Lexical and Sublexical Skills in Children’s Literacy

Joana Acha¹,², Gorka Ibaibarriaga¹, Nuria Rodríguez¹, and Manuel Perea³,⁴

Abstract
Letter knowledge and word identification are key skills for reading and spelling. Letter knowledge facilitates the application of sublexical letter-sound mappings to decode words. With reading experience, word identification becomes a key lexical skill to support decoding. In transparent orthographies, however, letter knowledge might be an enduring predictor of decoding and spelling, even in children with some reading experience. This study investigated the association of children’s sublexical (letter knowledge) and lexical skills (word identification and vocabulary) with word decoding and spelling accuracy in Spanish, which is a transparent orthography. The sample consisted of 117 Spanish-speaking children, aged 8 to 10. Results revealed that (1) letter knowledge and word identification were independently associated with children’s word spelling; (2) word identification was uniquely associated with word decoding; and (3) children’s vocabulary level was associated with word identification. The implications of these findings were examined within the framework of reading models and the characteristics of a transparent orthography.

Keywords
decoding, spelling, children, literacy processes

Introduction
Achieving accurate reading and writing requires putting multiple cognitive skills into practice. In beginning readers, letter knowledge (i.e., the knowledge of each letter’s name or sound) has been shown to be a robust and early predictor of children’s reading and spelling accuracy as it facilitates learning letter-sound mappings

¹Universidad del País Vasco UPV/EHU, San Sebastian, Spain
²Biodonostia, Health Research Institute, San Sebastian, Spain
³Universitat de València, València, Spain
⁴Universidad Antonio de Nebrija, Madrid, Spain

Corresponding Author:
Joana Acha, Department of Basic Cognitive Processes and its Development, Universidad del País Vasco UPV/EHU, Tolosa Hiribidea, 20018, Donostia, Spain.
Email: joana.acha@ehu.eus
(e.g., Bara et al., 2016; Foulin, 2005). After some reading experience, children acquire knowledge about the orthographic characteristics of words. Consequently, lexical whole-word identification skills replace letter-sound mappings as an effective strategy for decoding and spelling, particularly in opaque orthographies with inconsistent mappings between letters and sounds (Conrad et al., 2013; Rothe et al., 2015).

In transparent orthographies, in which mappings between letters and sounds are consistent and univocal, the interplay between sublexical and lexical skills in childhood might differ. On the one hand, letter knowledge is strongly related to accurate word decoding and spelling during the first years of primary school (Caravolas et al., 2012; De la Calle et al., 2018; Kim & Pallante, 2012; Ziegler et al., 2010). Hence, letter knowledge could be a long-lasting predictor of word decoding and spelling in these orthographies (Ardila & Cuetos, 2016), even when word identification skills are in place. On the other hand, the extreme transparency in letter-sound mappings in Spanish facilitates the construction of word representations (Suárez-Coalla et al., 2016) and, therefore, the early use of word identification skills. To date, the specific association of letter knowledge and word identification skills to decoding and spelling in a transparent orthography like Spanish has not been addressed. Understanding these associations has important implications for theoretical and educational levels because there is a current debate about whether reading instruction in primary school should focus mainly on letter-sound mappings or word identification skills (Suárez-Coalla et al., 2022; Villar & Vieiro, 2015).

First, in this introduction, we will review the role of letter knowledge in reading, in particular with developing readers. Second, we focus on the transition from letter to word knowledge and emphasize the importance of word identification skills in reading. Third, we explore the potential distinct roles of sublexical processes at the letter level and lexical processes at the word level in the context of transparent orthographies. Finally, we outline the goals of the current study—which was conducted in a transparent orthography (namely, Spanish)—and the predictions, which are framed within the context of leading models of reading.

**Letter Knowledge and Its Role in Word Decoding and Spelling**

Letter knowledge is one of the strongest predictors of word decoding and spelling from an early age. Previous studies have repeatedly shown that knowledge of the names or sounds of letters in the alphabet helps children isolate specific sounds in the language, understand the relation with the written units that represent them (e.g., see Cardoso-Martins et al., 2011; Share, 2004), and learn letter-sound associations (Treiman & Kessler, 2003). Because letter knowledge supports children’s grapheme–phoneme mapping ability, it becomes the driving force of decoding and spelling accuracy during the early years of reading instruction (Piasta & Wagner, 2010; Thompson et al., 2015). However, the period in which this variable maintains its predictive strength varies across orthographies.

