

Original

## El conocimiento de las grafías y la secuencia de aprendizaje de los grafemas en español: precursores de la lectura temprana



Ana María De la Calle<sup>a,\*</sup>, Fernando Guzmán-Simón<sup>b</sup>, y Eduardo García-Jiménez<sup>c</sup>

<sup>a</sup> Departamento de Psicología Evolutiva, Facultad de Ciencias de la Educación, Universidad de Cádiz, Cádiz, España

<sup>b</sup> Departamento de Didáctica de la Lengua y la Literatura y Filologías Integradas, Facultad de Ciencias de la Educación, Universidad de Sevilla, Sevilla, España

<sup>c</sup> Departamento de Métodos de Investigación y Diagnóstico en Educación, Facultad de Ciencias de la Educación, Universidad de Sevilla, Sevilla, España

### INFORMACIÓN DEL ARTÍCULO

*Historia del artículo:*

Recibido el 3 de abril de 2017

Aceptado el 11 de octubre de 2017

On-line el 21 de noviembre de 2017

*Palabras clave:*

Grafías

Precursores

Lectura temprana

Rendimiento lector

Modelo de ecuaciones estructurales

### R E S U M E N

Este artículo aborda una investigación sobre el valor predictivo del conocimiento de las grafías para el rendimiento lector. Los estudios precedentes han mostrado una fuerte relación entre el conocimiento de las grafías en edades tempranas y los procesos de decodificación. Nuestro estudio indaga sobre el proceso de aprendizaje del código alfabético español centrado en los grafemas y analiza el poder predictivo del conocimiento de las grafías para las habilidades de decodificación en la lectura inicial. Para ello, se apoya en una metodología longitudinal prospectiva y en el uso de instrumentos estandarizados (PROLEC-R y BIL) aplicado a 362 alumnos y alumnas de 4 y 5 años. Los datos obtenidos se analizan mediante regresión múltiple, utilizando modelos de ecuaciones estructurales. Nuestra investigación subraya la relevancia de la secuencia de aprendizaje de los grafemas en edades tempranas, diferenciándose la secuencia de aprendizaje en español con respecto a la lengua inglesa. En español, la secuencia de aprendizaje de los grafemas es independiente de la secuencia de aprendizaje de los fonemas. Asimismo, este artículo subraya la importancia de una secuencia de aprendizaje de estas grafías para predecir el desarrollo de las habilidades de decodificación. Este estudio concluye que las prácticas educativas tempranas que consideran el nombre y fonema de las grafías y la secuencia determinada en el aprendizaje de los grafemas optimizan el rendimiento lector de los niños y niñas de lengua española.

© 2017 Universidad de País Vasco. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Letter Knowledge and Learning Sequence of Graphemes in Spanish: Precursors of Early Reading

### A B S T R A C T

The current article addresses a research on the predictive value of knowing the letters within the initial reading performance. Recent studies have shown a strong link between the knowledge of letters at early ages and the decoding processes. Our study deepens into the learning process of the Spanish alphabet code, focused on the graphemes, and analyzes the predictive power of knowing of letters for the decoding abilities in initial reading. To this end, the research relies on a longitudinal prospective methodology and makes use of standardized instruments (PROLEC-R and BIL) applied to 362 students aged 4 and 5. The data obtained are analyzed through multiple regression, using structural equation models. Our research outlines the relevance of learning the sequence of graphemes from early ages, differentiating the sequence of learning in Spanish with respect to English language. In Spanish, the learning sequence of the graphemes is independent of the learning sequence of phonemes. Moreover, this article emphasizes the importance of a learning sequence of these letters, in order to foresee the development of the decoding abilities. This study concludes that the early educational practices that take into account the letters name and phoneme, as well as the concrete sequence in graphemes learning, optimize the reading performance of Spanish speaking children.

© 2017 Universidad de País Vasco. Published by Elsevier España, S.L.U. All rights reserved.

*Keywords:*

Letters

Precursors

Early reading

Reading performance

Structural equation models

\* Autor para correspondencia.

Correo electrónico: [anamaria.decalleca@gmail.com](mailto:anamaria.decalleca@gmail.com) (A.M. De la Calle).

## Introducción

El conocimiento de las grafías posee un papel relevante en el aprendizaje del código alfabético y supone el descubrimiento de las reglas que relacionan los fonemas y los grafemas. Las grafías representan las unidades mínimas de sonido de una lengua (fonemas) mediante signos ortográficos denominados grafemas (Perry, Ziegler, y Zorzi, 2013). Los grafemas son símbolos con unas características perceptivas determinadas por la orientación y la forma que permiten al lector realizar una discriminación visual (Lorenzo, 2001). Por consiguiente, las destrezas atencionales desempeñan un papel predominante en el conocimiento de las grafías (Helal y Weil-Barais, 2015).

El dominio de las correspondencias grafema/fonema es una de las claves del aprendizaje lector. La instrucción explícita de las correspondencias grafema/fonema garantiza la adquisición del principio alfabético (Earle y Sayeski, 2017; Schaars, Segers, y Verhoeven, 2017). El uso sistemático de dichas correspondencias tras la adquisición del principio alfabético potencia las habilidades de decodificación y, en consecuencia, el conocimiento de las grafías resulta una de las habilidades cognitivas que mejor predice el futuro rendimiento lector (Hammill, 2004; Huang, Tortorelli, y Ivernizzi, 2014; Storch y Whitehurst, 2002).

El carácter precursor del conocimiento del nombre y sonido de las grafías en las primeras etapas instruccionales se ha convertido en uno de los indicadores del rendimiento lector en numerosas investigaciones. El estudio de Muter, Hulme, Snowling, y Stevenson (2004) ha evaluado en 90 sujetos ingleses el conocimiento de las grafías a los 4 y 5 años y su relación con las habilidades de decodificación durante dos años. Esta investigación determina el 63% de la varianza de las habilidades de decodificación un año más tarde a partir del conocimiento del nombre y el sonido de las grafías. En esta misma línea, Evans, Bell, Shaw, Moretti, y Page (2006) evalúan en 149 sujetos canadienses de habla inglesa el conocimiento del nombre y sonido de las grafías a los 5 años y, diez meses después, las habilidades de decodificación. En esta ocasión, el conocimiento del nombre de las grafías en minúscula determina el 51% de la varianza del rendimiento lector de 1.º. Recientemente, Onochie-Quintanilla, Defior, y Simpson (2017) han evaluado en 100 sujetos españoles de Educación Infantil el conocimiento del nombre y sonido de las 27 grafías mayúsculas y minúsculas (y los dígrafos «ch» y «ll») a la edad promedio de 5.6 años y la exactitud en la lectura de palabras en 1.º de Educación Primaria. Sus resultados muestran la ausencia de valor predictivo del conocimiento de las grafías para la lectura de palabras meses más tarde.

