Revista de Psicodidáctica, 2018, 23 (2), 128–136

Revista de Psicodidáctica

www.elsevier.es/psicod

Original

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

Ana María De la Calle[a,*], Fernando Guzmán-Simón[b], and Eduardo García-Jiménez[c]

[a] Departamento de Psicología Evolutiva, Facultad de Ciencias de la Educación, Universidad de Cádiz, Cádiz, Spain
[b] Departamento de Didáctica de la Lengua y la Literatura y Filologías Integradas, Facultad de Ciencias de la Educación, Universidad de Sevilla, Sevilla, Spain
[c] Departamento de Métodos de Investigación y Diagnóstico en Educación, Facultad de Ciencias de la Educación, Universidad de Sevilla, Sevilla, Spain

ARTICLE INFO

Article history:
Received 3 April 2017
Accepted 11 October 2017
Available online 1 December 2017

Keywords:
Letters
Precursors
Early reading
Reading performance
Structural equation models

ABSTRACT

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 standardised instruments (PROLEC-R and BIL) applied to 362 students aged 4 and 5. The data obtained are analysed 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, optimise the reading performance of Spanish speaking children.

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

Palabras clave:
Grafías
Precursores
Lectura temprana
Rendimiento lector
Modelo de ecuaciones estructurales

RESUMEN

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.

PII of original article:S1136-1034(17)30071-0.

☆ Please cite this article as: de la Calle AM, Guzmán-Simón F, García-Jiménez E. El conocimiento de las grafías y la secuencia de aprendizaje de los grafemas en español: Precursores de la lectura temprana. Rev Psicodidáct. 2018;23:128–136. http://dx.doi.org/10.1016/j.psicod.2017.10.002

* Corresponding author.
E-mail address: anamaria.decalleca@gmail.com (A.M. De la Calle).

Introduction

Letter knowledge is important for learning the alphabet and implies the discovery of the rules relating *phonemes* to *graphemes*. *Letters* represent the minimal units of sound of a language (*phonemes*) with orthographic signs termed *graphemes* (Perry, Ziegler, & Zorzi, 2013). *Graphemes* are symbols perceptually determined by orientation and shape, which enables the reader to visually discriminate them (Lorenzo, 2001). Therefore, attentional skills are predominant in letter knowledge (Helal & Weil-Barais, 2015).

The mastery of *grapheme/phoneme* correspondences is key to reading achievement. The explicit instruction of *grapheme/phoneme* correspondences supports the acquisition of the alphabetic principle (Earle & Sayeski, 2017; Schaars, Segers, & Verhoeven, 2017). The systematic use of such correspondences after acquiring the alphabetic principle improves *decoding skills.* Therefore, *letter knowledge* is one of the cognitive skills that best predict future *reading performance* (Hammill, 2004; Huang, Tortorelli, & Invernizzi, 2014; Storch & Whitehurst, 2002).

The precursor nature of *letter-name and letter-sound knowledge* in early educational stages is regarded as an indicator of *reading performance* in many studies. Muter, Hulme, Snowling, and Stevenson (2004) evaluated *letter knowledge* at 4 and 5 years of age in 90 English subjects and its relation to *decoding skills* for 2 years. The research determined 63% of the variance of *decoding skills* 1 year after the acquisition of *letter-name and letter-sound knowledge*. Similarly, Evans, Bell, Shaw, Moretti, and Page (2006) evaluated *letter-name and letter-sound knowledge* in 149 English-speaking Canadian subjects at 5 years of age and, ten months later, *decoding skills*. In this case, lower-case *letter-name knowledge* explained 51% of the variance of *reading performance* of 1st-year students. Recently, Onochie-Quintanilla, Defior, and Simpson (2017) evaluated *letter-name and letter-sound knowledge* of 27 upper- and lower-case letters (and the diletters "ch" and "ll") in 100 Spanish early childhood education students (average age: 5.6 years) and *word-reading accuracy* in 1st-year primary education students. Their results reveal a lack of predictive value of *letter knowledge* for subsequent (i.e., months later) word reading.

The cited studies highlight the importance of research on *grapheme/phoneme* correspondence and its predictive value for later *reading development*. The findings of Caravolas, Lervåg, Defior, Seidlová Málková, and Hulme (2013); Fricke, Szczerbinski, Fox-Boyer, and Stackhouse (2016); Snel, Aarnoutse, Terwel, van Leeuwe, and van der Veld (2016) suggest *letter knowledge* as an early precursor of reading common to different orthographic systems (English, Spanish, Czech, German, and Dutch). However, our research considers that *grapheme/phoneme* correspondences vary according to the alphabetic code because the level of transparency of a language influences alphabet learning (Ziegler et al., 2010). In particular, our research focuses on the predictive value of *single-letter knowledge* for early *reading development* in the Spanish alphabet.

