Journal of Child Language (2018), **45**, 858–877
doi:10.1017/S0305000917000514

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

The role of sublexical variables in reading fluency development among Spanish children

Marta ÁLVAREZ-CAÑIZO*, Paz SUÁREZ-COALLA, and Fernando CUETOS

Departament of Psychology, University of Oviedo, Asturias, Spain
*Corresponding author. Department of Psychology, University of Oviedo, Pza Feijóo s/n, Oviedo 33003, Spain. E-mail: alvarezcanmarta@uniovi.es

(Received 22 June 2016; revised 13 February 2017; accepted 26 September 2017; first published online 19 February 2018)

Abstract
Several studies have found that, after repeated exposure to new words, children form orthographic representations that allow them to read those words faster and more fluently. However, these studies did not take into account variables related to the words. The aim of this study was to investigate the influence of sublexical variables on the formation of orthographic representations of words by Spanish children. The first experiment used pseudo-words of varying syllabic structure and syllabic frequency. The stimuli for the second experiment were formed with or without context-dependent graphemes. We found that formation of orthographic representations was influenced by syllabic structure (easier for words with simple syllabic structure) and the context-dependency of graphemes (easier in the absence of context-dependent graphemes), but not syllabic frequency. These results indicate that the easier it is to read a word, the easier it is to form an orthographic representation of it.

Keywords: orthographic representations; context-dependent graphemes; syllabic frequency and syllabic structure

In order to read fluently, children need to form orthographic representations of words that appear frequently in the texts they read. When they are learning to read they have to use a sublexical strategy: reading the letters of words one by one (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). This is a very slow method of reading, especially for long words with many letters to decode. An important step in learning to read is learning to read words by identifying the letters in parallel. This amounts to a lexical strategy that allows the word to be read as a whole. This kind of strategy depends on having representations of the words in an orthographic lexicon (Share, 1995). Each successful reading of a new word is an opportunity for the child to acquire orthographic information about that word, and thus gradually form orthographic representations (Share, 1995) that allow them to recognise that word as a unit and read it faster. Therefore, orthographic representations are considered to be a trace stored in the orthographic lexicon, which allow for reading the words as a whole. The self-teaching hypothesis (Share, 1995) assumes that lexical and sublexical reading processes co-exist throughout life. Through self-teaching, children develop orthographic knowledge of new words independently; instruction is not required and self-teaching begins when children start reading.

Various methods have been used to determine whether children have formed orthographic representations of new words. In some studies with English children, participants were asked to write the new words after a training period and correct spelling was taken as an indication of orthographic learning (Cunningham, Perry, Stanovich, & Share, 2002). Other studies, with Hebrew children, assessed recognition of a new word by requiring participants to select it from four options with similar spellings (Share, 1999, 2004); yet others analysed reading times, based on the assumption that a significant reduction in reading time signifies orthographic learning in English children (Ehri & Saltmarsh, 1995). However, the problem with all these methods is that the results, which are assumed to indicate orthographic learning, might also be due to an episodic memory trace or to familiarity with the new words as a result of repeated exposure (Qiao, Forster, & Witzel, 2009). Orthographic learning means that the orthographic representation has been formed in the visual lexicon; however, repeated exposure to words could make them familiar or keep them in episodic memory, which also reduces reading times. To avoid these problems, some authors have begun to use a different method, based on the decrease in the length effect on reaction or articulation times that occurs after many exposures to words in reading tasks. Length effect can be described as the difference between short and long stimuli, where long stimuli have longer reading times and also less accuracy than short stimuli (Just & Carpenter, 1980).

Previous research has shown that length effect is affected in different ways for pseudo-words, low-frequency words, and high-frequency words in English adults (Weekes, 1997). In high-frequency words, no length effect was found, while in pseudo-words the effect was significant, the reaction times for long pseudo-words being longer than those for short stimuli. The effect was also studied in repeated reading, using the same stimuli (Maloney, Risko, O'Malley, & Besner, 2009). These authors found that the length effect diminished across the blocks, as reading times for short and long pseudo-words converged. Thus, this method assumes that the difference in reaction time for short and long words (the length effect) will be greater on first exposure to new words, because at this point children are reading the words using a sublexical route and therefore need more time to decode each letter into a sound. The length effect disappears when they start reading using the lexical route, which involves decoding all the letters of a word in parallel. A decrease in the length effect is thus an indication of the formation of orthographic representations of new words and the use of a lexical route. Consequently, the use of the length effect ensures that what is being measured is the formation of orthographic representations rather than familiarity or episodic memory effects; since familiarity and episodic memory traces would similarly affect reaction times for both short and long new words, the convergence in reaction times of short and long new words would indicate a parallel identification of letters, which corresponds to lexical reading. As a consequence of these results, length effect as a measure of orthographic representations has been used by various authors (Kwok, Cuetos, Avdyli, & Ellis, 2017; Kwok & Ellis, 2014, 2015; Suárez-Coalla, Álvarez-Cañizo, & Cuetos, 2016; Suárez-Coalla, Avdyli, & Cuetos, 2014; Suárez-Coalla, Ramos, Álvarez-Cañizo, & Cuetos, 2014). Using the decrease in length effect measure in Spanish, Suárez-Coalla *et al.* (2016), found that children in primary education started forming orthographic representations of new words, presented either in isolation or within a text, after six exposures.

However, few studies have analysed sublexical variables that might influence orthographic learning. Despite some studies exploring the role of orthographic consistency in reading times (Barca, Ellis, & Burani, 2007; Peereman, Dufour, & Burt, 2009; Ventura, Kolinsky, Pattamadilok, & Morais, 2008), we know of only two studies of English-reading children that investigated the role of orthographic consistency in the formation of orthographic representations. In the first (Wang, Castles, & Nickels, 2012), second-graders were presented with eight new regular and irregular pseudo-words. After training in the phonology and meaning of the new words, children were introduced to their spellings. Ten days later, their orthographic learning was evaluated using an orthographic decision task and a spelling task. In the orthographic decision task, the experimental words were presented with three distracters and the child was required to identify the experimental words. The results revealed that children formed orthographic representations of regular words more easily, since irregular words were both coded (in the spelling task) and decoded (in the orthographic decision task) less precisely. In a subsequent study, Wang, Nickels, Nation, and Castles (2013) followed a similar procedure in order to explore how children's language skills were associated with orthographic learning and how these factors were influenced by word regularity. They found that reading and language skills, such as phonetic decoding skills, orthographic knowledge, and vocabulary knowledge, were related to orthographic learning of both regular and irregular words. However, for orthographic learning of irregular words, knowledge of meaning and pronunciation of the new words made their learning easier. These studies are very relevant because they allow us to understand the regularity effect on the formation of orthographic representations in an opaque language, where the grapheme–phoneme correspondence is low.