Longitudinal studies conducted in opaque orthographies (i.e., orthographies in which letters and sounds do not have univocal associations) have shown that the
predictive role of letter knowledge stands up to grade two (Catts et al., 2015; Schatschneider et al., 2004; Sunde et al., 2020), although the effects are stronger at early ages (Caravolas et al., 2013; Catts et al., 2015; Muter et al., 2004). Importantly, this developmental period accounts for children’s readiness to improve accuracy in reading words in opaque orthographies (Caravolas et al., 2013). Notably, in transparent orthographies (i.e., orthographies with a clear one-to-one association between letters and sounds), letter knowledge seems to be strongly associated with children’s decoding and spelling abilities for a more extended period (see Caravolas et al., 2013; Kim & Pallante, 2012; Ritchey & Speece, 2006). For example, in a longitudinal study in Finnish, which has a very transparent orthography, Leppänen et al. (2008) and Torppa et al. (2016) reported that knowledge of the letter name at age five predicted reading accuracy among the same children until fourth grade (age 9). Similarly, Lervåg et al. (2009) found that Norwegian children’s letter name knowledge at age 6 predicted decoding abilities until age 8. It is important to note that in the mentioned Finnish studies, letter names were equivalent to letter sounds, which also occurs in Spanish with few exceptions (“f”, “l”, “m”, “n”, and “s”). In Spanish, Martínez and Goikoetxea (2020) found that for children between 6 and 7 years of age, knowledge of letter names at the beginning of the school year was the most robust predictor of word reading and spelling accuracy 1 year later when children were in second and third grade, respectively.

The same relation seems to apply to phoneme–grapheme mappings in spelling. There is strong evidence for a significant relationship between letter knowledge and spelling in children between 4 and 5 years old (Paige et al., 2018; Puranik et al., 2011; Sunde et al., 2020; Wang et al., 2014), and recent studies suggest that the relation between both skills in Spanish extends beyond 7-year-olds (Martínez & Goikoetxea, 2020). However, this relation has not been fully explored in this orthography. Several reasons explain the long-lasting relationship between children’s letter knowledge and their decoding and spelling skills in transparent orthographies. First, letter knowledge can make the child aware of phonemes as critical linguistic units facilitating early mastery of the sublexical grapheme–phoneme mapping strategy sufficient to read and spell any word (Suárez-Coalla et al., 2014). Second, letter knowledge can facilitate the encoding of letters within words and the storage of accurate word representations necessary for accurate reading and spelling (Share, 2004; Suárez-Coalla et al., 2016).

Word Identification and Development of Lexical Strategies for Reading and Spelling

Another key aspect of reading development is that as the reading experience increases, children store word representations and can retrieve them automatically from memory (Mol & Bus, 2011). Therefore children experience a transition from sublexical reading and spelling strategies supported by letter knowledge to lexical strategies supported by word identification skills. One explanation for this transition comes from the self-teaching hypothesis (Share, 2008), which posits that once children master letter-sound mappings, they only need to read a word a few times to store its orthographic form.
This mastery results in faster and more accurate decoding and spelling of words that have been encountered several times (Martens & de Jong, 2006; Tucker et al., 2016). Interestingly, a number of studies have found that once a word’s orthographic form is attained—between grades two and three—lexical strategies become more important than sublexical strategies in supporting decoding and spelling (Lété et al., 2008; Mesquita et al., 2020).

The emergence of lexicality effects in reading and lexical decision tasks (identifying known words faster and more accurately than unknown pseudowords) has indicated that the child is employing lexical word identification strategies for reading (e.g., see van Viersen et al., 2022; Zoccolotti et al., 2009). For instance, training studies have shown that children exposed to a set of pseudowords for decoding can identify these trained pseudowords more accurately than nontrained pseudowords in a post-test lexical decision task (Bosse et al., 2015; Ricketts et al., 2011). Interestingly, recent empirical evidence has sustained the finding that the word learning process is easier in transparent orthographies as opposed to opaque orthographies: children need very few exposures to correctly identify trained unknown words compared to nontrained, unknown words (see Suárez-Coalla et al., 2016). These findings suggested that after some reading experience, children’s word identification skills might support word decoding and spelling. This might occur because the whole word representation is retrieved from memory, even given the different nature of decoding (translating a visual array of letters into sounds) and spelling (translating a spoken word into its corresponding letters within the visual array).

An important issue is that when lexical strategies are put in place, word identification skill seems to be aided by the child’s knowledge of the word’s meaning both in opaque (Ouellette & Beers, 2010; Ricketts et al., 2007) and in transparent orthographies (Álvarez-Cañizo et al., 2019). For instance, Ouellette (2006) found that while expressive vocabulary predicted English word identification in fourth-grade children (see also Chiappe et al., 2004; Lindsey et al., 2003), receptive vocabulary, which taps into children’s access to meaning, predicted decoding (see also Carlson et al., 2013). Some evidence has shown that receptive vocabulary in fourth-grade children shares variance with English word decoding and that both skills are related to word identification (Metsala, 1997; Tunmer & Chapman, 2012). Receptive vocabulary has also been related to spelling accuracy in primary school children (Mackenzie & Hemmings, 2014; Sumner et al., 2016). These findings have suggested that at least in opaque orthographies, receptive vocabulary can facilitate word decoding and spelling. The logic is that children’s knowledge of the word’s meaning could facilitate their access to the correct word representation.