Learning sequence of graphemes

A number of studies that address the learning of the English alphabet have developed a *learning sequence of letters*. The relationship between each *single letter* and its *grapheme* facilitates evaluating the *knowledge of their associated graphemes* by studying *single-letter knowledge* at early reading stages. Thus, the previously mentioned studies show that the order of *letters* in the English alphabet and phonological development affects the *learning sequence of graphemes* (Jones, Clark, & Reutzel, 2013, Justice, Pence, Bowles, & Wiggins, 2006; Treiman, Kessler, & Pollo, 2006).

In particular, Justice et al. (2006) analysed knowledge of the 26 *letters* of the English alphabet in children between 3 and 5 years of age. Their findings reveal the *grapheme learning sequence* of the English alphabet based on a frequency analysis of *letter-name knowledge*. This sequence was determined through *graphemes* recognised by 50% of the subjects. Justice et al. (2006) also considered the relationship between the *learning sequence of graphemes* determined in their results and the *learning sequence of phonemes* determined by Sander (1972). The conclusions of Justice et al. (2006) emphasise the relationship between phonological and graphemic representations. Native English-speaking children first learn *graphemes* whose corresponding *phonemes* are acquired earlier in phonological development. The authors term this finding "the hypothesis of the order of consonants".

Learning sequence of phonemes

Studies on the learning of *phonemes* drew increasing research attention at the end of the 20th century. These studies analysed the phonetic features of *phonemes* (such as sonority and the point and mode of articulation) and age in relation to *phoneme* acquisition, primarily in English (Sander, 1972; Smit, Hand, Freilinger, Bernthal, & Bird, 1990).

Studies by Bosch (2004), Camargo (2006), and Vivar and León (2009) investigated Spanish phonological development. These studies adopted different methods with respect to the subject age ranges they considered and different criteria to determine the *learning sequences of phonemes*. Regarding the sample, Bosch (1983, 2004) and Vivar and León (2009) analysed the phonological repertoire from 3 years of age, whereas Camargo (2006) examined phonological development from the first months of life. Bosch (1983, 2004) and Vivar and León (2009) also studied the learning of *phonemes* up to 6 years. However, Camargo (2006) established the last phase in the learning of *phonemes* from 3 years of age. Regarding their results, the research by Bosch (1983, 2004) describes the acquisition of phonemic inventory between 3 and 7 years of age, classifying *phonemes* according to success rates in average-age groups (3 years, 4 years, 5 years, 6 years, and 7 years). Camargo (2006) classifies *phoneme* learning into 4 groups between 0 and 3 years. Finally, Vivar and León (2009) determine 5 groups in *phoneme* learning that range from 3 years to 5 years and 11 months.

The groups delimited by the different models implicitly define the *learning sequence of phonemes*, and the data analysis by these authors suggests discrepancies in data collection. More specifically, Bosch (1983, 2004) considered the position of *phonemes* within syllables in direct, inverse, or mixed position. In contrast, Camargo (2006) studied the position of *phonemes* within words (initial, middle, or final). Vivar and León (2009) excluded these criteria in their analysis.

Table 1 compares the *phoneme learning sequences* proposed by the previously mentioned authors. *Phonemes* are allocated to different groups when they are learned in any syllable or word position. Thus, the clusters of the cited models are comparable because there is a tendency to learn certain *phonemes* before others. However, the *phoneme learning sequences* in the different studies are not identical.

Our research on the Spanish alphabet

Our study analyses the relationships between *single-letter names, phonemes*, and their associated *graphemes* in the process of learning the Spanish alphabet. In addition, we investigate *single-letter knowledge* as a precursor of *reading performance.* Based on the

Table 1
Comparison of the learning of phonemes in Spanish according to the models of Bosch (1983, 2004), Camargo (2006), and Vivar and León (2009)

	Bosch (1983, 2004) (3:0–7:11)	Camargo (2006) (0:0–3:0)	Vivar and León (2009)(3:0–5:11)
Group1	/m/, /n/, /ɲ/, /p/, /b/, /k/, /g/, /t/, /f/, /x/	/p/, /b/, /m/, /t/, /k/, /g/, /s/	/p/, /m/, /t/
Group 2	/l/, /d/, /y/	/n/, /ɲ/, /d/, /x/, /f/, /θ/, /l/, /ks/	/b/, /n/, /ɲ/
Group 3	/s/	/r/, /r̄/	/x/, /k/
Group 4	/r/, /r̄/	–	/l/
Group 5	–	–	/d/, /r/
Group 6	–	–	/r̄/, /g/, /s/, /f/, /ks/

Table 2
Cohorts of the sample

	Early childhood education at age 4 (4:0–5:4)	Early childhood education at age 5 (5:0–6:4)	1st Grade(6:0–6:10)
Cohort 1	180 (t_1)	180 (t_2)	
Cohort 2		182 (t_1)	182 (t_2)

literature and our understanding of the issue, we seek to answer the following questions:

1. Is it possible to define a learning sequence of the Spanish alphabet focused on *graphemes*?
2. Is there a relationship between the *grapheme learning sequence* and the *phoneme learning sequence* in Spanish?
3. What predictive value does *letter knowledge* have for early *reading performance*?