Spanish has a transparent orthographic system, meaning that once a child has learned the grapheme–phoneme conversion rules, it is possible for him or her to pronounce almost any Spanish word. In fact, after one year of learning to read a transparent language, children read with an accuracy of 95% (Seymour, Aro, & Erskine, 2003). Because of its transparent orthography, a clear understanding of the syllable boundaries in Spanish is established in the first years of literacy acquisition (Álvarez, García-Saavedra, Luque, & Taft, 2017). Syllables are fundamental units of processing in the Spanish language, being phonological sublexical units in visual word recognition (Álvarez, Carreiras, & Perea, 2004; Jiménez, García, O'Shanahan, & Rojas, 2010). Spanish has some exceptions for grapheme–phoneme correspondence, since the pronunciation of 'c' and 'g' graphemes is dependent on the following vowel, which means that the reader must advance to the next grapheme in order to pronounce them correctly ('c' followed by 'a', 'o', or 'u' is pronounced /k/, but followed by 'e' or 'i', it is pronounced /θ/; 'g' followed by 'a', 'o', or 'u' is pronounced /ɣ /, but followed by 'e' or 'i', it is pronounced /x/).

There has been some research into the influence of these syllabic variables on reading speed. Studies that manipulated syllabic frequency in lexical decision tasks in Spanish (Álvarez, Carreiras, & Taft, 2001; Carreiras, Álvarez, & de Vega, 1993; Perea & Carreiras, 1995, 1998) have shown that words are recognised more slowly if they have a high-frequency first syllable. Carreiras and Perea (2004) found that in naming tasks using Spanish words with a high frequency, the first syllable was read faster than in other words; when the high-frequency syllable was in second position there was no effect on reading speed.

Syllabic structure is also important in reading. We define syllabic structure as the different options of vowels and consonants organised within a syllable (e.g., CV, VC, CCV, or CVC). A simple structure was shown to facilitate visual recognition of words in a lexical decision task with English adults (Taft, 1979). A similar effect has

https://doi.org/10.1017/S0305000917000514 Published online by Cambridge University Press

been observed in French, where children found syllables with a simple structure easier to read and spell (Sprenger-Charolles & Siegel, 1997).

The role of context dependency in reading has been investigated in several transparent languages. In Portuguese, children had shorter reading latencies in a lexical decision task when pseudo-words consisted of context-independent graphemes (Ventura *et al.*, 2008). Similar results were found for reading in other languages: reading latencies were shorter in French adults and Italian children for stimuli comprising context-independent graphemes (Barca *et al.*, 2007; Peereman *et al.*, 2009).

Given that, in transparent languages, the presence of context-dependent graphemes influences reading times, it seems likely that it also plays an important role in the formation of orthographic representations. This could be because, as the self-teaching hypothesis (Share, 1995) postulates, orthographic representations are based on phonological decoding. The decoding of a context-dependent grapheme depends on the next letter, which complicates the grapheme–phoneme conversion process and hence makes formation of orthographic representations more difficult. A similar argument can be applied to syllabic variables; because they influence reading it is plausible that they also influence the formation of orthographic representations. We have seen that syllabic structure influences reading times in opaque languages and that syllabic frequency does so in Spanish. Taking into account the transparency of Spanish and the importance of the syllable in the processing of written words, we hypothesise that a complex or low-frequency syllable could render decoding in reading more difficult, which in turn causes difficulty when forming orthographic representations.

The aim of the experiments in this paper was to study the role of several sublexical variables, such as syllabic frequency, syllabic structure, and the context dependency of the graphemes, in the formation of orthographic representations in third-grade children; some of the reviewed studies above have shown that these variables affect reading times in Spanish or orthographic learning in opaque languages. We constructed two reading aloud tasks of isolated pseudo-words repeated along six blocks. In the first one, stimuli were constructed by manipulating syllabic frequency and syllabic structure. The stimuli in the second task were created with or without context-dependent graphemes. We selected children in the third grade, in common with other studies of this type, because it is an important grade in the learning-to-read process (Cunningham & Stanovich, 1990; Kyte & Johnson, 2006) as children start to acquire the reading skills needed for fluency.

Experiment 1

With this experiment we sought to explore the role of syllabic frequency and syllabic structure in the formation of orthographic representations of third-grade children. We constructed a reading aloud task consisting of six blocks of a set of 16 pseudo-words varying in length, syllabic frequency, and syllabic structure. We predicted that simple structure and high frequency would facilitate formation of orthographic representations, since this kind of syllable facilitates reading. Based on the self-teaching hypothesis (Share, 1995), a correct reading of new words when repeated allows for the formation of orthographic representations. Therefore, we expected that the length effect would disappear sooner in the pseudo-words formed by simple structure and high-frequency syllables. Consequently, we predicted that children would form orthographic representations after six exposures to the new words (as was found in previous studies: Bowey & Muller, 2005; Kwok & Ellis, 2014;

Maloney *et al.*, 2009; Suárez-Coalla *et al.*, 2016), which would be indicated by a significant Block * Length interaction (i.e., the length effect decreases significantly along exposures). Moreover, the aim of our study was to identify the possible influence of syllabic variables in the formation of orthographic representations. This would be shown by a significant three-way interaction Block * Length * Syllabic structure/frequency; this interaction would mean that the syllabic variable modulates the decrease of the length effect along exposures to new words.

Method

Participants

The study was conducted with 34 children in the third grade (14 boys and 20 girls; age $M = 8.6$ years, $SD = 0.34$) from a state-assisted school in Oviedo. Selection of participants was based on tests of reading of words and non-words in PROLEC-R (Cuetos, Rodríguez, Ruano, & Arribas, 2007). We only included children with normal results according to the battery's established age ranges. PROLEC-R is a standardised battery for the assessment of reading in Spanish children aged between six and twelve years. It consists of nine tasks and assesses the four processes involved in reading: letter identification, word recognition, syntactic processing, and semantic processing. None of the children had any known cognitive impairments or visual or motor disorders; all were native Spanish speakers.