**From Letters to Words in Transparent Orthographies**

A crucial gap in previous research is that despite the importance of letter knowledge, word identification, and vocabulary skills for decoding and spelling success, the evidence about their specific role in children learning to read in a transparent orthography has been scarce and inconclusive. Because letter knowledge in transparent orthographies is a strong longitudinal predictor of reading accuracy until grade four
(Leppänen et al., 2008), applying letter-sound mappings becomes a reliable strategy for children to read and spell accurately almost any word (Caravolas et al., 2012); therefore, children in transparent orthographies might rely to a great extent on their letter knowledge. Indeed, word identification skills have seemed less supportive in transparent orthographies (Defior et al., 2009; Zoccolotti et al., 2009), even when word representations are easily formed (Suárez-Coalla et al., 2016). Several studies have found that children between the second and fifth grades who showed decoding difficulties in opaque orthographies improved their decoding ability after an intervention based on word identification compared to children who did not receive this intervention (Cunningham, 2006; van Gorp et al., 2017). However, this result has not been replicated with third- and fifth-grade children in a transparent orthography, such as Norwegian, suggesting that these children may rely more on letter-by-letter decoding than word identification strategies (Brinchmann et al., 2016). This empirical discrepancy has raised the question of whether, in transparent orthographies, even after some years of reading experience, letter knowledge still plays a role either by facilitating word decoding and spelling directly or by its association with children’s word identification skills.

This question has important implications for models of visual-word recognition. In the leading models of word recognition and reading, letter encoding is a sine qua non condition of learning words. However, the models have made different claims about how letter and word information is processed to support reading. For instance, some models assumed that there is one single procedure for reading (e.g., the triangle model of Seidenberg & McClelland, 1989), which entailed the simultaneous activation of a word’s orthography, phonology, and semantics. Particularly in transparent orthographies (i.e., orthographies with a one-to-one correspondence between phonology and orthography), the activation of a phonological representation at the sound-word level can strongly activate its orthographic representation at the letter-word level in the model. Critically, the association between different representations is reciprocal (i.e., there is activation from letters and their sounds to words and their pronunciations and meanings and vice versa). Hence, single route models like the triangle model (Seidenberg & McClelland, 1989) assumed an association between sublexical and lexical processes because all representations (i.e., letter, word, and semantic) were activated when a word is encountered for decoding or spelling.

Some other leading models, however, assumed that there are two distinct procedures for word reading (e.g., dual-route cascaded model [Coltheart et al., 2001]; connectionist dual process model [Zorzi, 2010]). One procedure consisted of mapping letters to their sounds. Another procedure consisted of activating the orthographic information of the whole word stored in memory and retrieving its phonology from there. This last procedure was the usual one for words that have been read several times and which can be identified automatically, aided by the reader’s knowledge of a word’s meaning. Both procedures shared an initial (letter identification) and final stage (whole word identification). The difference was that one procedure was serial and slow, although effective for reading unknown words, and the other was automatic and effective for quickly identifying known words and their meanings.

Similarly, the multiple route model (Grainger et al., 2012) accommodated two distinct procedures for reading. In the initial stages of learning to read, letters were used as
basic, elementary processing units, resulting in a “fine-grained” processing style. With reading experience, complete words were retained, leading to a “coarse-grained” processing style, which helped identify the stimulus as a whole word. Thus, according to the family of dual route models, lexical, and sublexical procedures for reading were distinct and independent, and lexical word-level processes were expected to be put in place when readers had some reading experience. However, as in transparent orthographies, letter-sound mappings were enough for reading words accurately, so readers could only rely on this strategy for successful decoding even after some reading experience. According to dual-route models, a weak association between letter knowledge and word identification and semantics was expected because letters either activate the whole word’s phonological form (lexical procedure) or activate each letter’s sound in a serial manner (sublexical procedure).

Although single-route and dual-route models have made significant contributions to the understanding of the mechanisms involved in reading, there is still a debate about which type of models best represents the characteristics of readers in transparent orthographies. (Note that these models were initially implemented in English, which has a deep orthography.) Thus, it could be that in children with some reading experience, a single procedure was more efficient due to the direct mapping between orthography and phonology at the letter and word levels. Alternatively, it could also have been that even in children with some reading experience, a system with two procedures was more efficient, depending on the child’s word knowledge (only when the child’s word identification skill is good will the lexical procedure overcome the sublexical one) or the task demand (spelling or decoding).