To answer these questions, our study examines the *single-letter* knowledge (i.e., associated with a single *grapheme*) of children 4 and 5 years of age and its predictive value for *decoding skills* in early reading achievement. First, it is necessary to specify the *grapheme learning sequence* in Spanish based on *letter knowledge*. Subsequently, the *phoneme learning sequences* proposed by Bosch (1983, 2004), Camargo (2006), and Vivar and León (2009) are compared with the *grapheme learning sequence* determined in our study. Finally, the predictive value of *letter knowledge* is analysed as criterion variable for *decoding skills* (i.e., a dependent variable).

Our research focuses on the process of learning *graphemes*, the relation of that knowledge to later *reading performance*, and on clarifying how the Spanish alphabet is learned. The distinctive character of this research compared to the studies of Evans et al. (2006), Muter et al. (2004), and Onochie-Quintanilla et al. (2017) lies in its delimitation of the *learning sequence of graphemes* in Spanish and its comparison with phonological development.

Method

Participants

The study involved 362 students enrolled in three schools in the province of Cádiz. These schools were selected because they have the same socio-economic indicator (SEI) and similar characteristics. That is, they are public schools of childhood and primary education under the same regulations with experienced teaching staffs. Special educational needs were considered exclusion criteria for participation in the study. In addition, student selection required the informed consent of the students' parents. This study follows the regulations for social sciences of the Research Ethical Committees at the universities of Cádiz and Seville.

The participants were grouped into cohorts. Each cohort was individually evaluated at two different time points (t_1 and t_2). The first cohort consisted of 180 early childhood education students who were 4 years of age (*mean* age = 4.7; 45.6% female), and the second cohort consisted of 182 early childhood education students who were 5 years of age (*mean* age = 5.7; 48.6% girls) (Table 2). Regarding the assessment timeline, t_1 corresponds to the second trimester of the school year, when *letter knowledge* was measured, and t_2 corresponds to the beginning of the first trimester of the following school year (6 months after t_1), when *decoding skills* were measured.

Instruments

At the first time point (t_1), *letter knowledge* subtests of two instruments were administered.

1. Letter knowledge subtest of the Start Reading Battery (BIL 3–6) (Sellés, Martínez, & Vidal-Abarca, 2010). The authors determined a Cronbach's alpha coefficient of .97 for this subtest, which tests knowledge of the five lower-case vowels. In the current study, Cronbach's alpha (α) is .77, McDonald's omega (Ω) is .83, and the average variance extracted (AVE) is .50.
2. Letter recognition subtest of the Reading Processes Assessment Battery (PROLEC-R) (Cuetos, Rodríguez, Ruano, & Arribas, 2007). For this subtest, the authors found a Cronbach's alpha of .49, which was exceeded in our measurements (α = .76). This subtest measures knowledge of 20 lower-case letters. However, our study considers the correctness of the consonants (maximum score = 19). The value of Ω in our study is .92, while the AVE is .41.

At the second time point (t_2), the *word-reading* subtest (s_1) and *pseudoword-reading* subtest (s_2) of PROLEC-R were administered. The authors determined a Cronbach's alpha of .74 for s_1 and .68 for s_2. These subtests consist of reading 40 words and 40 pseudowords, both lower case. Factor analysis of the total scores of these subtests using the main components extraction method determined the *reading accuracy* index (one component; s_1 = .994 and s_2 = .994). The Cronbach's alpha in our measurements for reading accuracy is .93. The value of Ω is .99, while the AVE is .98.

Procedure

Assessments were performed by one researcher in a private space under appropriate environmental conditions. Because all participants were individually assessed in two different sessions (session t_1 and session t_2), the assessments were performed over several months. Session t_1 measured the *knowledge* of 24 of the 30 *letters* of the Spanish alphabet, corresponding to *single letters* associated with a single grapheme (except "h", "k", and "w"). We chose to assess *lower-case letter recognition* because these letters

Table 3
Classification of students according to grapheme learning sequence

		At 4 age (4:0–5:4)		At 5 age (5:0–6:4)	
		Frequency	Percentage	Frequency	Percentage
Cluster	1	127	70.9%	18	10.1%
	2	50	27.9%	161	89.9%
	Atypical value (−1)	2	1.1%	0	.0%
	Combined	179	100.0%	179	100.0%

appear more frequently in texts (Jones & Mewhort, 2004). BIL and PROLEC-R batteries facilitate using a larger number of evaluated items with respect to single vowels and consonant *letters* of the Spanish alphabet. The response was considered correct in *letter recognition* if the child identified the *letter name* or *phoneme*, or the *phoneme* only in the syllable phonological structure. The last criterion was considered because *phoneme* identification in syllable structure is easier to discriminate and produce at these ages by Spanish subjects (Defior & Serrano, 2014). Session t_2 measured *lower-case word and pseudoword decoding skills* by PROLEC-R. The selection of the same instrument reduces the effects of calligraphy in the recognition of the perceptual characteristics of *graphemes*.