The study was approved by the Ethics Committee of the Psychology Department of the University of Oviedo. Before children participated in the experimental tasks, their families were given information about the purpose of the study and the type and duration of the tasks. Parents or guardians were then asked to provide written consent for their children's participation.

Materials

We constructed 16 pseudo-words, which varied systematically with respect to three variables: syllabic frequency, syllabic structure, and length. We created eight stimuli consisting of high-frequency syllables, four of which had a simple syllabic structure (i.e., short pseudo-words: *fibe* and *doño*; long pseudo-words: *nochola* and *farraña*), and four a complex syllabic structure (i.e., short: *siem* and *nues*; long: *tiempre* and *trabien*). We also created eight pseudo-words consisting of low-frequency syllables, half with a simple syllabic structure (i.e., short: *bifa* and *vomu;* long: *rullefo* and *bullipe*), and half with a complex syllabic structure (i.e., short: *blan* and *cons;* long: *crisdar* and *tridian*). This block of 16 pseudo-words was repeated six times.

Information about syllabic frequency was derived from several databases (Alameda & Cuetos, 1995; Davis & Perea, 2005; Duñabeitia, Cholin, Corral, Perea, & Carreiras, 2010), to be sure that all the selected syllables in the same frequency category had a similar token frequency (the number of times the child encounters the word when reading; see Table 1). Syllabic structure was classified as follows (C = consonant, V = vowel): structures CCV, CVVC, and CCVC were classified as complex, and structure CV as simple. Digraphs (two letters that make one sound) were classed as consonants. We thus had a set of eight long (seven-letter) pseudo-words and eight short (four-letter) pseudo-words.

Procedure

The experiment was carried out in a room free of noise and distractions. Children were assessed individually on the pseudo-word reading aloud task. Stimuli were presented on

Table 1. Mean Values of Token Syllabic Frequency for Pseudo-words

	Alameda & Cuetos, 1995	Duñabeitia *et al.*, 2010	Davis & Perea, 2005
High-frequency syllables	1882.68	1103.54	724.6
Low-frequency syllables	243.5	265.68	72.87

a laptop using the DMDX programme (Forster & Forster, 2003). The block of 16 stimuli was repeated six times. All children received the same instructions, which were displayed on the screen and reinforced orally. Children did not receive any feedback on their response, nor were they corrected if they misread the pseudo-words. In this way we were trying to simulate the natural conditions of individual reading, based on self-teaching.

Children's responses were recorded in WAV format using the DMDX program, and analysed with CheckVocal software (Protopapas, 2007) to calculate the number of correct responses and reaction times (RTs).

Results

Accuracy

There were 165 mistakes (mispronunciations) out of a total of 3,264 responses (16 pseudo-words × 6 blocks × 34 children), an error rate of 5.05%. This error rate decreased along the blocks (Block 1: 27.3%; Block 6: 11.5%). Although it was not a key to addressing our initial research questions, we nevertheless looked at the type errors made by children, and how these compared for our different types of experimental stimuli. We classified them as follows: letter inversion (i.e., changing the order of two or more letters of the word), letter substitution (i.e., changing one or more letters of the word for another), omission (i.e., missing one or more letters of the word when reading), addition (i.e., adding a letter to the word), repetition (i.e., repeating the first syllable before reading the whole word), and lexicalisation (i.e., reading a real word instead of the written pseudo-word; they are often orthographically similar). We used this classification of mistakes as it has been widely used in psycholinguistic studies for both oral and reading errors (Boder, 1970; Bušta, Hlavácková, Jakubícek, & Pala, 2009; Cano-Tobíasa, Granados-Ramosb, & Alcaraz-Romeroc, 2014; Fogarty, Dabbish, Steck, & Mostow, 2001; Khalid, Buari, & Chen, 2017; Payne, 1930; Williams & Reiter, 2004). The most common type of error was letter substitution, comprising 43.3% of the total errors, followed by letter inversions at 18.3%. The type of error varied according to the features of pseudo-words, as Table 2 indicates. The most common errors in simple high and low syllabic frequency were letter substitution (e.g., *mochola* instead of *nochola*) and letter inversion (e.g., *bife* instead of *fibe*). In complex high and low syllabic frequency, the usual errors were letter substitution (e.g., *tiembre* instead of *tiempre*) and addition (e.g., *trambien* instead of *trabien*).

Reaction times

The mean RTs for each kind of stimuli are shown in Table 3. This study used a repeated measures design: the same items were presented to multiple participants. With this design, mixed-effects modelling is required to estimate both fixed effects, i.e., replicable effects of theoretical interest: syllabic frequency (high–low frequency),

Table 2. Percentage of Each Kind of Mistake according to the Type of Pseudo-word

	High syllabic frequency	Low syllabic frequency	Simple syllabic structure	Complex syllabic structure
Letter inversion	22.86	0	31.65	6.90
Letter substitution	48.57	28.21	53.16	27.59
Omission	0	5.13	0	20.69
Addition	5.71	56.41	0	24.14
Repetition	14.29	5.13	13.92	13.79
Lexicalisation	8.57	5.13	1.27	6.90

syllabic structure (simple–complex), block (1 to 6), and length (short–long) and random effects, i.e., unexplained effects due to random variation between items or participants (Baayen, 2008; Baayen, Davidson, & Bates, 2008). We therefore used the lme4 (Bates, Mächler, Bolker, & Walker, 2015) and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2015) packages of R (R Core Team, 2017) in order to conduct these analyses. As the effects of the fixed effects do not have to be the same for all subjects and items, we used a random slope model, taking into account the different influence of each fixed effect studied on the subjects and items.