Los estudios precedentes subrayan la importancia de las investigaciones sobre la correspondencia grafema/fonema y su carácter predictivo en el desarrollo lector posterior. Los hallazgos de Caravolas, Lervåg, Defior, Seidlová-Málková, y Hulme (2013), de Fricke, Szczerbinski, Fox-Boyer, y Stackhouse (2016) y de Snel, Aarnoutse, Terwel, van Leeuwe, y van der Veld (2016) sitúan al conocimiento de las grafías como un precursor temprano de la lectura común en distintos sistemas ortográficos (inglés, español, checo, alemán y neerlandés). No obstante, nuestra investigación debe considerar que las correspondencias de los grafemas y fonemas varían en función del código alfabético, puesto que el proceso de aprendizaje del código alfabético está condicionado por el grado de transparencia de la lengua (Ziegler et al., 2010). En particular, nuestra investigación se centra en el valor predictivo del conocimiento de las grafías simples para el desarrollo lector inicial en el código alfabético español.

### Secuencia de aprendizaje de los grafemas

Algunos estudios que abordan el aprendizaje del código alfabético inglés han elaborado una secuencia de aprendizaje de las grafías.

La relación existente entre cada grafía simple y su grafema correspondiente permite que se evalúe el conocimiento de los grafemas asociados a ellas al estudiar el conocimiento de las grafías simples en las etapas iniciales de la lectura. Así, las investigaciones aludidas demuestran que el orden de las grafías en el alfabeto inglés y el desarrollo fonológico condicionan la secuencia de aprendizaje de los grafemas (Jones, Clark, y Reutzel, 2013; Justice, Pence, Bowles, y Wiggins, 2006; Treiman, Kessler, y Pollo, 2006).

En particular, Justice et al. (2006) analizan el conocimiento de las 26 grafías del alfabeto inglés de los 3 a los 5 años de edad. Los hallazgos de su investigación muestran la secuencia de aprendizaje de los grafemas del código alfabético inglés mediante el análisis de frecuencias en el conocimiento del nombre de las grafías. Dicha secuencia es determinada a través de los grafemas reconocidos por el 50% de los sujetos. Justice et al. (2006) abordan la relación entre la secuencia de aprendizaje de los grafemas determinada en sus resultados y la secuencia de aprendizaje de los fonemas determinada por Sander (1972). Las conclusiones de Justice et al. (2006) subrayan la existencia de relación entre las dimensiones fonológica y grafémica. Los escolares de habla inglesa aprenden antes aquellos grafemas cuyos fonemas correspondientes se adquieren antes en desarrollo fonológico. Los autores denominan a este hallazgo «la hipótesis del orden de las consonantes».

### Secuencia de aprendizaje de los fonemas

Los estudios sobre el aprendizaje de los fonemas adquieren relevancia a finales del siglo xx. Estas investigaciones han analizado los rasgos fonéticos de los fonemas (como la sonoridad, punto y modo de articulación) y la edad en relación con la adquisición de los fonemas en inglés principalmente (Sander, 1972; Smit, Hand, Freilinger, Bernthal, y Bird, 1990).

Los estudios de Bosch (2004), Camargo (2006) y Vivar y León (2009) investigan el desarrollo fonológico español y muestran distintas decisiones metodológicas, puesto que consideran los rangos de edad en los sujetos y los diferentes criterios para determinar las secuencias de aprendizaje de los fonemas. Con relación a la muestra, los estudios de Bosch (1983, 2004) y Vivar y León (2009) analizan el repertorio fonológico a partir de los 3 años, y Camargo (2006), el desarrollo fonológico desde los primeros meses de vida. Asimismo, los estudios de Bosch (1983, 2004) y Vivar y León (2009) valoran el aprendizaje de los fonemas hasta los 6 años; sin embargo, Camargo (2006) establece la última fase en el aprendizaje de los fonemas a partir de los 3 años. En cuanto a sus resultados, la investigación de Bosch (1983, 2004) describe la adquisición del inventario fonémico de los 3 a los 7 años clasificando los fonemas según porcentajes de acierto en grupos con promedio de edades (3 años, 4 años, 5 años, 6 años y 7 años); Camargo (2006) clasifica el aprendizaje de los fonemas en cuatro grupos comprendidos entre los 0 y los 3 años; por último, Vivar y León (2009) determinan cinco grupos en el aprendizaje de los fonemas entre los 3 y los 5 años y 11 meses.

Los grupos delimitados por los distintos modelos definen implícitamente la secuencia de aprendizaje de los fonemas y el análisis de los datos entre estos autores evidencia ciertas discrepancias en la obtención de los datos. En particular, Bosch (1983, 2004) tiene en consideración la posición de los fonemas dentro de las sílabas en posición directa, inversa o mixta; en cambio, Camargo (2006) estudia la posición de los fonemas dentro de las palabras (inicial, media o final) y Vivar y León (2009) excluyen estos criterios en sus análisis.

La Tabla 1 muestra un cuadro comparativo donde se representa la secuencia de aprendizaje de fonemas propuesta por los autores comentados anteriormente. La adscripción de los fonemas a los distintos grupos se realiza cuando estos son aprendidos en cualquier posición de la sílaba o la palabra. De este modo, las agrupaciones de los modelos referenciados son comparables, pues se comprueba

**Tabla 1**  
Comparativa del aprendizaje de los fonemas en español según los modelos de Bosch (1983, 2004), Camargo (2006) y Vivar y León (2009)

	Bosch (1983, 2004) (3:0-7:11)	Camargo (2006) (0:0-3:0)	Vivar y León (2009) (3:0-5:11)
Grupo 1	/m/, /n/, /ɲ/, /p/, /b/, /k/, /g/, /t/, /f/, /x/	p/, /b/, /m/, /t/, /k/, /g/, /s/	/p/, /m/, /t/
Grupo 2	/l/, /d/, /y/	/n/, /ɲ/, /d/, /x/, /f/, /θ/, /l/, /ks/	/b/, /n/, /ɲ /
Grupo 3	/s/	/r/, /r̄ /	/x/, /k/
Grupo 4	/r/, /r̄ /	-	/l/
Grupo 5	-	-	/d/, /r/
Grupo 6	-	-	/r̄ /, /g/, /s/, /f/, /ks/

que existe una tendencia a aprender unos fonemas antes que otros, aunque dichas secuencias de aprendizaje de los fonemas de los distintos estudios no sean idénticas.

### Nuestra investigación sobre el código alfabético español

Nuestro estudio analiza las relaciones entre los nombres y fonemas de las grafías simples y los grafemas asociados a estas en el proceso de aprendizaje del código alfabético español, y estudia el conocimiento de estas grafías como precursor del rendimiento lector. Tras la revisión del estado de la cuestión, hemos tratado de responder a las siguientes preguntas:

1. ¿Es posible definir una secuencia del proceso de aprendizaje del código alfabético español centrado en los grafemas?
2. ¿Existe relación entre la secuencia de aprendizaje de los grafemas y la secuencia de aprendizaje de los fonemas en español?
3. ¿Qué valor predictivo tiene el conocimiento de las grafías para el rendimiento lector inicial?