Data analysis

To answer the research questions, different types of analysis were developed using SPSS Amos version 23. Inferential data analysis required non-parametric tests because not all assumptions of normality were met. First, two-stage cluster analysis was performed by dichotomising the data (1 = error, 2 = success). From this analysis, possible significant differences between age groups were established by classifying the subjects based on the results obtained for the recognition of each letter. Second, the *letters* that differentiated the subjects according to age were delimited by discriminant analysis supported by the stepwise inclusion method.

To determine *letter knowledge* accuracy and the sequence in which the associated *graphemes* are learned, frequency analysis was performed for the group of 4-year-olds, the group of 5-year-olds, and for both groups together. Then, *letters* were grouped according to frequency intervals (quintiles), which enabled the sequence of *grapheme* knowledge accuracy to be determined. *Graphemes* were considered to be known by the age groups when recognised by 50% of the member children (Justice et al., 2006). To confirm the *grapheme learning sequence*, derived from frequency analysis, a hierarchical cluster was created using Ward's method and the squared Euclidean distance. Next, a comparison between the *grapheme learning sequence* and the *phoneme learning sequence* was performed based on the Bosch (1983, 2004), Camargo (2006), and Vivar and León (2009) models. Finally, we analysed the relationship between *letter knowledge* in early childhood and *decoding skills* in early reading achievement using structural equation modelling.

Results

Alphabet learning focused on graphemes

Two-stage cluster analysis identified two different sequences in *letter* knowledge according to age group. *Letter knowledge*—assessed with 24 *single letters*—correctly classified 80.4% of the sample under the criterion of error or success according to age group (Table 3). Cluster 1 included the sequence of the 4-year-old students (70.9%), and cluster 2 included that of the 5-year-old students (90%).

Discriminant analysis enabled letter classification according to the age group of the participants. Canonical discriminant functions exhibited the following values: eigenvalue = 1.323, canonical correlation coefficient = .755, Wilks's lambda = .430 (p = .001). The value of the discriminant function was 1147 for the 4-year olds and the opposite for the 5-year olds (−1147). Figure 1 shows the structure matrix values, which indicate the weight of each *grapheme* when age acts as a discriminant function in *letter knowledge*. According to age group, the *graphemes* that best discriminate subjects are those associated with ten *letters*: "t", "n", "d", "m", "ñ", "r", "a", "y", "u", and "j".

Subsequently, we analysed the extent to which *letter knowledge* reflects the *grapheme learning sequence* in the children of an age group. Table 4 shows the frequency distribution of *letter knowledge* with five groups in *grapheme* learning. Because the differences between the highest frequencies of *graphemes* in different quintiles were greater than the group interval (20%), the assignment of *graphemes* to each group is justified. Children of different age groups displayed at least consolidated learning of recognised *graphemes* in the first two groups (≥50%).

The frequency distribution results were confirmed from the results of the hierarchical cluster analysis. The dendogram shows *letter knowledge* at 4 and 5 years of age, and the *graphemes* are classified into 4 main groups according to distance (Figure 2).

Table 5 shows that the distance analysis of *letter* pairs identifies the *learning sequence of graphemes* in four groups. *Grapheme* grouping by distance analysis maintains the sequence determined by the frequency analysis but exhibits better accuracy by grouping the *graphemes* in four rather than five groups. Thus, distance analysis definitively classified the *graphemes* into four groups.

Letter knowledge and decoding skills

First-year primary education students (i.e., 5-year-old early childhood education students at t_1) had greater *decoding skills* than the 4-year-old students (4-year-old early childhood education students at t_1). The average of words and pseudowords correctly read by the 5-year-old students was 3 of 80 words, while the first-year primary education students read 50 of 80 words. *Decoding skills* increased significantly with age (Table 6).

Word and pseudoword decoding skills in 5-year-old early childhood education students and first-year primary education students were determined by prior *letter* knowledge (6 months earlier). Structural equation modelling was used to determine the relationship between *letter knowledge* at 4 and 5 years of age and *decoding skills* in early reading achievement. A first model considered the predictive value of all *letters* for *reading accuracy* at 5 and 6 years