From the data, and using the linear mixed-effects model fit by REML, the final model (with the lowest AIC: Akaike's Information Criterion) was as follows: RTs ~ Block * Syllabic frequency * Syllabic structure * Length + (Block + Length + 1|Subject) + (1|Item). Using the maximal random effects approach we tested various models, trying to find which fixed effects would affect subjects and items, using the AIC score (i.e., a measure of model fit; Sakamoto, Ishiguro, & Kitagawa, 1986). The final random slope model (i.e., the most informative model) indicated that block and length had different effects on the subjects. We fitted the model using the REML estimation method, and performed null hypothesis significance testing of individual effects using the ANOVA model function call to obtain Type III sums of squares *F*-tests (see Table 4). There were significant main effects of Block (i.e., RTs were lower in the last block, $M = 866$, $SD = 27$, compared with the first one, $M = 977$, $SD = 23.3$), Syllabic Frequency (i.e., pseudo-words formed with high-frequency syllables had lower RTs than those with low-frequency syllables), and Length (i.e., short

Table 3. Mean RTs (SD) and Error Rate for Each Kind of Stimuli

	M (*SD*)	Error rate
Stimuli with high-frequency syllables	874 (21.9)	53.9 %
Stimuli with low-frequency syllables	895 (21.8)	46 %
Stimuli with simple structured syllables	895 (21.9)	59.4 %
Stimuli with complex structured syllables	874 (21.9)	40.6 %
Short stimuli	807 (20.7)	20 %
Long stimuli	961 (23.7)	80 %

https://doi.org/10.1017/S0305000917000514 Published online by Cambridge University Press

pseudo-words had lower RTs than long pseudo-words). The following significant interactions were also found: Block by Length, Syllabic Frequency by Length, and Syllabic Structure by Length, and Block by Syllabic Structure by Length.

We then carried out post-hoc analyses of the significant interactions in order to explain them. To perform the post-hoc analyses we used the lsmeans package (Lenth, 2016), and tested the length effect across blocks in both simple and complex structured pseudo-words using the Tukey adjustment. First, regarding to the Block * Length significant interaction, we found that the length effect (difference between long and short pseudo-words) decreased along the blocks (Block 1: $Estimate_{length\ effect} = 235$, $SD = 24$, $p < .001$; Block 6: $Estimate_{length\ effect} = 93$, $SD = 23.7$, $p = .056$). Second, our analysis of the Syllabic Frequency * Length significant interaction showed that the length effect, although significant in both types of pseudo-words, was lower in the pseudo-words formed with low-frequency syllables ($Estimate_{length\ effect} = 201$, $SD = 24.6$, $p < .001$) than in those with high-frequency syllables ($Estimate_{length\ effect} = 106$, $SD = 24.6$, $p = .007$). Third, the Syllabic Structure * Length interaction meant that the length effect, despite being significant in both types of stimuli, was lower in the pseudo-words with simple structured syllables ($Estimate_{length\ effect} = 118$, $SD = 24.6$, $p = .003$) than in those with complex syllables ($Estimate_{length\ effect} = 189$, $SD = 24.6$, $p < .001$). Finally, we carried out post-hoc analyses of the significant interaction Block * Syllabic structure * Length in order to discover when the length effect disappeared. We found that in the case of simple pseudo-words the length effect disappeared, whereas in the case of complex pseudo-words it remained present throughout the experiment (see Table 5 and Figure 1).

Table 4. Summary of Mixed-effects Models of RTs of Correct Responses

	F value	*p value*	*Cohen's d*
Block	10.61	< .001*	3.40
Syllabic frequency	1.63	= .037*	0.95
Syllabic structure	1.65	= .234	0.96
Length	69.72	< .001*	3.89
Block * Syllabic frequency	1.81	= .108	0.32
Block * Syllabic structure	0.57	= .727	0.12
Syllabic frequency * Syllabic structure	2.22	= .174	0.28
Block * Length	7.86	< .001*	2.23
Syllabic frequency * Length	8.55	= .019*	1.91
Syllabic structure * Length	4.79	= .059·	0.98
Block * Syllabic frequency * Syllabic structure	1.83	= .105	0.13
Block * Syllabic frequency * Length	1.42	= .213	0.34
Block * Syllabic structure * Length	5.32	= .043*	1.24
Syllabic frequency * Syllabic structure * Length	0.13	= .725	0.09
Block * Syllabic frequency * Syllabic structure * Length	1.12	= .347	0.15

Notes. * = significant effect or interaction; · = near-significant effect or interaction.

https://doi.org/10.1017/S0305000917000514 Published online by Cambridge University Press

Discussion

The aim of this experiment was to determine how frequency and syllabic complexity influenced the formation of orthographic representations of new words in third-grade children. To do this, we used a pseudo-word reading aloud task in which we manipulated the frequency and complexity of the syllables making up the pseudo-words, as well as their length. In order to evaluate the formation of orthographic representations we analysed reading times and pronunciation errors.

The results showed that children started to form orthographic representations as the length effect in pseudo-words (independent of the manipulated variables) decreased significantly along the blocks, as indicated by the significant interaction Block * Length. Moreover, syllabic structure influenced the formation of orthographic representations, as indicated by the significant interaction Syllabic structure * Block * Length. Post-hoc analyses demonstrated that the length effect disappeared sooner in the pseudo-words formed of structurally simple syllables, which means that the children started to use a lexical route for reading them. This is consistent with other studies showing that simple syllabic structure facilitates the recognition and reading of words (Jiménez *et al.*, 2010; Taft, 1979, 2001).

However, we did not find evidence that syllabic frequency influenced orthographic learning, as the length effect persisted throughout the six blocks in pseudo-words with both low-frequency and high-frequency syllables. Nevertheless, RTs were lower for pseudo-words with high-frequency syllables, as indicated by the significant effect of syllabic frequency. It seems that syllabic frequency only affected word recognition, which is in line with previous studies (Álvarez *et al.*, 2001; Carreiras *et al.*, 1993; Carreiras & Perea, 2014; Domínguez, Segui, & Vega, 2000; Perea & Carreiras, 1995, 1998). Therefore, it is possible that syllabic frequency affected the decoding (i.e., high-frequency syllables were decoded faster), which is the preceding step in the formation of orthographic representations.

Regarding the errors, we assessed the pseudo-words according to the most common type of error. On the one hand, in simple pseudo-words with both high and low syllabic frequencies, the main errors were letter inversion and letter substitution. On the other hand, in complex pseudo-words with high and low syllabic frequencies, the most typical errors were letter substitution and omission. It is possible to explain these errors according to the dual route model (Coltheart *et al.*, 2001). Letter substitutions could be due to a failure in the application of grapheme–phoneme conversion rules; that is,