Para responder a estas cuestiones, nuestro estudio tiene por objeto examinar el conocimiento que los escolares de 4 y 5 años tienen de las grafías simples (que se asocian con un solo grafema) y su valor predictivo para las habilidades de decodificación al inicio del aprendizaje lector. En primer lugar, se plantea la necesidad de precisar la secuencia de aprendizaje de los grafemas en español con base en el conocimiento de estas grafías. Con posterioridad, las secuencias de aprendizaje de los fonemas propuestas por los modelos de Bosch (1983, 2004), Camargo (2006) y Vivar y León (2009) son comparadas con la secuencia de aprendizaje de los grafemas determinada en nuestro estudio. Por último, se analiza el valor predictivo del conocimiento de las grafías, como variable criterio, para las habilidades de decodificación (variable dependiente).

Nuestro estudio sitúa el foco de la investigación tanto en el proceso de aprendizaje de los grafemas y en su relación con el rendimiento lector posterior como en clarificar el modo en que se produce el aprendizaje del código alfabético en español. El carácter distintivo de esta investigación respecto a los trabajos de Evans et al. (2006), Muter et al. (2004) y Onochie-Quintanilla et al. (2017) radica en la delimitación de la secuencia de aprendizaje de los grafemas en español y su comparativa con el desarrollo fonológico.

### Método

#### Participantes

En este estudio participan 362 alumnos y alumnas matriculados en tres escuelas de la provincia de Cádiz. Estos centros son seleccionados porque poseen el mismo ISE (indicador socioeconómico) y

unas características similares: centros públicos de Educación Infantil y Primaria sujetos a una misma regulación de sus actividades, y un profesorado funcionario experimentado. La presencia de necesidades educativas es considerada como criterio de exclusión para la participación en el estudio. Asimismo, la selección del alumnado está sujeta al consentimiento informado de sus progenitores o progenitoras. Este estudio sigue la regulación interna en Ciencias Sociales por los Comités Éticos de Experimentación de las universidades de Cádiz y de Sevilla.

Los participantes son agrupados en cohortes, cada una de ellas evaluadas individualmente en dos momentos diferentes ( $t_1$  y  $t_2$ ). La Tabla 2 ilustra que la primera cohorte está integrada por 180 alumnos y alumnas de Educación Infantil de 4 años (media de edad = 4.7; 45.6% niñas) y la segunda cohorte, por 182 alumnos y alumnas de Educación Infantil 5 años (media de edad = 5.7; 48.6% niñas). En relación con el desarrollo temporal de las evaluaciones, el  $t_1$  corresponde al segundo trimestre del curso escolar cuando se mide el conocimiento de las grafías, y el  $t_2$  tiene lugar al inicio del primer trimestre del año escolar siguiente (seis meses después del  $t_1$ ) cuando se miden las habilidades de decodificación.

#### Instrumentos

En el primer momento ( $t_1$ ), se hace uso de las subpruebas sobre conocimiento de las grafías de dos materiales. En concreto, se emplean las siguientes subpruebas:

*Subprueba del conocimiento de las letras de la Batería de inicio a la lectura* (BIL 3-6) (Sellés, Martínez, y Vidal-Abarca, 2010). Los autores hallan un coeficiente del alfa de Cronbach de .97 para esta subprueba. Mediante esta subprueba se evalúa el conocimiento de las cinco vocales minúsculas. En el estudio actual el alfa de Cronbach ( $\alpha$ ) calculado es .77, omega de McDonald ( $\Omega$ ) es .83 y la varianza media extractada (VME) es .50.

*Subprueba de reconocimiento de letras de la Batería de evaluación de los procesos lectores* (PROLEC-R) (Cuetos, Rodríguez, Ruano, y Arribas, 2007). Los autores hallan en esta subprueba un alfa de .49 que es superado en nuestras medidas ( $\alpha = .76$ ). Mediante esta subprueba se mide el conocimiento de 20 grafías minúsculas. No obstante, nuestro estudio considera los aciertos para las consonantes (puntuación máxima = 19). El valor de  $\Omega$  en nuestro estudio es .92 y la VME es .41.

En el segundo momento ( $t_2$ ), se aplica la subprueba de lectura de palabras ( $s_1$ ) y la subprueba de lectura de pseudopalabras ( $s_2$ ) de PROLEC-R. Los autores hallan un alfa de .74 para  $s_1$  y de .68, para  $s_2$ . Las subpruebas consisten en la lectura de 40 palabras y 40 pseudopalabras en minúsculas. El análisis factorial de las puntuaciones totales de estas subpruebas, mediante el método de extracción de componentes principales, determina el índice de exactitud lectora

**Tabla 2**  
Cohortes de la muestra

	Educación Infantil de 4 años (4:0-5:4)	Educación Infantil de 5 años (5:0-6:4)	1.º de Educación Primaria (6:0-6:10)
Cohorte 1	180 ( $t_1$ )	180 ( $t_2$ )	
Cohorte 2		182 ( $t_1$ )	182 ( $t_2$ )

**Tabla 3**  
Clasificación del alumnado en función de la secuencia de aprendizaje de los grafemas

Clúster		4 años (4:0-5:4)		5 años (5:0-6:4)	
		Frecuencia	Porcentaje	Frecuencia	Porcentaje
Clúster	1	127	70.9	18	10.1
	2	50	27.9	161	89.9
	Valor atípico (-1)	2	1.1	0	.0
	Combinado	179	100.0	179	100.0

(un componente;  $s_1 = .994$  y  $s_2 = .994$ ). El alfa en nuestras medidas es .93 para la exactitud lectora. El valor de  $\Omega$  es .99 y la VME es .98.

### Procedimiento

Las evaluaciones se han llevado a cabo por uno de los investigadores en un espacio privado que reúne las condiciones ambientales oportunas. Estas fueron realizadas durante varios meses, ya que se realizaron individualmente a todos los participantes en dos sesiones diferentes (la sesión  $t_1$  y la sesión  $t_2$ ). La sesión  $t_1$  consiste en la medición del *conocimiento* de 24 de las 30 *grafías* del alfabeto español, correspondientes a aquellas *grafías simples* que están asociadas a un único grafema (excepto «h», «k» y «w»). Se opta por evaluar el *reconocimiento de grafías minúsculas*, ya que estas son más frecuentes en los textos (Jones y Mewhort, 2004). El uso de BIL y de PROLEC-R permite el aumento de los ítems evaluados en alusión a las *grafías simples* vocálicas y consonánticas del código español. La respuesta es considerada correcta en el *reconocimiento de las grafías* si el niño identifica el *nombre* o el *fonema de la grafía*, o bien el *fonema* en la estructura fonológica silábica. La consideración del último criterio reside en que la identificación del *fonema* en la estructura silábica es más fácil de discriminar y producir a estas edades por los sujetos españoles (Defior y Serrano, 2014). La sesión  $t_2$  se encauza hacia la evaluación de las *habilidades de decodificación de palabras y pseudopalabras en minúsculas* de PROLEC-R. La selección del mismo instrumento aminora los efectos de la caligrafía en el reconocimiento de las características perceptivas de los *grafemas*.