The Present Study

The present study aimed to expand the scope of previous studies by examining the association of sublexical and lexical skills with word decoding and spelling in a transparent orthography (Spanish) with a sample of Spanish speakers (emergent bilinguals) in third grade. Note, that in this grade, children are expected to have good word knowledge due to reading experience, with letter knowledge still expected to be predictive of decoding and spelling.

We had two main goals. First, we wanted to identify the specific role of children’s sublexical (letter-sound knowledge) and lexical skills (word identification and vocabulary) on decoding and spelling accuracy in a transparent orthography. Second, we wanted to test the relation between letter knowledge and word identification skills in their contribution to reading and spelling under the predictions of the two main families of models of reading. If, due to the transparency of Spanish, children relied on letter knowledge to decode and spell words, and such knowledge facilitated the storage of word forms, both sublexical and lexical skills should play a role in reading and spelling accuracy. In this scenario, letter knowledge would be strongly associated with children’s word identification skills, as predicted by models that assume a single procedure for reading. Alternatively, if children’s word identification skills are primarily based on their reading experience and stored knowledge of word forms, their decoding and spelling abilities may be strongly associated with their word identification skills. As a result,
there would be little correlation between letter knowledge and word identification skill, as postulated by reading models involving two distinct procedures.

**Method**

**Participants**

A total of 117 primary school children between 8- and 10-years old (\(M\) age = 9.2, \(SD = 0.45\), range = 7.4–8.9; 62 females) participated in this experiment with the informed consent of their parents. The data were part of a larger longitudinal study focused on exploring the development of orthographic knowledge (learn REAĐy founded by the Spanish Ministry of Science and Innovation, 2020), and were collected from a school located in the urban area of Bilbao (Basque Country, Spain). The children who participated in the study met the following inclusion criteria: (a) they were enrolled in third grade; (b) they had no symptoms of neuropsychiatric (ADHD, autism spectrum disorders) and sensory problems; (c) they had no history of special education services or reading (or language) therapy; and (d) they had normal or corrected vision. No grade repetition was reported. The method of reading instruction used in the school was the phonic or synthetic method, which implied that all children were explicitly instructed in the alphabetic code from the first year of primary school. All participants were Spanish speakers (emergent Spanish-Basque bilinguals), with Spanish being the L1. They were all enrolled in the Spanish model; in these schools, Spanish was the language of schooling, and Basque was a subject of the curriculum. Basque is also a transparent orthography in which most grapheme–phoneme relationships are shared with Spanish. The home language exposure level was measured with a questionnaire completed by the families. All children showed a similar level of exposure to Spanish outside of school with preferential use of that language in a percentage of 80% (\(M = 79.8\), \(SD = 12.8\)) compared to 20% of Basque (\(M = 20.2\), \(SD = 11.3\)).

**Materials and Design**

**Control Task**  

*Nonverbal IQ*. The matrices task of the Kaufman Brief Intelligence Test (K-BIT, Kaufman & Kaufman, 1990) was used to control the nonverbal reasoning score across participants. According to the test manual, internal consistency estimates for the subtest ranged from .74 to .93. The task required the child to point to the missing figure from a logical sequence or set. The trials were grouped into eight sets of five items each. Testing was discontinued when the child responded incorrectly to all items in one set. Raw measures showed that all children were within the normal range (\(M = 107.27\); \(SD = 10.73\), range = 89–138).

**Experimental Tasks**  

*Letter Knowledge*. Letter knowledge was assessed using a computerized letter naming task with the DMDX software program (Forster & Forster, 2003). Twenty letters of the alphabet (t, u, b, f, n, v, c, r, x, z, j, s, q, ñ, y, p, d, l, g, and m) were
presented one by one in the center of the screen, with a display time of 100 milliseconds (ms). Each trial began with a central fixation point presented for 400 ms, followed by a letter that remained on the screen for 2000 ms. Once the letter disappeared from the screen, the screen remained blank for 150 ms before the appearance of the next fixation point. Participants were asked to produce the sound of each letter immediately after its presentation on the screen. The percentage of hits was analyzed.

**Receptive Vocabulary.** Children’s semantic knowledge was tested using the Peabody Picture Vocabulary Test-III (Dunn, 2006). It contained 175 cards with four pictures on each, and the child had to indicate which image represented the meaning of the word produced by the examiner. One score was given for each correctly identified picture. Raw scores were employed for experimental purposes. External validity with the Stanford-Binet vocabulary test ranged from 0.68 to 0.76, and test reliability was 0.80.