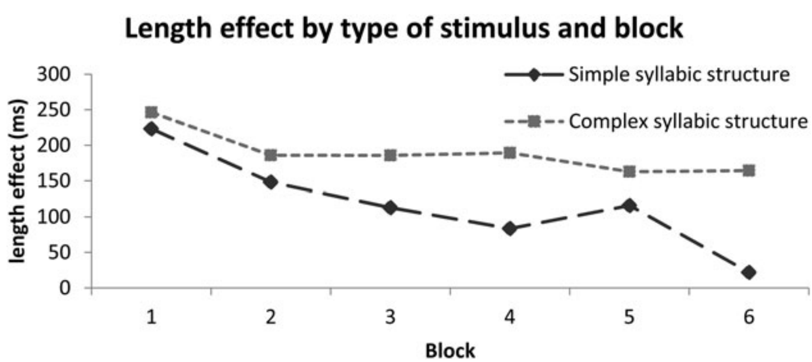

Figure 1. Length Effect in RTs of Pseudo-words by Syllabic Structure across Blocks

https://doi.org/10.1017/S0305000917000514 Published online by Cambridge University Press

Table 5. Post-hoc Analysis of the Block * Syllabic Structure * Length Significant Interaction

		Simple syllabic structure			Complex syllabic structure		
		estimate[1]	*SE*	*p value*	*estimate*[1]	*SE*	*p value*
Block 1	short-long	223.171	32.89	< .001*	246.41	32.93	< .001*
Block 2	short-long	148.282	32.62	= .003*	185.871	32.54	< .001*
Block 3	short-long	113.032	32.79	= .058·	185.732	32.36	< .001*
Block 4	short-long	83.65	32.53	= .336	189.411	32.39	< .001*
Block 5	short-long	116.007	32.58	= .045*	162.832	32.32	= .001*
Block 6	short-long	21.9	32.49	= .812	164.367	32.29	< .001*

Notes.[1] estimate mean difference between long and short pseudo-words in RTs; * = significant difference between short and long pseudo-words (i.e., significant length effect); · = near-significant length effect.

a failure in the decoding process. Letter inversions, as errors in the order of letter decoding, could be caused by a failure in visual analysis and/or in the order in which the stimuli's phonemes are stored. Finally, omissions might be produced by a fault in visual analysis, in which one or more letters are omitted.

Experiment 2

The aim of this part of the study was to explore the influence of context-dependent graphemes on the formation of orthographic representations in a language with transparent orthography, namely Spanish. We expected that the length effect would disappear sooner along the blocks in the pseudo-words formed by context-independent graphemes, since children formed orthographic representations of these pseudo-words more easily than those with context-dependent graphemes. Therefore, as in Experiment 1, the key finding we expected was a significant three-way interaction Block * Length * Context-dependency, indicating that the context-dependent variable modulates the decrease of the length effect along blocks; that is, the formation of orthographic representations.

Method

Participants

The study was conducted with 65 third-grade children (30 boys and 35 girls, age $M =$ 8.3 years, $SD = 0.28$) attending a state-assisted school in Oviedo. Selection of participants was based on tests of reading of words and non-words in PROLEC-R (Cuetos *et al.*, 2007). Only those children whose results were within the normal range for their age were included. None of the children had any known cognitive impairments or visual or motor disorders; all were native Spanish speakers.

The study was approved by the Ethics Committee of the Psychology Department of the University of Oviedo. Before children participated in the experimental tasks, their families were given information about the purpose of the study and the type and duration of the tasks. Parents or guardians were then asked to provide written consent for their children's participation.

Materials

Sixteen pseudo-words, half including context-dependent graphemes and half without, were created for a reading aloud task. Half of the stimuli in both conditions were short (two syllables, e.g., *guco*, *deto*) and half were long (three syllables, e.g., *gukato*, *dufemo*). Various syllables were used to create the pseudo-words, all of which had a paired syllable structure (CV). Stimuli obeyed Spanish spelling rules and could be read by following the rules for grapheme–phoneme conversion.

Procedure

Experiment 2 followed exactly the same procedure as Experiment 1.

Results

Accuracy

There were 731 errors (mispronunciations) out of a total of 6,240 responses (6 x 16 pseudo-words per 65 children), an error rate of 11.7%. The error rate declined across the blocks, with 20.7% of errors made in the first block and 15.7% in the last.

Just over half (53.4%) of the errors in reading stimuli with context-dependent graphemes were the result of consistent mispronunciation of a given stimulus. Response latencies were shorter for mispronunciations in the last block than in the first one.

As in Experiment 1, we analysed the type of error in each condition. Errors were divided into letter inversion, letter substitution, errors in orthographic rule application (i.e., mistakes made in reading context-dependent graphemes due to not correctly applying the grapheme–phoneme conversion rules), omission, addition, repetition, and lexicalisation. The most common errors in the pseudo-words with context-independent graphemes were letter substitution (e.g., *dufeno* instead of *dufemo*), followed by repetition (e.g., *mu-mu-mulepe*). In the stimulus formed from context-dependent graphemes, children made more errors of orthographic rule application (e.g., *zugue* instead of *zuge*), followed by letter substitution (e.g., *zapopa* instead of *gazopa*). Table 6 contains the percentage of the type of error for each kind of pseudo-word.

Reaction times

In Table 7 the mean RTs for each kind of stimuli are presented. This study used a repeated measures design, i.e., the same items were presented to multiple participants. As in the previous experiment, we estimated fixed effects: context-dependency (presence–absence), block (1 to 6), and length (short–long), and random effects (items and participants).

Using the linear mixed-effects model fit by REML (R Core Team, 2017), the final model was as follows: RTs ~ Context-dependency * Block * Length + (Context-dependency + Length + 1|subject) + (1|item). There were effects of Context-dependency (i.e., pseudo-words not containing context-dependent graphemes had lower RTs than those containing them, $M = 995$, $SD = 30.9$), Block (i.e., RTs in the last block were lower, $M = 938$, $SD = 24.5$, than in the first one, $M = 991$, $SD = 24.5$), and Length (i.e., short pseudo-words had lower RTs than long pseudo-words did). Also, some interactions were significant: Block by Length, and Context-dependency by Block by Length. The effects and interactions are presented in Table 8.