### Análisis de datos

Esta investigación desarrolla diferentes tipos de análisis para dar respuesta a las cuestiones planteadas mediante SPSS y AMOS en sus versiones 23. El análisis inferencial de los datos requiere el uso de pruebas no paramétricas, puesto que no se cumplen todos los supuestos de normalidad. En primer lugar, se lleva a cabo un análisis de clúster bietápico mediante la dicotomización de los datos (1 = error, 2 = acierto). A partir de dicho análisis, se establecen las posibles diferencias significativas entre grupos de edad mediante la clasificación de los sujetos a partir de los resultados obtenidos en el reconocimiento de cada grafía. En segundo lugar, se delimitan las *grafías* que diferencian a los sujetos según un criterio de edad mediante un análisis discriminante apoyado en el método de inclusión por pasos.

Con el propósito de determinar la precisión en el *conocimiento de las grafías* y la secuencia en la que son aprendidos los *grafemas* asociados a estas, se realiza un análisis de frecuencias tanto para el grupo de escolares de 4 años como para el grupo de 5 años, así como para ambos grupos de forma conjunta. Posteriormente, las *grafías* se agrupan según intervalos de frecuencias (quintiles) que permiten establecer la secuencia en la precisión con la que los sujetos conocen los *grafemas*. Asimismo, se considera que los *grafemas* son aprendidos por los grupos de edad, cuando son reconocidos por el 50% de los escolares (Justice et al., 2006). Para confirmar la *secuencia de aprendizaje de los grafemas*, obtenida mediante el análisis de frecuencias, se realiza un clúster jerárquico mediante

el método de Ward y utilizando la distancia euclídea al cuadrado, y, posteriormente, una comparativa entre la secuencia encontrada del aprendizaje de los *grafemas* y la *secuencia de aprendizaje de los fonemas* con base en los modelos de Bosch (1983, 2004), Camargo (2006) y Vivar y León (2009). Finalmente, se analiza la relación entre el *conocimiento de las grafías* a edades tempranas y las *habilidades de decodificación* en el aprendizaje inicial de la lectura, utilizando para ello modelos de ecuaciones estructurales.

## Resultados

### Aprendizaje del código alfabético centrado en los grafemas

El análisis de clúster bietápico identifica dos secuencias diferentes en el aprendizaje de las *grafías* al considerar el grupo de edad. El *conocimiento de las grafías* evaluadas (las 24 *grafías simples*) clasifica al 80.4% de la muestra correctamente bajo el criterio de error o acierto según su grupo de edad (véase Tabla 3). El clúster 1 recoge la secuencia seguida por el alumnado de 4 años (70.9%) y el clúster 2, por el alumnado de 5 años (90%).

El análisis discriminante ha permitido clasificar las *grafías* en función del grupo de edad de los escolares. Las funciones discriminantes canónicas muestran un autovalor = 1.323, un coeficiente de correlación canónica de .755, y un lambda de Wilks de .430 ( $p = .001$ ). La función discriminante es de 1.147 para el grupo de 4 años, y opuesta para el grupo de 5 años (-1.147). En la Figura 1, se han representado los valores de la matriz de estructura, que indican el peso de cada uno de los *grafemas* cuando el grupo de edad actúa como función discriminante en el *conocimiento de las grafías*. En función del grupo de edad, los *grafemas* que mejor discriminan a los sujetos son los asociados a las diez *grafías* siguientes: «t», «n», «d», «m», «ñ», «r», «a», «y», «u» y «j».

Posteriormente, se analiza en qué medida el *conocimiento de las grafías* es el reflejo de la *secuencia de aprendizaje de los grafemas* de los escolares de un grupo de edad. En la Tabla 4, la distribución de frecuencias en el *conocimiento de las grafías* muestra cinco grupos en el aprendizaje de los *grafemas*. La adscripción de los *grafemas* a cada grupo se justifica en que las diferencias entre las frecuencias de los *grafemas* con mayor valor de los distintos quintiles superan el rango contemplado en el intervalo del grupo (20%). Los escolares de los distintos grupos de edad poseen como mínimo un aprendizaje consolidado de los *grafemas* que reconocen en los dos primeros grupos ( $\geq 50\%$ ).

Los resultados de la distribución de frecuencias son corroborados a partir de los resultados obtenidos en el análisis de clúster jerárquico. El dendograma resultante muestra el *conocimiento de las grafías* a los 4 y 5 años, y la clasificación de los *grafemas* en cuatro grupos principales según distancias (Figura 2).

La Tabla 5 muestra que el análisis de distancias entre pares de *grafías* identifica la *secuencia de aprendizaje de los grafemas* en cuatro grupos. Puede observarse que la reagrupación de los *grafemas* mediante el análisis de las distancias mantiene la secuencia determinada según el análisis de frecuencias, si bien con mayor precisión al agrupar los *grafemas* en cuatro y no en cinco grupos. Así, el análisis según distancias establece definitivamente el agrupamiento de los *grafemas* en cuatro grupos.



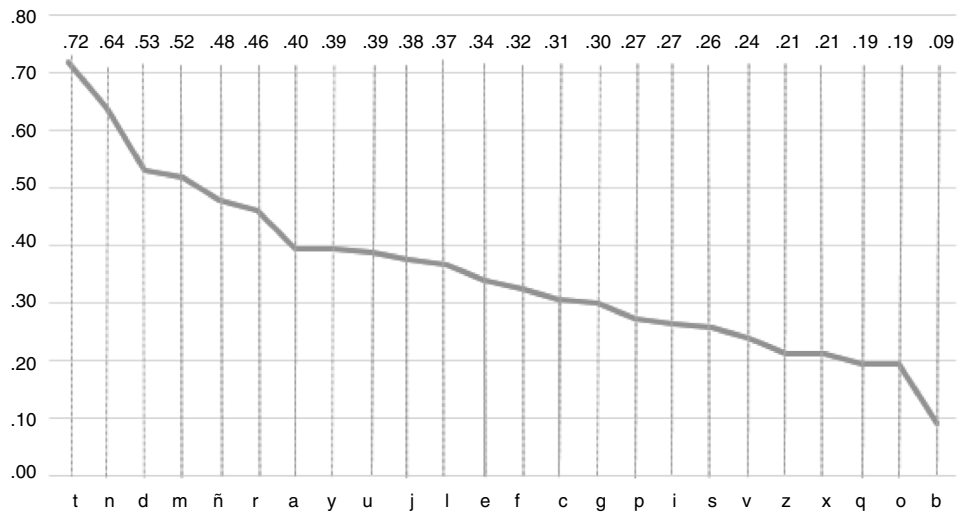


Figura 1. Representación de los valores de la matriz de estructura para el conocimiento de las grafías, en función del grupo de edad.

