J.E. and Belajouza, M. (Eds.), *L'apprentissage de la lecture. Perspectives comparatives* (pp. 51-67). Rennes, France: Presse Universitaires de Rennes.
- Alegría, J. y Mousty, P. (1996). The development of spelling procedures in French-speaking, normal and reading-disabled children: Effects of frequency and lexicality. *Journal of Experimental Child Psychology* 63(2), 312-338.
- Bruck, M., Treiman, R. y Caravolas, M. (1995). Role of the syllable in the processing of spoken English: Evidence from a nonword comparison task.
- Calvo, A. R. (1999). Adquisición de la lectura en lengua castellana: perfiles cognitivos de aprendices con dificultades. Tesis doctoral no publicada. Departamento de Psicología Básica y Metodología. Universidad de Murcia, España.
- Carrillo, M. S., Alegría, J. y Calvo, A. R. (2007,). Subtipos de dislexia en una ortografía transparente. Paper presented at the XXV Congreso Internacional de AESLA, Murcia (Spain).
- Coltheart, M., Curtis, B., Atkins, P. y Haller, M. (1993). Models of reading aloud: Dual-route and Parallel Distributed Processing Approaches. *Psychological Review* 100 589-608.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. y Ziegler, J. C. (2001). DRC: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review* 108(1), 204-256.
- Cuadro, A. (2006). Caracterización de los malos lectores en relación a los buenos lectores en niños uruguayos. Tesis doctoral no publicada. Universidad de Murcia. España
- Ehri, L. C. y Wilce, L. S. (1985). Movement into reading: is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly* 20 163-179.
- Frost, R., Feldman, L. B. y Katz, L. (1990). Phonological ambiguity and lexical ambiguity: Effects on visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16(4), 569-580.
- Frost, R. y Katz, L. (Eds.). (1992). *Orthography, phonology, morphology, and meaning*. Vol. 94). Oxford, England: North-Holland.
- Goswami, U. y Ziegler, J. C. (2006). A developmental perspective on the neural code for written words. *TRENDS in Cognitive Sciences* 10(4), 142-143.
- Jacobs, A. M. y Grainger, J. (1994). Models of Visual Word Recognition — Sampling the State of the Art. *Journal of Experimental Psychology: Human Perception and Performance* 20(6), 1311-1334.
- Jorm, A. F. y Share, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics* 4 103-147.
- Katz, L. y Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9(1), 157-166.
- Katz, L. y Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In Frost, R. and Katz, L. (Eds.), *Orthography, phonology, morphology, and meaning*. (pp. 67-84). Oxford, England: North Holland.
- Katz, L. y Frost, S. (2001). Phonology Constrains the Internal Orthographic Representation. *Reading and Writing* 14 297-332.
- Katz, L. y Frost, S. (2001). Phonology Constrains the Internal Orthographic Representation. *Reading and Writing* 14 297-332.
- Marín, J. y Carrillo, M. S. (1999). Test Colectivo de Eficacia Lectora (TECLE). Documento no publicado, Universidad de Murcia, Dep. Psicología Básica y Metodología.
- Marín, J., Alegría, J. y Carrillo, M. S. (remitido). First steps in the acquisition of a deep (French) and a shallow (Spanish) orthography.
- Marín, J., Aveledo, F., Pagán, A. y Cuadro, A. (2007). ¿Por qué es más complicado distinguir entre "berenjena" y "berengena" que entre "escabeche" y "escaveche"? Paper presented at the XXV Congreso Internacional de AESLA, Murcia.
- Marín, J., Aveledo, F., Pagán, A. y Pérez, M. A. (2006,). El papel de la Ambigüedad Fonológica en la Adquisición de Representaciones Ortográficas. Paper presented at the VI Congreso de la Sociedad Española de Psicología Experimental, Santiago de Compostela.
- Marín, J., Aveledo, F., Vera, F. y Alegría, J. (2004a,). El papel de la fonología en la adquisición del léxico ortográfico. El caso de los grafemas 'g' y 'c'. Comunicación presentada al V Congreso de la Sociedad Española de Psicología Experimental (SEPEX), Madrid.
- Marín, J., Aveledo, F., Vera, F. y Alegría, J. (2004b,). The acquisition of lexical orthography: The case of ambiguous graphemes. Paper presented at the British Psychological Society. XXI Annual Cognitive Section Conference, Leeds, UK.
- Marín, J., Pagán, A., Cuadro, A. y Aveledo, F. (2007). Adquisición de información ortográfica y mecanismos de recodificación fonológica: nuevas implicaciones para el modelo de autoaprendizaje (Share, 1995). Paper presented at the XXV Congreso Internacional de AESLA, Murcia, 19-21 de abril.
- Perfetti, C. A. y Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: evidence from backward masking and priming.

- Journal of Memory and Language 30(4), 473-485.
- Perry, C., Ziegler, J. C. y Coltheart, M. (2002). How predictable is spelling? Developing and testing metrics of phoneme-grapheme contingency. *Quarterly Journal of Experimental Psychology* 55A 43-73.
- Sebastian Galles, N., Martí Antonin, M. A., Carreiras, M. y Cuetos, F. (2000). *Lexesp. Léxico informatizado del español*. Barcelona, Spain: Edicions de la Universitat de Barcelona.
- Seymour, P. H. K. (2004). Literacy Acquisition in European Orthographies, from <http://www.dundee.ac.uk/psychology/collesrc/welcome.htm>.
- Seymour, P. H. K. (2005). Early Reading Development in European Orthographies. In Snowling, M.J. and Hulme, C. (Eds.), *The Science of reading: A handbook* (pp. 296-315). Malden, MA: USA: Blackwell Publishing.
- Seymour, P. H. K., Aro, M. y Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*(94), 143-174.
- Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition* 55 151-218.
- Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology* 87 267-298.
- Valle Arroyo, F. (1990). Spelling errors in Spanish. *Reading and Writing* 2(1), 83-98.
- Van Orden, G. C. (1987). A ROWS is a ROSE: spelling, sound, and reading. *Memory and Cognition* 15(3), 181-198.
- Vellutino, F. R. (1979). *Dyslexia: Theory and research*. Cambridge, MA: MIT Press.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J. y Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry* 45(1).
- Ziegler, J. C. y Goswami, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Developmental Science* 9(5), 429-453.
- Ziegler, J. C., Jacobs, A. M. y Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments and Computers* 28(4), 504-515.
- Ziegler, J. C., Stone, G. O. y Jacobs, A. M. (1997). What is the pronunciation for -ough and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments and Computers* 29(4), 600-618.