**Word Identification.** This skill was measured with a computerized lexical decision task in which word and pseudoword stimuli were presented, one by one, on the computer screen with the DMDX software in 14-point Courier New font, uppercase, in black on a white screen. The lexical decision task (see Meyer & Schvaneveldt [1971] and Rubenstein et al. [1970], for early research using this task) is one of the most employed tasks in word recognition research since it provides an online measure of the reader’s ability to identify words (Grainger & Jacobs, 1996). The task was to decide whether what appeared on a screen was a real word or not. Words and pseudowords were included and presented randomly in the task to test the subject’s ability to access a word from the mental lexicon automatically (hence, it was a time-limited task) and accurately (by identifying real words among some orthographically legal pseudowords that are similar to real words). The word stimuli consisted of 120 decodable words of high length ($M = 6.5$ letters; range 6–8 letters) and medium frequency and neighborhood (Frequency per million = 59; $N = 1.2$) according to the ONESC children frequency database (Martín & Pérez, 2008). Each trial consisted of a central presentation of a series of hashes (######) for 500 ms, followed by the target item being presented in all caps for 5000 ms or until a response was given. A set of 120 decodable pseudowords was created by changing two consonants from the word targets (see Appendix in the online, supplementary archive [provide URL by Sage]). The order of the items was random for each participant. We employed a go/no-go procedure of pressing the yes key when the stimulus was a word and refrain from responding when the stimulus was not a word. We chose the go/no-go procedure because it was easier than the two-choice procedure of pressing the yes key when the stimulus was a word and the no key when the stimulus was not a word. Our procedure made the task more manageable in developing readers (see Perea et al., 2013). Children were instructed to press the green-painted L key on the keyboard to indicate if what they saw was a word they knew and refrain from pressing it if the word was unknown to them. They were instructed that they should be quick to respond while maintaining the highest possible level of accuracy. The percentage of correct answers was recorded. The split-half reliability coefficient of the lexical decision task was 0.65.
**Word Decoding and Spelling.** For the decoding and spelling test, 180 medium-length words ($M = 4.8$ letters; range 4–6 letters; frequency per million = 257; $N = 12$) were selected from the ONESC children frequency database (Martín & Pérez, 2008). These words were different from the ones employed in the word identification task. To control for repetition effects, children only decoded or spelled half of the items. For this purpose, two counterbalanced sets were created. In each set, the 90 spelling items from list 1 were used as decoding items in list 2, and the 90 decoding items from list 1 were used as spelling items in list 2. The sets were balanced in length, frequency, and $N$ (number of orthographic neighbors; e.g., house and mouse, differing in just one letter, would be orthographic neighbors). In this way, decoding and spelling measures were obtained for all items avoiding facilitative effects from item repetition.

**Procedure.** The battery of tests was administered in three sessions. The first session lasted approximately half an hour for the spelling and decoding test. The experimenter dictated each word aloud and asked the children to write the words on a blank sheet of paper given to them for this purpose. In the second 1h session, children were assessed individually in letter knowledge, word identification, and nonverbal IQ. The receptive vocabulary test was also evaluated individually in the third half-hour session.

**Analytical Approach.** As a first step, a zero-order correlation analysis was performed to ensure that multicollinearity was not a problem in the study (Tabachnick & Fidell, 1996). Afterward, structural equation modeling (SEM, Satorra & Bentler, 2001) was employed to investigate the relationship between sublexical (letter knowledge) and lexical (word identification and vocabulary) variables with word decoding and spelling. This was done using the AMOS 28 statistical package. Due to the relatively large sample size, children from several classes took part in the study. Therefore, as a second step, multilevel methods were applied following the procedures described by Hox et al. (2017) so that each class outcome (level 1 of analysis) was nested into a group latent variable (level 2 of analysis) for each measure. However, if between-class differences were not found at level 1, within-group element analysis could be conducted at level 2 (group level) to simplify the model and to avoid creating unnecessary latent variables. Consequently, before testing the final SEM model, we analyzed whether significant variance existed at the class level for each variable. As this was not the case, no structural relations were hypothesized at level 1 (classes), and SEM analyses were finally performed at level 2 (i.e., whole group measures for each observed variable). One model was tested for decoding and spelling outcomes, respectively. To fit the specific predictions of reading models, independent paths from letter knowledge, word identification, and vocabulary were associated with the outcome in each model. A path was also added from letter knowledge to word identification and from vocabulary to word identification.

Model fit was evaluated using maximum likelihood estimation parameters. As a general rule, model fit was considered good if the chi-square was not significant, the root mean square error of approximation (RMSEA) was below .080, and the Tucker-Lewis index (TLI) and the comparative fit index (CFI) were above .90 (Kline, 2005).
Results

Descriptive Statistics and Correlational Analyses

The descriptive statistics for all the variables are shown in Table 1. Data were filtered to eliminate univariate extreme values, defined as those cases that presented values greater than two standard deviations above or below the sample mean. All children’s data complied with this criterion, so the final data set for further analysis consisted of 117 children.