Tabla 4. Secuencias de aprendizaje de los grafemas según análisis de frecuencias en el conocimiento de las grafías

	Educación Infantil de 4 años (4:0-5:4)	Educación Infantil de 4-5 años(4:0-6:4)	Educación Infantil de 5 años (5:0-6:4)
Grupo 1 [80-100%]	«O»	«O», «í», «S», «E», «U»	«O», «e», «u», «S», «í», «a», «m», «Z», «p», «t»
Grupo 2 [60-80%]	«i», «S», «Z», «e», «u», «a», «C», «p»	«Z», «a», «p», «C», «X», «m»	«C», «X», «n», «d», «ñ», «r», «V», «y»
Grupo 3 [40-60%]	«X», «v», «y», «m»	«v», «ñ», «y», «d», «t», «n», «r», «j»	«j», «f»
Grupo 4 [20-40%]	«ñ», «d», «j», «f», «r»	«f», «g», «l»	«l», «g», «b»
Grupo 5 [0-20%]	«n», «t», «b», «g», «l», «q»	«b», «q»	«q»

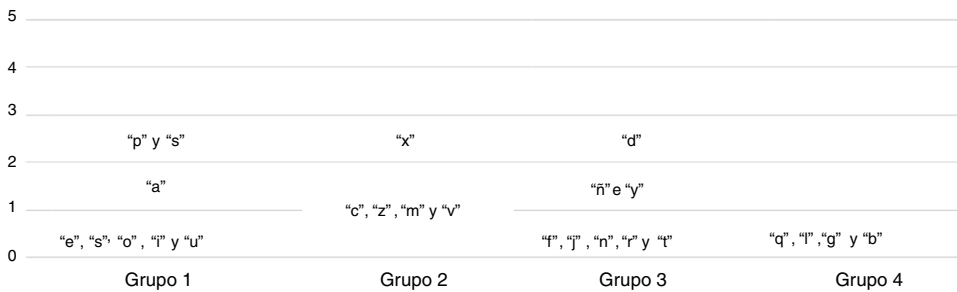


Figura 2. Dendrograma del conocimiento de las grafías a los cuatro y cinco años.

Tabla 5. Comparativa entre las secuencias de aprendizaje de los grafemas según análisis de frecuencias y distancias en el conocimiento de las grafías a los 4-5 años

	Según frecuencias	Según distancias
Grupo 1	«O», «í», «S», «E», «U»	«e», «S», «O», «í», «u», «a», «p»
Grupo 2	«Z», «a», «p», «C», «X», «m»	«C», «Z», «m», «v», «X»
Grupo 3	«v», «ñ», «y», «d», «t», «n», «r», «j»	«f», «j», «n», «r», «t», «ñ», «d», «y»
Grupo 4	«f», «g», «l»	«q», «l», «g», «b»
Grupo 5	«b», «q»	-

Conocimiento de las grafías y habilidades de decodificación

Los escolares de 1.º de Educación Primaria (alumnado de Educación Infantil de 5 años del t<sub>1</sub>) poseen mayores habilidades de decodificación que el alumnado de 5 años (alumnado de Educación Infantil de 4 años del t<sub>1</sub>). La media de palabras y pseudopalabras leídas correctamente por el alumnado de 5 años es 3 de 80 palabras, mientras que el alumnado de 1.º de Educación Primaria lee 50 de 80 palabras. Las habilidades de decodificación aumentan considerablemente con la edad (véase Tabla 6).

Las habilidades de decodificación de palabras y pseudopalabras en Educación Infantil de 5 años y 1.º de Educación Primaria son determinadas por el conocimiento previo de las grafías (seis meses antes). La configuración de modelos de ecuaciones estructurales determina la relación entre el conocimiento de las grafías a los 4 y 5 años y las habilidades de decodificación al inicio del aprendizaje lector. Un primer modelo considera el valor predictivo de todas las grafías para la exactitud lectora a los 5 y 6 años, y logra valores de bondad de ajuste inaceptables:  $\chi^2 = 826.291$   $gl = 275$ ,  $p = .000$ ,  $NFI = .797$ ,  $TLI = .826$ ,  $CFI = .853$ ,  $RMSEA = .075$  (ver Figura 3).

La Figura 4 ilustra un segundo modelo que considera las grafías con mayor valor predictivo para la decodificación. Estas grafías tienen asociados los grafemas que se aprenden en los grupos intermedios en la secuencia de aprendizaje delimitada según el análisis de frecuencias a los 4 y 5 años. Las grafías con mayor valor predictivo (a excepción de «f» y «m») son reconocidas por el 40-60% de los casos (véase la Tabla 4). En este caso se logra un buen ajuste del modelo:  $\chi^2 = 36.223$ ,  $gl = 26$ ,  $p = .088$ ,  $NFI = .979$ ,  $TLI = .992$ ,  $CFI = .994$ ,  $RMSEA = .033$ .