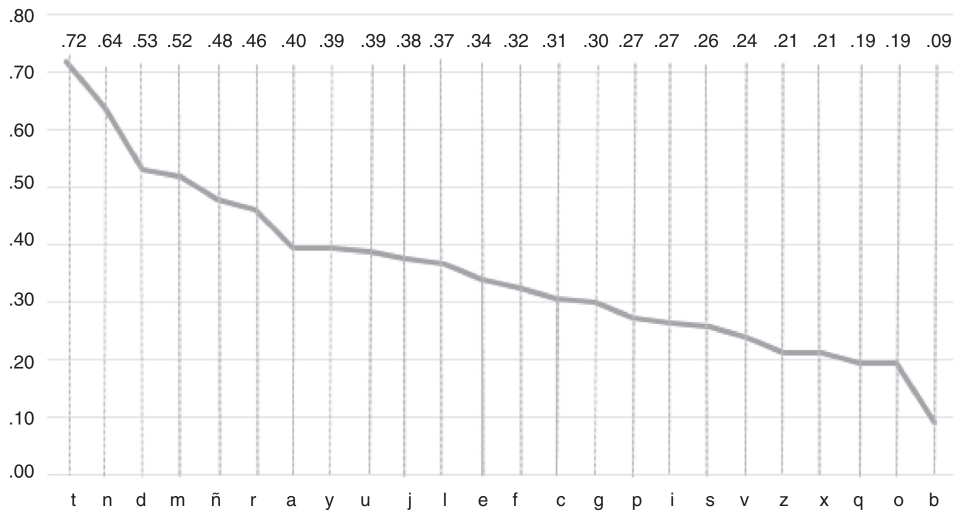

Figure 1. Structure matrix representation for letter knowledge, according to age.

Table 4
Learning sequences of graphemes according to frequency analysis of letter knowledge

	Early childhood education at age 4 (4:0–5:4)	Early childhood education at ages 4–5 (4:0–6:4)	Early childhood education at age 5 (5:0–6:4)
Group 1(80–100%)	"o"	"o", "i", "s", "e", "u"	"o", "e", "u", "s", "i", "a", "m", "z", "p", "t"
Group 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"
Group 3 (40–60%)	"x", "v", "y", "m"	"v", "ñ", "y", "d", "t", "n", "r", "j"	"j", "f"
Group 4 (20–40%)	"ñ", "d", "j", "f", "r"	"f", "g", "l"	"l", "g", "b"
Group 5 (0–20%)	"n", "t", "b", "g", "l", "q"	"b", "q"	"q"

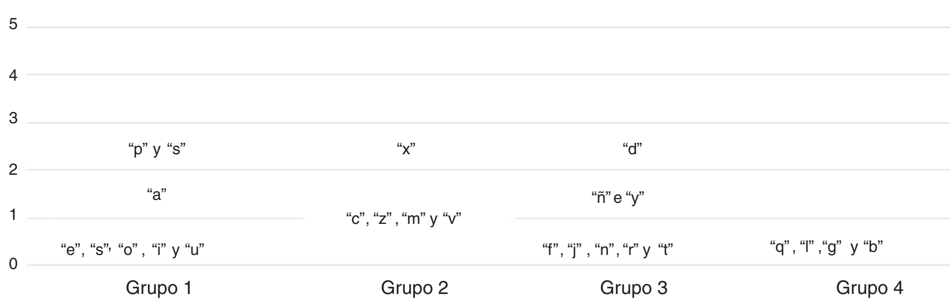

Figure 2. Dendogram of letter knowledge at 4 and 5 years of age.

Table 5
Comparison between learning sequences of graphemes according to frequency and distance analysis in letter knowledge at 4–5 years of age

	According to *frequencies*	According to *distances*
Group 1	"o", "i", "s", "e", "u"	"e", "s", "o", "i", "u", "a", "p"
Group 2	"z", "a", "p", "c", "x", "m"	"c", "z", "m", "v", "x"
Group 3	"v", "ñ", "y", "d", "t", "n", "r", "j"	"f", "j", "n", "r", "t", "ñ", "d", "y"
Group 4	"f", "g", "l"	"q", "l", "g", "b"
Group 5	"b", "q"	–

and achieved unacceptable goodness of fit values: $\chi^2 = 826.291$, $df = 275$, $p = .000$, NFI = .797, TLI = .826, CFI = .853, and RMSEA = .075 (Figure 3).

Figure 4 shows a second model that considered *letters* with the highest predictive value for *decoding*. These *letters* have associated *graphemes* that are learned in intermediate groups in the learning sequence, delimited according to frequency analysis at 4 and 5 years. *Letters* with the highest predictive value (except "f" and "m") were recognised by 40–60% of the participants (Table 4). In this case, good model fit was achieved: $\chi^2 = 36.223$, $df = 26$, $p = .088$, NFI = .979, TLI = .992, CFI = .994, and RMSEA = .033.