The correlation coefficients between the main measures reflected a significant association between word spelling and word identification, $r(115) = .41$, letter knowledge, $r(115) = .30$, and receptive vocabulary, $r(115) = .39$, all $p$ values <.001. Word decoding showed a significant relationship with word identification, $r(115) = .27$, and vocabulary, $r(115) = .23$.

Relations Among Components of Reading and Writing: Fitting Models

For the full sample of 117 children, we tested whether the predictive relations between the components of decoding and spelling fitted a model including independent sublexical (letter knowledge) and lexical (word identification, vocabulary) paths of influence as well as an association between sublexical and lexical skills according to the assumptions of leading models of reading. Age was included in each model as a control variable, given its relationship with reading and writing. Statistical power was calculated for SEM models (Westland, 2010) with a value of .82 for an optimal sample between 55 and 167. When word decoding was included as the outcome measure (see Figure 1), the model provided a good fit for the data, $\chi^2 = 1.313$, $df = 1$, $p = .252$, $CMIN = 1.31$, $CFI = .984$; $TLI = .844$, $RMSEA = .052$. This model revealed that in grade-three Spanish children, only word identification was associated with decoding (11% variance explained) and that vocabulary was associated with word identification (12% variance explained). Notably, the model did not reflect a significant relation between letter knowledge and word identification.

When spelling was included as the outcome measure, the model provided a slightly better fit for the data, $\chi^2 = 1.265$, $df = 1$, $p = .261$, $CMIN = 1.26$, $CFI = .994$; $TLI = .994$, $RMSEA = .047$. Specifically, this model showed that letter knowledge, word identification, and vocabulary were related, with vocabulary having a stronger association with spelling.

Table 1. Mean Age and Descriptive Statistics on Accuracy of Data for Sublexical and Lexical Outcomes in Children Learning to Read in a Transparent Orthography.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Range</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>8.40</td>
<td>0.84</td>
<td>7.1–10.0</td>
<td>.25</td>
<td>1.07</td>
</tr>
<tr>
<td>Letter knowledge (raw)</td>
<td>18.7</td>
<td>2.94</td>
<td>5–20</td>
<td>-2.6</td>
<td>2.97</td>
</tr>
<tr>
<td>Receptive vocabulary (raw)</td>
<td>117.4</td>
<td>12.4</td>
<td>157–174 vuelta</td>
<td>.75</td>
<td>2.90</td>
</tr>
<tr>
<td>Word decoding (raw)</td>
<td>172.6</td>
<td>3.44</td>
<td>162–180</td>
<td>-.603</td>
<td>.587</td>
</tr>
<tr>
<td>Word spelling (raw)</td>
<td>150.1</td>
<td>16.83</td>
<td>95–180</td>
<td>-1.38</td>
<td>1.55</td>
</tr>
<tr>
<td>Word identification (raw)</td>
<td>104.0</td>
<td>12.66</td>
<td>79–180</td>
<td>-1.58</td>
<td>2.72</td>
</tr>
</tbody>
</table>

Note. $M =$ mean, $SD =$ standard deviation.
and vocabulary were independently and directly associated with spelling (28% variance explained). Vocabulary was related to word identification (12% variance explained), and again there was no significant association between letter knowledge and word identification.

In sum, both decoding and spelling models provided a good fit, reflecting that letter knowledge and word identification are distinct and independent procedures/skills for word decoding and spelling in grade-three children of a transparent orthography (Spanish). The main difference between the decoding and spelling models was that we found an association between spelling and letter knowledge but not between decoding and letter knowledge. These findings are discussed in detail in the Discussion section. A summary of the main findings of the role of letter and word knowledge on children’s reading and spelling outcomes is presented in Table 2.

**Discussion**

This work aimed to examine the role of sublexical (letter knowledge) and lexical (word identification, vocabulary) skills on children’s decoding and spelling accuracy in a very transparent orthography, Spanish. According to previous empirical evidence, letter
knowledge facilitated learning letter-sound mappings involved in accurate decoding of words (Share, 2004). Because mastering letter-sound mappings is sufficient to read and spell accurately in Spanish, children’s letter knowledge may play a long-lasting role in both decoding and spelling outcomes in this orthography. Alternatively, the relative ease of storing word representations in Spanish (Suárez-Coalla et al., 2016) might contribute to children’s reliance on word identification skills after some reading experience.