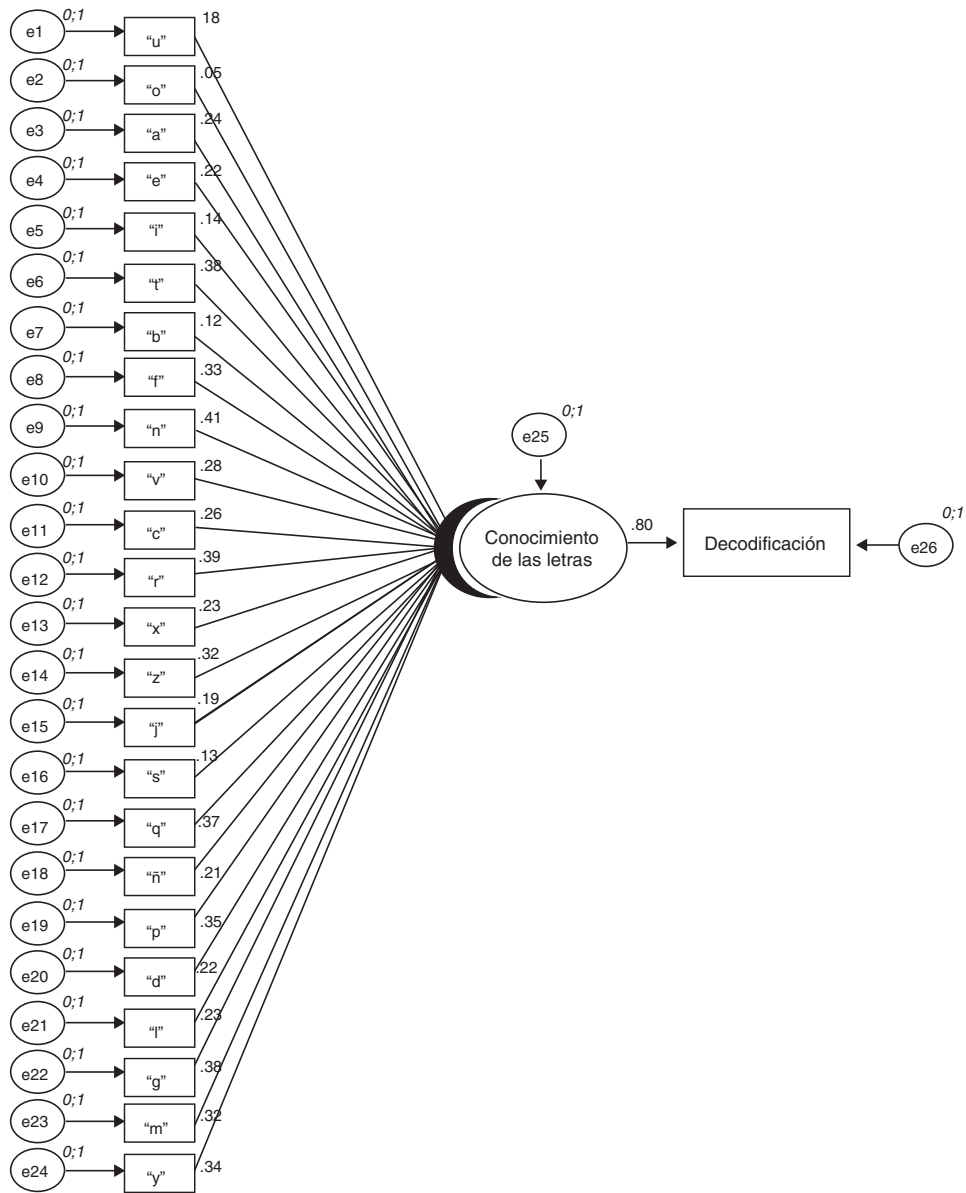


Figura 3. Modelo de predicción del conocimiento de letras para las habilidades de decodificación I.

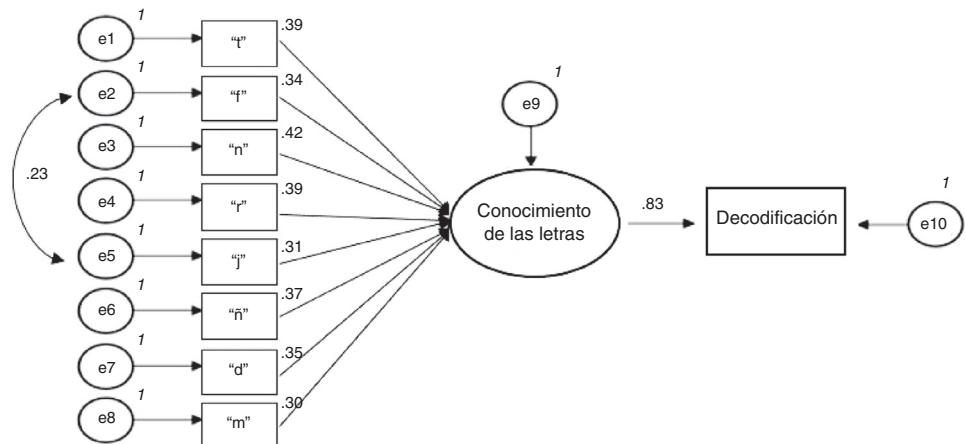


Figura 4. Modelo de predicción del conocimiento de letras para las habilidades de decodificación II.

**Tabla 6**  
Resultados medios en las habilidades de decodificación

	Media (DE)			Rango
	Educación Infantil de 5 años (5:0-6:4)	Educación Infantil de 5 años y 1.º de Educación Primaria (5:0-6:10)	1.º de Educación Primaria (6:0-6:10)	
Lectura de palabras	1.69 (6.62)	13.97 (16.13)	26.33 (13.15)	0-40
Lectura de pseudopalabras	1.30 (5.45)	12.53 (14.70)	28.81 (12.23)	0-40
Total	2.99 (11.85)	26.50 (30.64)	50.14 (25.04)	0-80

## Discusión

### Secuencia de aprendizaje de los grafemas

Nuestra investigación analiza la secuencia de aprendizaje del código alfabético español centrada en los *grafemas*. Las diferencias encontradas en el *conocimiento de las grafías simples*, según la edad de los participantes, señalan la existencia de distintos niveles de desarrollo en el proceso de aprendizaje de los *grafemas*. A los 4 años, los escolares tienen un aprendizaje consolidado, como mínimo, de todas las *vocales* y 5 *consonantes* («s», «z», «c», «p» y «x»); un año después, con 5 años consolidan 8 *consonantes* más («t», «m», «n», «d», «ñ», «r», «v» e «y»). El *conocimiento de las grafías* delimita el aprendizaje de los *grafemas* en cuatro grupos (véase *Tabla 5*). Las *vocales* forman parte de los *grafemas* que antes se aprenden; mientras que los *grafemas* «q», «l», «g» y «b» son los que más tarde se aprenden. Los *grafemas* «g» y «q» se corresponden con aquellas *grafías* que fueron menos reconocidas por su *nombre* y *sonido* a los 5 años en el estudio de *Evans et al. (2006)*.

La *secuencia de aprendizaje de los grafemas* en nuestra investigación difiere de los estudios que consideran relevante el orden de las *grafías* en la cadena alfabética y el desarrollo fonológico. Las divergencias encontradas en la *secuencia del aprendizaje de los grafemas* pueden estar motivadas por los distintos grados de transparencia ortográfica de las lenguas (*Defior y Serrano, 2014; Ziegler et al., 2010*). Los trabajos de *Jones et al. (2013)*, *Justice et al. (2006)* y *Treiman et al. (2006)* consideran relevante el orden de las *grafías* en la cadena alfabética en inglés. Sin embargo, nuestros resultados con sujetos españoles son opuestos. La exposición del alumnado de lengua inglesa a la enseñanza del alfabeto podría ser mayor que la del alumnado español, o bien, la exposición a la enseñanza del alfabeto podría no afectar a la secuencia de aprendizaje de los *grafemas* en la misma medida que en la lengua inglesa. Esta última explicación sería coherente con las conclusiones del estudio de *Treiman et al. (2006)* que señalan las diferencias en el *conocimiento alfabético* entre escolares de lengua inglesa y portuguesa. Dichas conclusiones aluden a que los sujetos estadounidenses cometen menos errores que los sujetos portugueses al reconocer las primeras letras del alfabeto. Por otro lado, nuestro estudio en la lengua española no confirma los hallazgos encontrados en la investigación de *Justice et al. (2006)*, en la que se relacionan el aprendizaje de los *grafemas* y el desarrollo fonológico en inglés. Mientras que el trabajo *Justice et al. (2006)* manifiesta que los escolares de habla inglesa aprenden antes aquellos *grafemas* cuyos *fonemas* correspondientes se adquieren antes en el repertorio fonológico, nuestro estudio delibera que los de habla española no aprenden los *grafemas* en correspondencia al desarrollo fonológico de sus *fonemas*.