Discussion

Learning sequence of graphemes

Our research studies the learning sequence of the Spanish alphabet focused on *graphemes*. *Single-letter knowledge* differ-

Table 6
Average results for decoding skills

	Mean (*SD*)			Rank
	Early childhood education at age 5 (5:0–6:4)	Early childhood education at age 5 and 1st Grade (5:0–6:10)	1st Grade (6:0–6:10)	
Words reading	1.69 (6.62)	13.97 (16.13)	26.33 (13.15)	0–40
Pseudowords reading	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

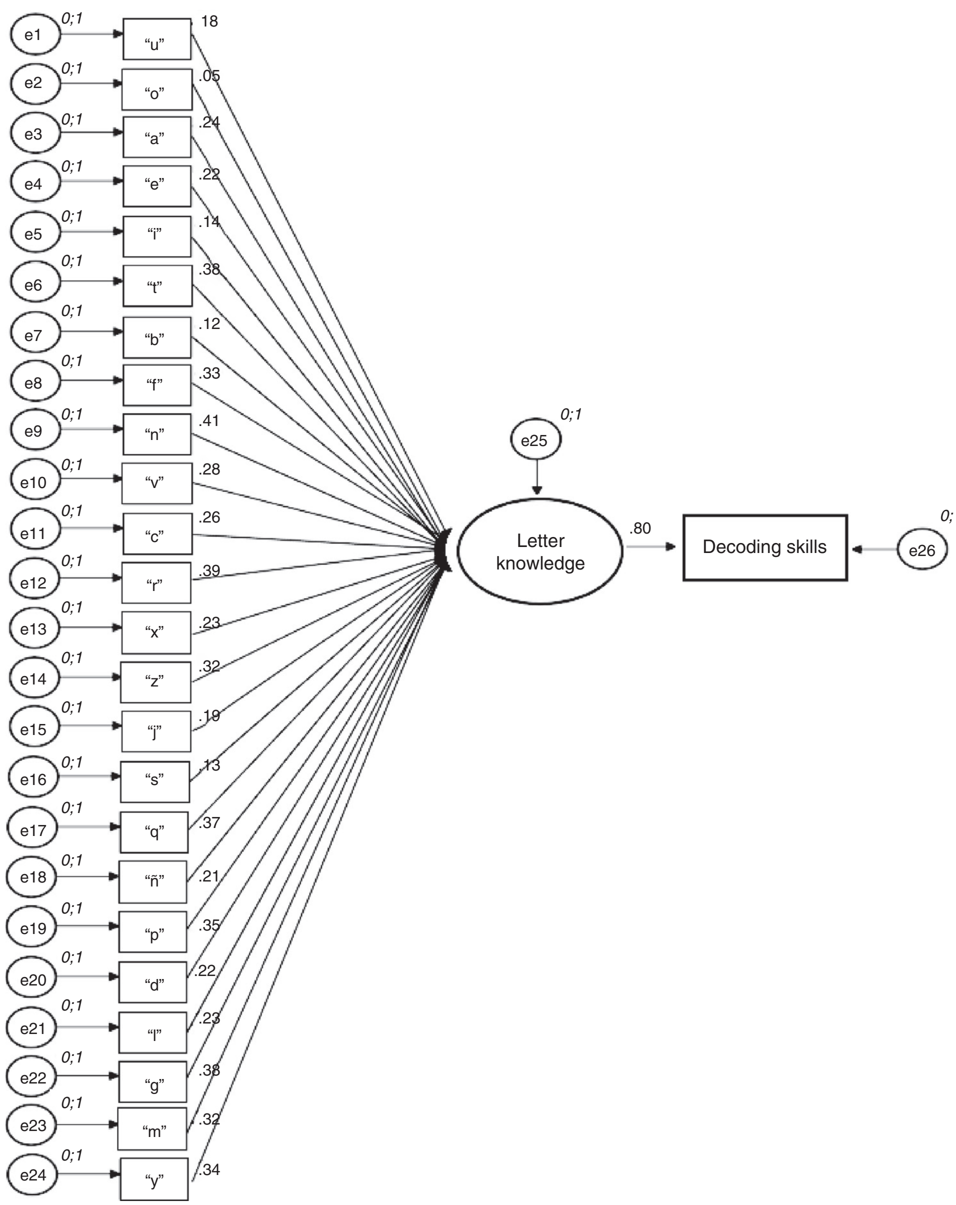

Figure 3. Letter knowledge predictive model for decoding skills I.

ences, according to age, suggest different levels of development in learning *graphemes*. At 4 years, the children exhibited consolidated learning of at least all *vowels* and five *consonants* ("s", "z", "c", "p", and "x"). One year later, the 5-year-old children consolidated eight more *consonants* ("t", "m", "n", "d", "ñ", "r", "v", and "y"). *Letter knowledge* delimited grapheme learning in four groups (Table 5). *Vowels* were among the first-learned *graphemes*, while the *graphemes* "q", "l", "g", and "b" were later-learned *graphemes*. The *graphemes* "g" and "q" correspond to *letters* less recognised by their *name and sound* at 5 years of age in Evans et al. (2006).