This was the first study to disentangle this issue in third-grade children exposed to a transparent orthography, employing both decoding and spelling outcomes. Two hypotheses were tested based on the assumptions of leading models of reading. For those models assuming one single procedure of reading (e.g., triangle model, Seidenberg & McClelland, 1989), letter knowledge and word identification should be associated with children’s decoding and spelling skills as well as with each other. Instead, for those models that account for two distinct reading procedures, letter knowledge, and word identification should show a unique and distinct association with children’s decoding and spelling accuracy, and little association was expected between letter knowledge and word identification. More specifically, dual-route models proposed that sublexical procedures take precedence over lexical procedures when children use letter-sound mappings for reading and spelling. On the other hand, when children relied on word identification skills, lexical procedures took precedence over sublexical ones. Overall, the present study revealed that (1) word identification was related to both word decoding and spelling, while letter knowledge was only related to spelling; and that (2) no association was observed between letter knowledge and word identification either for decoding or for spelling. These findings are discussed in detail below.

The Role of Letter Knowledge and Word Identification in Transparent Orthographies

Concerning our first question, whether sublexical (letter knowledge) and lexical skills (word identification and vocabulary) were related to third-grade children’s decoding and spelling accuracy, our results revealed that these associations differed for each outcome. While letter knowledge, word identification, and vocabulary skill were associated with spelling accuracy, only word identification skill was related to decoding accuracy. This outcome suggested that children employed their sublexical skills (the ability to associate letters and sounds) and lexical skills (the ability to retrieve the word form and access the word’s semantic meaning) to write the words. However, they primarily relied on lexical or word identification skills to decode words accurately.

One potential explanation for these findings was that we examined letter knowledge in conjunction with lexical skills, such as word identification and vocabulary. Previous studies that reported a long-lasting relationship between letter knowledge and decoding (Caravolas et al., 2013; De la Calle et al., 2018; Leppänen et al., 2008; Martínez & Goikoetxea, 2020) examined the role of letter knowledge in conjunction with pre-reading sublexical skills (phonological awareness). Some exceptions were the studies of Torppa et al. (2016) and Sunde et al. (2020) that included pre-reading sublexical skills and vocabulary as predictors of reading accuracy and comprehension.
These studies reported a link between vocabulary and reading comprehension, and replicated the association between letter knowledge and reading accuracy. However, our data revealed that when lexical skills are taken into account, Spanish children’s accurate decoding seemed to depend strongly on the word representations stored in their memory rather than on the knowledge of each letter and its corresponding sound. It is likely that the inclusion of word identification skills as a predictor in our study masked any effect of letter knowledge on word decoding. The lack of association between letter knowledge and decoding suggested that word identification skills may underlie individual differences in children’s reading accuracy rather than letter knowledge skills.

Word identification skills were also important for spelling. In this case, letter knowledge skills were also associated with children’s accurate spelling. The nature of spelling tasks might explain both this result and the associations found in previous studies between children’s spelling accuracy and letter knowledge (Paige et al., 2018; Sunde et al., 2020) or word identification skill (Lété et al., 2008; Mesquita et al., 2020). When a written word was encountered for reading, the child could activate the whole letter string in parallel and check whether that string matched a word representation in memory. However, spelling requires listening to the word, extracting the sequence of sounds to be spelled, and translating it into graphemes, even when the orthographic representation of a word was activated. The sequential nature of this task might support the direct involvement of letter knowledge in spelling outcomes.

Notwithstanding, the larger contribution of word identification to spelling (16%) as opposed to letter knowledge (9%) suggested that although letter knowledge is a foundational skill for reading and spelling in transparent orthographies, it becomes less relevant once children have been exposed to written words. Thus, our data aligned well with previous studies showing that lexical skills are involved in the accurate decoding and spelling of Spanish children between second and third grades (Cuetos & Suárez-Coalla, 2009; Suárez-Coalla et al., 2016). Although in transparent orthographies, applying letter-sound mappings could, in principle, be a suitable strategy for accurate reading (Rothe et al., 2015), our results suggested that children employ their word and semantic knowledge to decode and spell words accurately (Ouellette, 2006; Ouellette & Beers, 2010). This knowledge could facilitate the development of detailed mental representations and their accessibility for accurate reading and spelling (Walker et al., 2020; Walley et al., 2003). In sum, the present findings supported the view that as Spanish children progress through their first years of primary school, word identification and vocabulary should be specific targets of intervention in class rather than, or at least combined, with letter knowledge-based instruction.

**Letter Knowledge and Word Identification in Transparent Orthographies: Are They Related?**

Two different predictions were made regarding the second question about the relation between letter knowledge and word identification. One possibility was that the
association between letter knowledge and decoding was mediated by word identification skills. This association was expected within single reading procedure models in Spanish because letters and sounds are stable building blocks for words and their pronunciations in transparent orthographies. In single-route models (e.g., Seidenberg & McClelland, 1989), the simultaneous and reciprocal relationship between orthographic, phonological, and semantic representations at the letter and the word level would enhance reading accuracy. As the relationship between phonology and orthography in Spanish is straightforward, the activation of a letter or sound representation should directly activate the word’s orthographic and phonological representation according to this type of model. Consequently, an association should emerge between letter knowledge and word identification.