We then carried out post-hoc analyses of each significant interaction. First, and similar to Experiment 1, the Block * Length significant interaction meant that the length effect (difference between long and short pseudo-words) decreased significantly from first ($Estimate_{length\ effect} = 176$; $SD = 37.4$, $p = .007$) to last exposure ($Estimate_{length\ effect} = 51$; $SD = 37.3$, $p = .955$). We also conducted a post-hoc analysis

Table 6. Percentage of Type of Error by Context-dependency Variable

	Context-independent	Context-dependent
Letter inversion	1.8	0.7
Letter substitution	67.3	6.7
Errors in orthographic rule application	0	89.7
Omission	0	0
Addition	3.6	0.2
Repetition	21.8	1.7
Lexicalisation	5.5	1.2

Table 7. Mean RTs (SD) and Error Rate of Each Kind of Stimuli

	M (*SD*)	Error rate
Stimuli with context-dependent graphemes	995 (30.9)	90.7 %
Stimuli without context-dependent graphemes	903 (28.3)	9.3 %
Short stimuli	899 (28.7)	63.1 %
Long stimuli	999 (30.2)	36.9 %

on the interaction between Context-dependency (with and without context-dependent graphemes), Length, and Block to determine when the length effect disappeared. We found that, in the case of pseudo-words without context-dependent graphemes, the length effect disappeared in the last block, but persisted throughout the experiment for pseudo-words containing context-dependent graphemes (see Table 9 and Figure 2).

Discussion

The aim of this part of the study was to see whether the context-dependency of graphemes affected the formation of orthographic representations in third-grade children. We used a reading aloud task in which the stimuli were pseudo-words that contained, or did not contain, context-dependent graphemes; the stimuli were read six times. Learning was evaluated by analysing RTs and mispronunciations.

Our results again indicated that children formed orthographic representations, as indicated by the significant interaction Block * Length. Moreover, there was a difference in the formation of orthographic representations between stimuli with and without context-dependent graphemes, as suggested by the significant Context-dependency * Block * Length interaction. It seems that children started forming orthographic representations of pseudo-words with context-independent graphemes after six exposures: the length effect disappeared in the fifth block. However, they did not form orthographic representations of the pseudo-words containing context-dependent graphemes, as the length effect persisted throughout all six blocks, although reading times did decrease. This pattern of results is consistent with our hypothesis that children would have more difficulty with orthographic learning in

Table 8. Summary of Mixed-effects Models of RTs of Correct Responses

	F value	*p value*	*Cohen's d*
Block	12.11	< .001*	2.17
Context-dependency	6.67	= .021*	3.11
Length	8.15	= .013*	3.40
Block * Context-dependency	0.90	= .477	0.13
Block * Length	12.91	< .001*	2.76
Context-dependency * Length	0.04	= .837	0.53
Block * Context-dependency * Length	2.51	= .028*	1.26

Note. * = significant effect or interaction.

https://doi.org/10.1017/S0305000917000514 Published online by Cambridge University Press

Table 9. Length Effect in Each Condition and Block

		Context-independent			Context-dependent		
		estimate[1]	*SE*	*p value*	*estimate*[1]	*SE*	*p value*
Block 1	short-long	191.89	20.59	< .001*	156.41	23.51	< .001*
Block 2	short-long	128.93	20.73	< .001*	157.06	22.68	< .001*
Block 3	short-long	96.15	20.63	< .001*	118.47	22.76	< .001*
Block 4	short-long	77.35	20.68	= .01*	109.78	22.7	< .001*
Block 5	short-long	42.05	20.47	= .65	72.77	22.74	= .047*
Block 6	short-long	22.93	20.66	= .88	71.89	22.67	= .049*

Notes.[1] estimate mean difference between long and short pseudo-words in RTs; * = significant difference between short and long pseudo-words (i.e., significant length effect).

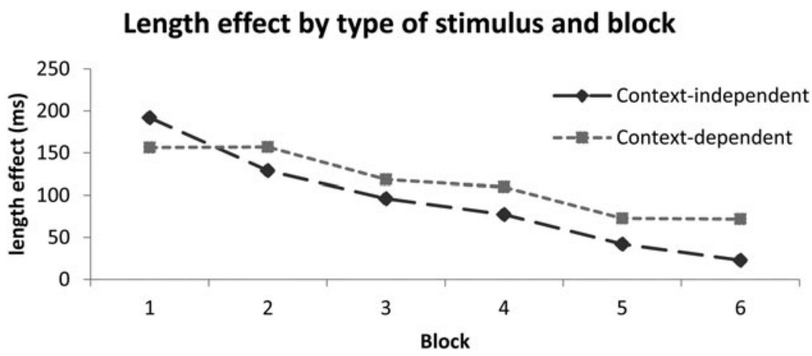

Figure 2. Length Effect in Each Condition and Block.

the case of words containing context-dependent graphemes. Our results are consistent with those found for reading in English (Wang *et al.*, 2012), despite the difference in language transparency.

Regarding the type of errors, a failure to use the grapheme–phoneme conversion rule was the most typical error for pseudo-words formed with context-dependent graphemes. The remaining errors were distributed similarly across both types of pseudo-words. The vast majority of mispronunciations occurred in reading pseudo-words containing context-dependent graphemes, and many pronunciation errors were repeated consistently at every exposure. Reading times for consistent mispronunciations decreased across blocks. This could be due to the children's failure to use phonological representation. This issue should be studied further.

General Discussion

The aim of our study was to determine how several sublexical variables influence the formation of orthographic representations in Spanish third-grade children. We carried out two experiments, manipulating variables such as syllabic frequency, syllabic structure, and the presence of context-dependent graphemes. Our results indicate that children started forming orthographic representations after six exposures, which corresponds with previous studies (Kwok & Ellis, 2014, 2015; Kwok *et al.*, 2017; Share, 1999; Suárez-Coalla *et al.*, 2016), although not all the variables studied facilitated this orthographic learning.

In the first experiment, children had to read pseudo-words that varied systematically with respect to syllabic frequency and syllabic complexity. We found that only syllabic complexity influenced the formation of orthographic representations; this was indicated by a reduction in the length effect along the blocks for pseudo-words with a simple structure, an effect that disappeared sooner than the length effect in complex-structured pseudo-words. Our results are in line with studies showing easier facilitation of reading simple words in opaque languages (Sprenger-Charolles & Siegel, 1997; Taft, 1979, 2001). In Spanish, a transparent language in which there is clear correspondence between grapheme and phoneme, it is more difficult to read and form orthographic representations of complex words than of simple words.

We did not find any evidence of the facilitation of orthographic learning for pseudo-words with high-frequency syllables, at least after six exposures. However,

https://doi.org/10.1017/S0305000917000514 Published online by Cambridge University Press

RTs were lower for pseudo-words with high-frequency syllables, which suggests that syllabic frequency may influence recognition and reading, as reported in previous studies (Álvarez *et al.*, 2001; Carreiras *et al.*, 1993; Perea & Carreiras, 1995, 1998). Therefore, there was no evidence that syllabic frequency influences the formation of orthographic representations, as the length effect did not disappear sooner in pseudo-words with high syllabic frequency than in those with low syllabic frequency. However, this variable affected reading times, since high-frequency syllables had smaller RTs. Perhaps syllabic frequency affected the decoding, which is the previous step in the formation of orthographic representations. In summary, our first experiment showed that syllabic structure, a variable that has hitherto been little studied, influences the formation of orthographic representations of new words.