### Secuencia de aprendizaje de los grafemas versus secuencia de aprendizaje de los fonemas

El presente estudio aborda la *secuencia de aprendizaje de los grafemas* en español entre los 4 y los 6 años, y establece su carácter independiente respecto a su desarrollo fonológico. En la *Tabla 7* se comparan las secuencias en el aprendizaje de los *grafemas* y los *fonemas*, de acuerdo con nuestros resultados y los modelos de

*Bosch (1983, 2004)*, *Camargo (2006)* y *Vivar y León (2009)*. Puede observarse, por ejemplo, que el *grafema* «s» se aprende en primer lugar; sin embargo, el *fonema* /s/ es adquirido por el niño con posterioridad, según el modelo de *Bosch (1983, 2004)* y *Vivar y León (2009)*. También, el *grafema* «b» es el último en aprenderse, a pesar de que el *fonema* /b/ es asimilado tempranamente (*Bosch, 1983, 2004; Camargo, 2006; Vivar y León, 2009*).

La secuencia hallada en el aprendizaje de los *grafemas* no coincide con las secuencias en el aprendizaje de los *fonemas* de los modelos de *Bosch (1983, 2004)*, *Camargo (2006)* y *Vivar y León (2009)*. La edad de los participantes nos permite comparar nuestra *secuencia de aprendizaje de los grafemas* con el repertorio fonológico de los escolares de 3 a 6 años (*Vivar y León, 2009*), y revelar que la *secuencia de aprendizaje de los grafemas* en español no reside en el desarrollo fonológico.

### Conocimiento de las grafías y habilidades de decodificación

El *conocimiento de las grafías* a edades tempranas condiciona el *rendimiento lector* posterior, según se ha podido comprobar en nuestro modelo predictivo (véase *Figura 3*). Nuestro estudio evidencia que el *conocimiento del nombre y fonema de las grafías* es un buen precursor del *rendimiento lector* al inicio del aprendizaje de la lectura, algo que también ha sido confirmado por los trabajos de *Evans et al. (2006)* y *Muter et al. (2004)*. Los resultados de nuestro análisis referido a la lengua española permiten establecer una relación causal entre el *conocimiento de las grafías* a edades tempranas y el desarrollo de las *habilidades de decodificación* hacia el aprendizaje inicial de la lectura. El *conocimiento de los grafemas* «t», «f», «n», «r», «j», «ñ», «d» y «m» a los 4 y 5 años explican el 83% de la varianza de las *habilidades de decodificación* seis meses después. Los resultados de nuestro estudio muestran que estos son los *grafemas* con mayor valor predictivo y que son aprendidos (a excepción de «f» y «m») por entre el 40 y el 60% de los casos. Este dato es coherente con el criterio (50%) establecido por *Justice et al. (2006)* para la delimitación de la *secuencia de aprendizaje de los grafemas*. La correlación existente entre los errores en el *conocimiento de las grafías* «f» y «j» vendría a indicar que existe algún criterio subyacente que las une. Una explicación posible es que estas *grafías* resulten más confusas en su discriminación ortovisual. *Huang e Invernizzi (2012)* demostraron que las *grafías* «b», «d», «p» y «q» son nombradas incorrectamente con un 8% más de probabilidad que las *grafías* menos confusas visualmente.

El valor predictivo de las *habilidades de decodificación* para el aprendizaje de la lectura no reside únicamente en el *conocimiento de las grafías*, sino en el aprendizaje de determinadas *grafemas* en una determinada secuencia. Los estudios de *Evans et al. (2006)* y *Muter et al. (2004)* consideran el valor predictivo de las 26 *grafías* del código alfabético inglés para las *habilidades de decodificación* posteriores. El 51 y el 63% de la varianza de las *habilidades de decodificación* es explicado por el *conocimiento de todas las grafías*. Asimismo, el trabajo de *Onochie-Quintanilla et al. (2017)* no considera el valor predictivo de las 27 *grafías simples* y dos dígrafos del código alfabético español para la varianza de las *habilidades de decodificación* posteriores. Sin embargo, nuestro estudio considera el valor predictivo de determinadas *grafías* y alcanza mejor índice

**Tabla 7**

Comparativa entre las secuencias evolutivas del aprendizaje de los grafemas y los fonemas de acuerdo con nuestros resultados y los modelos de Bosch (1983, 2004), Camargo (2006) y Vivar y León (2009)

Secuencia de aprendizaje de los grafemas (4:0-6:4)		Secuencia de aprendizaje de los fonemas		
		Bosch (1983, 2004) (3:0-7:11)	Camargo (2006) (0:0-3:0)	Vivar y León (2009) (3:0-5:11)
Grupo 1	«s», «p»	/m/, /n/, /ɲ/, /p/, /b/, /k/, /g/, /t/, /f/, /x/	/p/, /b/, /m/, /t/, /k/, /g/, /s/	/p/, /m/, /t/
Grupo 2	«c», «z», «m», «v», «x»	/l/, /d/, /y/	/n/, /ɲ/, /d/, /x/, /f/, /θ/, /l/, /ks/	/b/, /n/, /p/
Grupo 3	«f», «j», «n», «r», «t», «ñ», «d», «y»	/s/	/r/, /r̄/	/x/, /k/
Grupo 4	«q», «l», «g», «b»	/r/, /r̄/	-	/l/
Grupo 5	-	-	-	/d/, /r/
Grupo 6	-	-	-	/rr/, /g/, /s/ /f/, /ks/

de varianza explicada (83%). El conocimiento de las grafías resulta ser mejor precursor cuando se consideran las grafías de acuerdo con la secuencia de aprendizaje de los grafemas, es decir, cuando se consideran los grafemas que han sido aprendidos en torno a la mitad de los casos.

Nuestro estudio confirma que las habilidades de decodificación son potenciadas cuando los sujetos adquieren el principio alfabético (Huang et al., 2014). Del mismo modo, estos hallazgos reafirman que el desconocimiento de las grafías, que deben estar adquiridas en correspondencia a la etapa evolutiva en que se encuentra el sujeto, supone un factor de riesgo en el aprendizaje de la lectura (Hammill, 2004; Storch y Whitehurst, 2002). De acuerdo con este estudio, los grafemas en español son consolidados por los sujetos siguiendo una secuencia de aprendizaje independiente del desarrollo fonológico. Además, el conocimiento de las grafías a edades tempranas, en correspondencia con la secuencia de aprendizaje de los grafemas, es un buen precursor del rendimiento lector posterior. En consecuencia, las estrategias educativas en la enseñanza del código alfabético español optimizan el rendimiento en la lectura inicial cuando incorporan en el aprendizaje el nombre y fonema de las grafías y la secuencia de aprendizaje de los grafemas. Por consiguiente, el presente estudio puede ser el inicio de una nueva línea de investigación en el aprendizaje del código alfabético español centrado en los grafemas, debido a la importancia de la secuencia de aprendizaje de estos para el éxito lector.

#### Limitaciones y líneas de investigaciones futuras

Una de las limitaciones de nuestro estudio está relacionada con el hecho de que no se han tomado en consideración expresamente las características perceptivas de los grafemas y su relación con la capacidad cognitiva de los sujetos estudiados, lo que debería ser objeto de una futura investigación.