The *grapheme learning sequence* in our research differs from that of studies that consider the order of *letters* in the alphabetic chain and phonological development to be relevant. The divergences found in the *grapheme learning sequence* may be caused by the different levels of orthographic transparency among the studied languages (Defior & Serrano, 2014; Ziegler et al., 2010). The studies of Jones et al. (2013), Justice et al. (2006), and

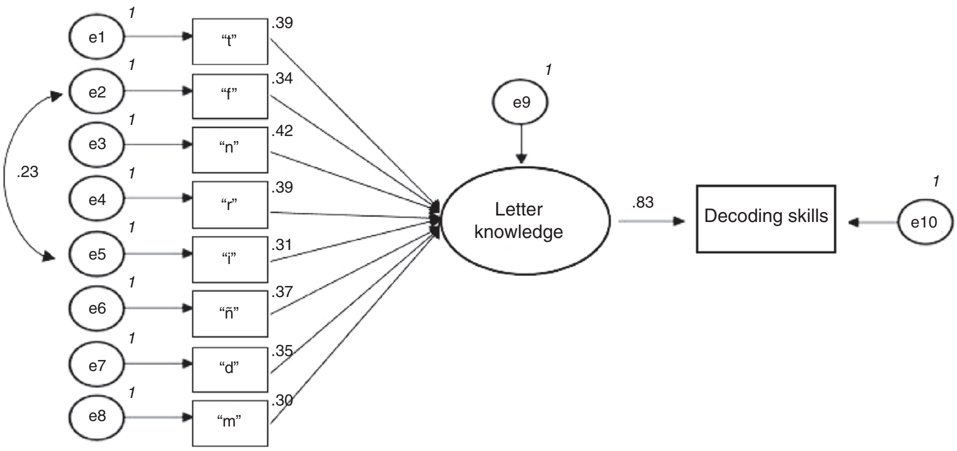

Figure 4. Letter knowledge predictive model for decoding skills II.

Treiman et al. (2006) consider the order of *letters* in the alphabetic chain in English to be relevant. However, our results with Spanish subjects are the opposite. The exposure of English-speaking students to the teaching of the alphabet could be greater than that of Spanish students. Alternatively, the exposure to the teaching of the alphabet might not affect the learning sequence of *graphemes* to the same extent as in English. This last explanation is consistent with the conclusions of Treiman et al. (2006), who noted differences in alphabet knowledge between English and Portuguese children and concluded that American subjects made fewer errors than Portuguese subjects in recognising the first letters of the alphabet. Our study on the Spanish language contradicts research findings by Justice et al. (2006), where *grapheme* learning is found to be related to phonological development in English. While Justice et al. state that English-speaking children first learn those graphemes whose corresponding *phonemes* are acquired earlier in the phonological repertoire, our study argues that Spanish-speaking children do not learn *graphemes* in correspondence to their *phonemic* development.

Learning sequence of graphemes versus learning sequence of phonemes

This study examined the *learning sequence of graphemes* in Spanish speakers between 4 and 6 years of age and established its independent character with respect to phonological development. Table 7 compares the learning sequences of *graphemes* and *phonemes* according to our results with the models of Bosch (1983, 2004), Camargo (2006), and Vivar and León (2009). For example, the *grapheme* “s” is learned first. However, according to Bosch (1983, 2004) and Vivar and León (2009), the *phoneme* /s/ is acquired later by the child. In addition, the *grapheme* “b” is the last to be learned, although the *phoneme* /b/ is assimilated early (Bosch, 1983, 2004; Vivar & León, 2009).

The learning sequence of *graphemes* differs from the learning sequence of *phonemes* in the models of Bosch (1983, 2004), Camargo (2006), and Vivar and León (2009). The age of the participants enables us to compare our *grapheme learning sequence* with the phonological repertoire of children between 3 and 6 years of age (Vivar & León, 2009) and to state that the *grapheme learning sequence* in Spanish is not based on phonological development.

Letter knowledge and decoding skills

Letter knowledge in early childhood influences later *reading performance*, as verified in our predictive model (Figure 3). Our study confirms that *letter-name and letter-phoneme knowledge* is a good precursor of *reading performance* in early reading achievement, which is also confirmed by Evans et al. (2006) and Muter et al. (2004). The results of our analysis of the Spanish language enable us to establish a causal relationship between *letter knowledge* in early childhood and the development of *decoding skills* in early reading achievement. Knowledge of the *graphemes* “t”, “f”, “n”, “r”, “j”, “ñ”, “d”, and “m” at 4 and 5 years explains 83% of the variance of *decoding skills* 6 months later. The results of our study reveal that these

Table 7
Comparison between learning sequence development of graphemes and phonemes according to our results and the models of Bosch (1983, 2004), Camargo (2006), and Vivar and León (2009)