The second experiment was carried out to investigate how the presence of context-dependent graphemes affected the formation of orthographic representations. The results indicate that the absence of context-dependent graphemes facilitated the formation of orthographic representations, since the length effect of this group of pseudo-words disappeared earlier than the length effect of the pseudo-words with context-dependent graphemes. We also found that reading latencies were shorter for stimuli without context-dependent graphemes. We conclude that an absence of context-dependent graphemes makes it easier to form an orthographic representation of the word, as there was a bigger reduction in reading latency over the course of repeated exposures for stimuli that did not contain context-dependent graphemes.

Our results are in line with studies of other transparent languages showing that reading latencies are shorter for pseudo-words based solely on context-independent graphemes. Some of those studies used a similar method to ours, measuring response latencies on a reading aloud task (Barca *et al.*, 2007); others used categorisation tasks (Peereman *et al.*, 2009) or lexical decision tasks (Ventura *et al.*, 2008). In all cases, words containing only context-independent graphemes were recognised and read faster.

We also found evidence that the context-independence of graphemes facilitates orthographic learning. This result is consistent with a study of learning English, which found that orthographic consistency facilitated the formation of orthographic representations in children (Wang *et al.*, 2012). English and Spanish have very different orthographic systems, since English is an opaque language with many irregularities, i.e., spellings are not reliably related to phonology, whereas Spanish is very transparent and has very regular grapheme–phoneme correspondences and only two graphemes whose pronunciation is context-dependent. The two studies used different methods: the English study (Wang *et al.*, 2012) used correct responses on an orthographic choice task as an indicator of the formation of an orthographic representation of the item concerned, whereas we used reduction in the length effect, which is assumed to indicate a shift from a sublexical to a lexical reading strategy. Despite these differences, the two studies produced similar results.

Although we did not conduct inferential statistics over the types of errors, we note that a considerable percentage of mistakes reflected errors that were repeated consistently at every exposure to a given context-dependent grapheme. This may be because, regardless of whether or not the reader has an orthographic representation of these pseudo-words, there is an error in the representation of the word in his or her phonological lexicon. If this is the case, then the presence of context-dependent graphemes in a word affects the formation of the phonological representation rather than the orthographic representation. It has previously been suggested that context-dependent phonological rules are more difficult to learn and automate

(Barca *et al.*, 2007; Rastle & Coltheart, 1998). These results have important educational implications, but we believe further investigation is needed to draw firm conclusions.

In summary, our study confirmed that third-grade children started forming orthographic representations of words after six exposures, which has already been demonstrated by other authors. There are several variables that are associated with more rapid formation of orthographic representations, namely a simple syllabic structure and the absence of context-dependent graphemes. The formation of orthographic representations has been investigated before, but those earlier studies did not typically consider the effects of sublexical variables. Finally, we found that a significant percentage of errors were consistent; this is a finding worth exploring in future research as it may reflect a problem with phonological representations. Our findings provide new evidence about the way in which children form orthographic representations, and allows us to identify some important points to take into account when considering how children learn to read. On the one hand, automating the grapheme–phoneme conversion rules is important, as this will allow children to form orthographic representations easily, which implies faster and more accurate reading. The results of our studies suggested that children have more difficulty in forming orthographic representations of words with context-dependent graphemes and structurally complex syllables. Therefore, it is possible that emphasising the decoding of these types of units in order to automate the process would reduce difficulties in the formation of orthographic representations of words containing these units. On the other hand, it seems to be important to prevent children making mistakes when they read, thus facilitating the formation of correct orthographic representations; children's mistakes while reading will probably persist in following readings, as shown in the error analysis of the second experiment. It is possible, but needs further study, that a persistent error in reading a word makes children form a phonological representation that does not correspond with the written word. This incorrect phonological representation would result in future incorrect readings that lead to an incorrect orthographic representation.

Supplementary materials. For supplementary material for this paper, please visit https://doi.org/10.1017/S0305000917000514.

Acknowledgements. This study was funded by Grant PSI2012-31913 and PSI2015-64174-P from the Spanish government and supported by a predoctoral grant (BP14-038) from the Foundation for the Promotion of Applied Scientific Research and Technology in Asturias (FICYT). We want to thank the reviewers and editors for providing helpful comments and suggestions that improved and clarified this manuscript.

References

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.

Álvarez, C., Carreiras, M., & Perea, M. (2004). Are syllables phonological units in visual word recognition? *Language and Cognitive Processes*, *19*(3), 427–52.

Álvarez, C., Carreiras, M., & Taft, M., (2001). Syllables and morphemes: contrasting frequency effects in Spanish. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*(2), 545–55.

Álvarez, C. J., García-Saavedra, G., Luque, J. L., & Taft, M. (2017). Syllabic parsing in children: a developmental study using visual word-spotting in Spanish. *Journal of Child Language*, *44*(2), 380–401.

Baayen, R. H. (2008) *Analyzing linguistic data: a practical introduction to statistics*. Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Barca, L., Ellis, A. W., & Burani, C. (2007). Context-sensitive rules and word naming in Italian children. *Reading and Writing*, *20*(5), 495–509.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Boder, E. (1970). Developmental dyslexia: a new diagnostic approach based on the identification of three subtypes. *Journal of School Health*, *40*, 289–90.

Bowey, J. A., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: a critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *92* (3), 203–19.

Bušta, J., Hlaváckoviá, D., Jakubícek, M., & Pala, K. (2009). Classification of errors in text. In P. Sojka & A. Horák (Eds.), *Recent advances in Slavonic natural language processing RASLAN*. (pp. 109–20). Brno: Masaryk University.

Cano-Tobíasa, G., Granados-Ramosb, D. E., & Alcaraz-Romeroc, V. M. (2014). Eye movement recordings during reading tasks in children with mixed dyslexia. *International Journal of Arts and Commerce*, *3*(5), 63–74.

Carreiras, M., Alvarez, C., & De Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, *32*, 766–80.