Otra limitación de los resultados de esta investigación está relacionada con la evaluación de las grafías simples «h», «k» y «w». Asimismo, la evaluación del conocimiento de las grafías debería realizarse en intervalos de edad más pequeños a los considerados en el actual estudio para una delimitación más precisa de la secuencia de aprendizaje de los grafemas en español.

Por último, nuestra investigación requiere de un análisis del repertorio grafémico tanto a edades más tempranas como otras posteriores, con el fin de realizar un estudio comparado bajo el mismo criterio de edad. Esto permitiría un cotejo más exhaustivo de nuestra investigación con los modelos desarrollados por Bosch (1983, 2004) y Camargo (2006). Del mismo modo, las investigaciones futuras deberán indagar sobre la configuración y los factores contextuales que inciden en el aprendizaje de los grafemas en español y el valor predictivo del conocimiento de las grafías, como precursor del rendimiento lector a lo largo del tiempo mediante el seguimiento longitudinal en distintos momentos.

#### Referencias

- Bosch, L. (1983). La evaluación del desarrollo fonológico en niños de 4 y 7 años. *Anuario de Psicología*, 28, 87–114.
- Bosch, L. (2004). *Evaluación fonológica del habla infantil*. Barcelona: Masson.
- Camargo, M. (2006). *Desarrollo fonético-fonológico del español en niños*. *Umbra Científico*, 9, 5–9.
- Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., y Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science*, 24(8), 1398–1407. <http://dx.doi.org/10.1177/0956797612473122>
- Cuetos, F., Rodríguez, B., Ruano, E., y Arribas, D. (2007). *PROLEC-R*. Madrid: TEA ediciones.
- Defior, S., y Serrano, F. (2014). Aspectos diacrónicos y sincrónicos del español: relación con la adquisición del lenguaje escrito. *Estudios de Psicología*, 35(3), 450–475. <http://dx.doi.org/10.1080/02109395.2014.974422>
- Earle, G. A., y Sayeski, K. L. (2017). Systematic instruction in phoneme-grapheme correspondence for students with reading disabilities. *Intervention in School & Clinic*, 52(5), 262–269. <http://dx.doi.org/10.1177/1053451216676798>
- Evans, M. A., Bell, M., Shaw, D., Moretti, S., y Page, J. (2006). Letter names, letter sounds and phonological awareness: An examination of kindergarten children across letters and of letters across children. *Reading and Writing*, 19(9), 959–989. <http://dx.doi.org/10.1007/s11145-006-9026-x>
- Fricke, S., Szczerbinski, M., Fox-Boyer, A., y Stackhouse, J. (2016). Preschool predictors of early literacy acquisition in German-speaking children. *Reading Research Quarterly*, 51(1), 29–53. <http://dx.doi.org/10.1002/rq.116>
- Hammill, D. D. (2004). *What we know about correlates of reading*. *Exceptional Children*, 70(4), 453–469.
- Helal, S., y Weil-Barais, A. (2015). Cognitive determinants of early letter knowledge. *European Early Childhood Education Research Journal*, 23(1), 86–98. <http://dx.doi.org/10.1080/1350293X.2014.991097>
- Huang, F. L., y Invernizzi, M. A. (2012). Factors associated with lowercase alphabet naming in kindergartners. *Applied Psycholinguistics*, 35(06), 943–968. <http://dx.doi.org/10.1017/S0142716412000604>
- Huang, F. L., Tortorelli, L. S., y Invernizzi, M. A. (2014). An investigation of factors associated with letter-sound knowledge at kindergarten entry. *Early Childhood Research Quarterly*, 29(2), 182–192. <http://dx.doi.org/10.1016/j.ecresq.2014.02.001>
- Jones, C. D., Clark, S. K., y Reutzel, D. R. (2013). Enhancing alphabet knowledge instruction: Research implications and practical strategies for early childhood educators. *Early Childhood Education Journal*, 41(2), 81–89. <http://dx.doi.org/10.1007/s10643-012-0534-9>
- Jones, M. N., y Mewhort, D. J. (2004). Case-sensitive letter and bigram frequency counts from large-scale English corpora. *Behavior Research Methods, Instruments, & Computers*, 36(3), 388–396.
- Justice, L. M., Pence, K., Bowles, R. B., y Wiggins, A. (2006). An investigation of four hypotheses concerning the order by which 4-year-old children learn the alphabet letters. *Early Childhood Research Quarterly*, 21(3), 374–389. <http://dx.doi.org/10.1016/j.ecresq.2006.07.010>
- Lorenzo, J. R. (2001). Procesos cognitivos básicos relacionados con la lectura. *Primera parte: la conciencia fonológica*. *Interdisciplinaria*, 18(1), 1–33.
- Muter, V., Hulme, C., Snowling, M. J., y 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. <http://dx.doi.org/10.1037/0012-1649.40.5.665>
- Onochie-Quintanilla, E., Defior, S., y Simpson, I. C. (2017). Visual multi-element processing as a pre-reading predictor of decoding skill. *Journal of Memory and Language*, 94(1), 134–148. <http://dx.doi.org/10.1016/j.jml.2016.11.003>
- Perry, C., Ziegler, J. C., y Zorzi, M. (2013). A computational and empirical investigation of graphemes in reading. *Cognitive Science*, 37(5), 800–828. <http://dx.doi.org/10.1111/cogs.12030>
- Sander, E. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders*, 37(1), 55–63.
- Schaars, M. M. H., Segers, E., y Verhoeven, L. (2017). Word decoding development in incremental phonics instruction in a transparent orthography. *Reading and Writing*, 30(7), 1529–1550. <http://dx.doi.org/10.1007/s11145-017-9735-3>
- Sellés, P., Martínez, T., y Vidal-Abarca, E. (2010). *Batería de inicio a la lectura: diseño y características psicométricas*. *Bordón*, 2(2), 137–160.



- Smit, A., Hand, L., Freilinger, J., Bernthal, J., y Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*(4), 779–798.
- Snel, M. J., Aarnoutse, C. A. J., Terwel, J., van Leeuwe, J. F. J., y van der Veld, W. M. (2016). Prediction of word recognition in the first half of grade 1. *European Early Childhood Education Research Journal, 24*(2), 229–238. <http://dx.doi.org/10.1080/1350293X.2016.1143260>
- Storch, S. A., y Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934–947. <http://dx.doi.org/10.1037//0012-1649.38.6.934>
- Treiman, R., Kessler, B., y Pollo, T. C. (2006). Learning about the letter name subset of the vocabulary: Evidence from US and Brazilian preschoolers. *Applied Psycholinguistics, 27*(2), 211–227, [10.1017.S0142716406060255](https://doi.org/10.1017/S0142716406060255).
- Vivar, P., y León, H. (2009). Desarrollo fonológico-fonético en un grupo de niños entre 3 y 5, 11 años. *CEFAC, 11*(2), 190–198. <http://dx.doi.org/10.1590/S1516-18462009000200003>
- Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Fáiça, L., . . . Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*(4), 551–559. <http://dx.doi.org/10.1177/0956797610363406>