Another possibility was that even in transparent orthographies, children could identify words by retrieving the whole word information without any need to process the letters and their respective sounds. As children in such orthographies store exhaustive word forms after only a few reads (Suárez-Coalla et al., 2016), reading and spelling could rely on lexical strategies without any implication of sublexical skills. Consistent with this claim, we found that once children have acquired some reading experience, even in a transparent orthography such as Spanish, they decode and spell words on a lexical (word-level) basis rather than a sublexical (letter-level) basis. The lack of association between letter- and word-level processes—together with the independent relation of the letter and word-related skills to decoding and spelling—is consistent with models that assume two distinct procedures for reading. Indeed, the structural equation model (see Figure 1) reflects separate and independent paths, one that operates analytically to process letters with their sounds and the other that works globally to process words with their meanings. Further, these paths reveal a unique link between word identification and decoding and an independent association between letter knowledge, word identification, and vocabulary to spelling. Additionally, we found a strong and stable relationship between semantic knowledge and word identification in decoding and spelling, which aligned with the principles of models that incorporate two procedures for reading and writing (Schubert & McCloskey, 2015).

An important implication of this finding concerns the value placed on letter knowledge in primary school. It is true that the transparency of Spanish orthography boosts the employment of letter-sound mapping strategies in beginning readers (Cuetos & Suárez-Coalla, 2009) and facilitates the accurate reading of most words by first grade (Seymour et al., 2003). But it is also true that early accurate reading may boost children’s learning of a word’s orthographic forms as hypothesized by the self-teaching hypothesis (Share, 2008). In other words, the transparency of Spanish orthography might boost children’s early reliance on lexical strategies. This hypothesis explains our finding that by third grade, Spanish children rely primarily on their word and meaning knowledge for reading and spelling.

In sum, our study revealed that lexical strategies primarily support decoding and spelling accuracy in Spanish third-grade children, which could be explained by the high level of orthographic knowledge at the sample age (Share & Shalev, 2004). These results extend the scope of previous findings in exploring the combined role
of letter knowledge, word identification, and vocabulary skills on decoding and spelling measures in a transparent orthography. Together with the theoretical implications of the present study, our findings also have an applied side. Specifically, our results support that word identification skills should be incorporated into intervention programs for primary school children learning a transparent orthography like Spanish. Most intervention programs in this language emphasize pre-reading skills because of their role in later reading accuracy (Soriano et al., 2011). However, including discrimination and word identification training activities can have beneficial effects from second and third grades onwards to help children achieve the accuracy and speed necessary for reading and spelling success. There are several limitations, however, that should be noted. First, we employed exogenous variables because only one task was used to test each skill. It would be worthwhile for further studies to use latent variables with several measures for each skill in a larger sample size. Second, this was a cross-sectional study; therefore, it did not provide longitudinal data on the same children. It would be interesting to evaluate the same relations explored in this study at different ages to better understand the developmental trajectory of specific skills involved in reading and spelling.

Conclusions

The present study revealed the relative contribution of letter knowledge, word identification, and semantic skills in the reading and writing of primary school Spanish children. Our findings build upon previous evidence reported in Spanish (e.g., Cuetos & Suárez-Coalla, 2009) by demonstrating that children in transparent orthographies rely on their lexical and semantic knowledge not only for decoding but also for spelling. Additionally, our results suggested that knowledge of letter sounds is primarily employed as a residual strategy exclusively for word spelling.

Furthermore, our findings support the relative independence of sublexical and lexical strategies for reading and writing. Therefore, interventions starting in the second year of primary school should promote the expansion of vocabulary and the consolidation of orthographic representations to guarantee accurate and rapid access to the lexicon during reading and writing practice. In sum, we believe that the present data provide valuable information to design beneficial training to ensure progress in reading skills from the second year of primary school and optimize teaching strategies at this developmental stage.

Ethical Approval

This study was conducted under the guidelines of the ethical committee of the University of the Basque Country, project approval reference M10-2017-158.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by grant number PID2020-116740 GB-I00 (funded by the MCIN/AEI/10.13039/501100011033) from the Spanish Ministry of Science and Innovation and grant CIAICO/2021/172 from the Valencian Government.

ORCID iDs
Joana Acha https://orcid.org/0000-0002-4977-3413
Manuel Perea https://orcid.org/0000-0002-3291-1365

Supplemental Material
The Appendix referenced in this article and abstracts in languages other than English are available at http://journals.sagepub.com/doi/suppl/10.1177/1086296X241226476

References


Revista de Investigación en Logopedia, 12(2), 6. ISSN-e 2174-5218 https://doi.org/10.5209/rlog.78445


**Manuscript submitted:** 7 December 2022  
**Manuscript accepted:** 2 October 2023  
**Article published online:**