Sequence development of graphemes (4:0–6:4)		Sequence development of phonemes		
		Bosch (1983, 2004) (3:0–7:11)	Camargo (2006) (0:0–3:0)	Vivar and León (2009) (3:0–5:11)
Group 1	“s”, “p”	/m/, /n/, /ɲ/, /p/, /b/, /k/, /g/, /t/, /f/, /x/	/p/, /b/, /m/, /t/, /k/, /g/, /s/	/p/, /m/, /t/
Group 2	“c”, “z”, “m”, “v”, “x”	/l/, /d/, /y/	/n/, /ɲ/, /d/, /x/, /f/, /θ/, /l/, /ks/	/b/, /n/, /ɲ/
Group 3	“f”, “j”, “n”, “r”, “t”, “ñ”, “d”, “y”	/s/	/r/, /r̄/	/x/, /k/
Group 4	“q”, “l”, “g”, “b”	/r/, /r̄/	–	/l/
Group 5	–	–	–	/d/, /r/
Group 6	–	–	–	/rr/, /g/, /s/, /f/, /ks/

graphemes have the highest predictive value and (except “f” and “m”) are learned by 40–60% of the subjects. This result is consistent with the criterion (50%) established by Justice et al. (2006) to delimit the *grapheme learning sequence*. The correlation between errors in the knowledge of the *letters* “f” and “j” suggest they are related by an underlying factor. One possible explanation is that these *letters* are more difficult to discriminate orthographically. Huang and Invernizzi (2012) showed that the *letters* “b”, “d”, “p”, and “q” are 8% more likely to be incorrectly named than less visually confusing *letters*.

The predictive value of *decoding skills* for reading achievement lies not only in *letter knowledge* but also in learning a given sequence of *graphemes*. The studies of Evans et al. (2006) and Muter et al. (2004) consider the predictive value of the 26 *letters* of the English alphabet for later *decoding skills*. Knowledge of all *letters* explained 51% and 63% of the variance of *decoding skills*. In addition, the study of Onochie-Quintanilla et al. (2017) does not consider the predictive value of the 27 *single letters* and two diletters of the Spanish alphabet for the variance of later *decoding skills*. However, our study considers the predictive value of several *letters* and can explain a greater amount of variance (83%). *Letter knowledge* is a better precursor when *letters* are considered according to the *learning sequence of graphemes*, that is, by considering *graphemes* that have been learned in approximately half of the cases.

Our study confirms that *decoding skills* improve when subjects acquire the alphabetic principle (Huang et al., 2014) Similarly, these findings reaffirm that the lack of *letter* knowledge, which must be acquired in correspondence with the developmental stage of the subject, is a risk factor for reading achievement (Hammill, 2004; Storch & Whitehurst, 2002). According to this study, Spanish-speaking children consolidate *graphemes* following a learning sequence independent of phonological development. In addition, *letter knowledge* in early childhood, in correspondence with the *grapheme learning sequence*, is a good precursor of later *reading performance*. Therefore, educational strategies in the teaching of the Spanish alphabet optimise early reading performance when they incorporate *letter names, phonemes*, and the *learning sequence of graphemes* in learning. As a result, this study may represent the beginning of a new line of research in the learning of the Spanish alphabet focused on *graphemes* because of the importance of the *grapheme learning sequence* for reading attainment.

Limitations and lines of future research

One limitation of our study is that the perceptual characteristics of *graphemes* and their relation with the cognitive capacity of the studied subjects were not expressly considered, which should be the object of future research.

Another limitation of this research is the lack of assessment of the *single letters* “h”, “k”, and “w”. In addition, assessing *letter knowledge* should be performed in smaller age intervals than those considered here to more precisely delimit the *learning sequence of graphemes* in Spanish.

Finally, to perform a comparative study under the same age criterion, our research requires an analysis of the grapheme repertoire at earlier and later ages. This analysis would facilitate a more comprehensive comparison of our results with the models developed by Bosch (1983, 2004) and Camargo (2006). In addition, future research should investigate the methodological approach and contextual factors that influence *grapheme* learning in Spanish and the predictive value of *letter knowledge* as a precursor of *reading performance* over time through longitudinal tracking.

References

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. *Umbral Científico*, *9*, 5–9.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & 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., & Arribas, D. (2007). *PROLEC-R*. Madrid: TEA ediciones.

Defior, S., & 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., & 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., & 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., & 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/rrq.116

Hammill, D. D. (2004). What we know about correlates of reading. *Exceptional Children*, *70*(4), 453–469.

Helal, S., & 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., & Invernizzi, M. A. (2012). Factors associated with lowercase alphabet naming in kindergarteners. *Applied Psycholinguistics*, *35*(6), 943–968. http://dx.doi.org/10.1017/S0142716412000604

Huang, F. L., Tortorelli, L. S., & 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., & 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., & 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., & 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., & 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., & 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., & 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., & 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., & 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., & 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., & 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., & 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., & 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. http://dx.doi.org/10.1017/S0142716406060255

Vivar, P., & 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., Faísca, 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