Carreiras, M., & Perea, M., (2004). Effects of syllable neighbourhood frequency in visual word recognition and reading: cross-task comparisons. In L. Ferrand & J. Grainger (Eds.), *Psycholinguistique Cognitive*. (pp. 233–51). Bruxelles : De Broeck Université.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–56.

Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). *Prolec-R: batería de evaluación de los procesos lectores, Revisada*. Madrid: TEA.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: examining the role of self-teaching. *Journal of Experimental Child Psychology*, *82*(3), 185–99.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: a quick measure of reading experience. *Journal of Educational Psychology*, *82*(4), 733–40.

Davis, C. J., & Perea, M. (2005). BuscaPalabras: a program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, *37*, 665–71.

Domínguez, A., Segui, J., & Vega, F. C. (2000). Syllable-frequency effect in visual word recognition: a review with particular reference to Spanish data. *Psicológica: Revista de metodología y psicología experimental*, *21*(3), 375–402.

Duñabeitia, J. A., Cholin, J., Corral, J., Perea, M., & Carreiras, M. (2010). SYLLABARIUM: an online application for deriving complete statistics for Basque and Spanish syllables. *Behavior Research Methods*, *42*, 118–25.

Ehri, L., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing: An Interdisciplinary Journal*, *7*, 295–326.

Fogarty, J., Dabbish, L., Steck, D., & Mostow, J. (2001). Mining a database of reading mistakes: For what should an automated reading tutor listen? In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Artificial intelligence in education: AI-ED in the wired and wireless future* (pp. 422–33). Amsterdam: IOS Press.

Forster, K. I., & Forster, J. C. (2003). DMDX: a Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–24.

Jiménez, J. E., García, E., O'Shanahan, I., & Rojas, E. (2010). Do Spanish children use the syllable in visual word recognition in learning to read? *Spanish Journal of Psychology*, *13*(1), 6374.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*(4), 329–54.

Khalid, N. M., Buari, N. H., & Chen, A.-H. (2017). Comparison of oral reading errors between contextual sentences and random words among schoolchildren. *International Education Studies*, *10*(1), 47–55.

Kwok, R. K. W., Cuetos, F., Avdyli, R., & Ellis, A. W. (2017). Reading and lexicalization in opaque and transparent orthographies: word naming and word learning in English and Spanish. *Quarterly Journal of Experimental Psychology*, *70*(10), 2105–29.

Kwok, R. K. W., & Ellis, A. W. (2014). Visual word learning in adults with dyslexia. *Frontiers in human neuroscience*, *8*, 264.

Kwok, R. K. W., & Ellis, A. W. (2015). Visual word learning in skilled readers of English. *Quarterly Journal of Experimental Psychology*, *68*(2), 326–49.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmerTest: tests in linear mixed effects models. R package version 2.0-33. Retrieved from: <https://CRAN.R-project.org/package=lmerTest>.

Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, *93*(2), 166–85.

Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33.

Maloney, E., Risko, E. F., O'Malley, S., & Besner, D. (2009). Tracking the transition from sublexical to lexical processing: on the creation of orthographic and phonological lexical representations. *Quarterly Journal of Experimental Psychology*, *62*(5), 858–67.

Payne, C. S. (1930). The classification of errors in oral reading. *Elementary School Journal*, *31*(2), 142–6.

Peereman, R., Dufour, S., & Burt, J. (2009). Orthographic influences in spoken-word recognition: the consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin & Review*, *16*, 363–8.

Perea, M., & Carreiras, M. (1995). Efectos de frecuencia silábica en tareas de identificación. *Psicológica*, *16*, 483–96.

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *24* (1), 134–44.

Protopapas, A. (2007). CheckVocal: a program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–62.

Qiao, X., Forster, K., & Witzel, N. (2009). Is Banara really a word? *Cognition*, *113*(2), 254–7.

R Core Team (2017). R: a language and environment for statistical computing. Retrieved from R Foundation for Statistical Computing, Vienna <https://www.R-project.org/>.

Rastle, K., & Coltheart, M. (1998). Whammies and double whammies: the length effect on nonword reading. *Psychonomic Bulletin & Review* *5*, 277–82.

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike information criterion statistics*. Dordrecht: D. Reidel.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–74.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218.

Share, D. L. (1999). Phonological recoding and orthographic learning: a direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 95–129.

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*, 267–98.

Sprenger-Charolles, L., & Siegel, L. S. (1997). A longitudinal study of the effects of syllabic structure on the development of reading and spelling skills in French. *Applied Psycholinguistics*, *18*(4), 485–505.

Suárez-Coalla, P., Álvarez-Cañizo, M., & Cuetos, F. (2016). Orthographic learning in Spanish children. *Journal of Research in Reading*, *39*, 292–311.

Suárez-Coalla, P., Avdyli, R., & Cuetos, F. (2014a). Influence of context-sensitive rules in the formation of orthographic representations in Spanish dyslexic children. *Frontiers in Psychology*, *5*:1409. doi:10.3389/fpsyg.2014.01409.

Suárez-Coalla, P., Ramos, S., Álvarez-Cañizo, M., & Cuetos, F. (2014b). Orthographic learning in dyslexic Spanish children. *Annals of Dyslexia*, *64*(2), 166–81.

Taft, M. (1979). Lexical access via an orthographic code: the Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, *18*, 21–39.

Taft, M. (2001). Processing of orthographic structure by adults of different reading ability. *Language and Speech*, *44*, 351–76.

Ventura, P., Kolinsky, R., Pattamadilok, C., & Morais, J. (2008). The developmental turn point of orthographic consistency effects in speech recognition. *Journal of Experimental Child Psychology*, *100*, 135–45.

Wang, H.-C., Castles, A., & Nickels, L. (2012). Word regularity affects orthographic learning. *Quarterly Journal of Experimental Psychology*, *65*, 856–64.

Wang, H.-C., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading*, *17*(5), 369–84.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology*, *50A*(2), 439–56.

Williams, S., & Reiter, E. (2004). Reading errors made by skilled and unskilled readers: evaluating a system that generates reports for people with poor literacy. In: *Fourteenth Annual Meeting of the Society for Text and Discourse* (p. 21), 1–4 August, 2004, Chicago, IL.

Cite this article: Álvarez-Cañizo M, Suárez-Coalla P, Cuetos F (2018). The role of sublexical variables in reading fluency development among Spanish children. *Journal of Child Language* **45**, 858–877. https://doi.org/10.1017/S0305000917000514