The role of letter knowledge acquisition ability on children’s decoding and word identification: Evidence from an artificial orthography

Joana Acha
Universidad del País Vasco UPV/EHU, San Sebastian, Spain and Biodonostia, Health Research Institute, San Sebastian, Spain

Nuria Rodriguez
Universidad del País Vasco UPV/EHU, San Sebastian, Spain

Manuel Perea
Universitat de Valéncia, Valencia, Spain and Universidad Antonio de Nebrija, Madrid, Spain

Background: Letter knowledge is crucial in the first stages of reading development. It supports learning letter-sound mappings and the identification of the letters that make up words. Previous studies have investigated the longitudinal impact of early letter knowledge on children’s further word reading abilities. This study employed an artificial orthography learning paradigm to explore whether the rate of letter learning modulates children’s reading and word identification skills.

Methods: In an initial training phase, 8-year-old Spanish children (N = 30) learned nine artificial letters and their corresponding sounds (two vowels and six consonants). The letter learning rate was set according to the number of attempts needed to name at least seven letters (i.e., 80% correct). These ranged from 1 to 4. In a second training phase, children visualized words made up of the trained letters while listening to their pronunciations. Some words included a context-dependent syllable (i.e., leading to grapheme-to-phoneme inconsistency), and others had an inconsistent syllable (i.e., phoneme-to-grapheme inconsistency). The post-test consisted of a reading aloud task and an orthographic-choice task in which the target word was presented with a distractor equal to the target except for the substitution of a letter.

Results: Children showed a high accuracy rate in the post-test tasks, regardless of whether words contained context-dependent or inconsistent syllables. Critically, the letter learning rate predicted both reading aloud and identification accuracy of words in the artificial orthography.

Conclusions: We provide evidence for the vital role of letter knowledge acquisition ability in children’s decoding and word identification skills. Training children on this ability facilitates serial letter-sound mapping and word identification skills. Artificial orthography paradigms are optimal for exploring children’s potential to achieve specific literacy skills.
Keywords: artificial orthography, learning, reading development, letter knowledge

Highlights

What is already known about this topic

- Children’s understanding of the alphabetic code in a given language is key to developing reading abilities.
- Letter knowledge is involved in word decoding by fostering letter-sound associations.
- Children’s letter knowledge is a crucial predictor of further decoding skills.

What this paper adds

- We employed an artificial orthography learning paradigm to mimic children’s process of reading acquisition, minimizing previous knowledge effects.
- The speed at which children acquire letter knowledge modulates not only decoding but also word identification outcomes.
- The results obtained in the artificial orthography paradigm replicate and extend the findings obtained in developmental studies with real orthographies.

Implications for theory, policy or practice

- Letter knowledge training is an important means to improve children’s decoding abilities and to facilitate the accurate identification of words.
- Children who struggle to learn letter sounds in the early years might be at risk of developing reading difficulties and should receive specific letter training.
- Artificial learning paradigms can be an important tool to test children’s learning ability when they have already been exposed to letters and words in their orthography.

Letters are key linguistic units for reading acquisition. Children must learn to identify letter symbols and link them with sounds to decode words accurately. They also have to remember how letters are arranged within the letter string to identify and discriminate words in their orthography. The close association between letter knowledge – a preliteracy skill involving knowledge about letters’ visual features and their names or sounds – and the development of reading skills is well established (Treiman & Kessler, 2003). For instance, letter knowledge has been identified as a critical early predictor of reading accuracy (Leppänen et al., 2008). Likewise, children with weak letter knowledge typically experience difficulty decoding words accurately. These findings suggest that identifying and naming letters are essential indicators of the child’s subsequent decoding ability across alphabetic orthographies (Caravolas et al., 2013; Snel et al., 2016).

The mechanism by which letter knowledge impacts reading ability stands at the letter level. It is based on the fact that it facilitates the application of the mappings between letters and sounds. However, while the empirical evidence supporting the relationship between
letter knowledge and reading is well established in opaque (Catts et al., 2015; Schatschneider et al., 2004) and transparent orthographies (de la Calle et al., 2018; Torppa et al., 2016), the question on whether letter knowledge also influences word identification remains a matter of interest (see Piasta & Wagner, 2010). Some studies suggest that the mechanism by which letter knowledge impacts word identification stands at the word level by optimizing the visual encoding of letters within whole letter arrays (Grainger et al., 2012). Evidence for this view is that children’s letter-sound knowledge correlates with the ability to discriminate a real word from an orthographic distractor (e.g., door-doar) in orthographic-choice tasks (Wang et al., 2014). Additionally, the accuracy by which primary school children discriminate a string that contains letters in high-frequency positions in word-likeness tasks (e.g., blay-btay) is related to their sight word recognition abilities (O’Brien, 2014). Thus, these studies suggest that the role of letter knowledge in word identification could arise from the child’s ability to integrate the knowledge of individual letters and print-based codes, such as a specific letter sequence on a letter string that constitutes a word.

Importantly, previous research has reported that fast acquisition of letter knowledge accounts for reading fluency in Finnish—a very transparent orthography—even after phonological abilities have been controlled for (Torppa et al., 2006). This raises the possibility that a child’s letter knowledge acquisition speed underpins word reading and identification skills. We believe this research question needs further research, particularly in transparent orthographies, where unambiguous letter-sound correspondences enable children to quickly identify and encode letters within words. The present work explores this issue using an artificial orthography learning paradigm.

The Role of Letter Knowledge in Reading Development

Longitudinal studies have consistently shown that, together with phonological skills, letter knowledge—typically measured by asking the child about letter’s names or sounds—in kindergarten is a strong predictor of children’s reading skills in primary school, both in opaque (Catts et al., 2015; Muter et al., 2004) and in transparent orthographies (Kim & Pallante, 2012; Leppänen et al., 2008). Similarly, a deficit in letter knowledge has been longitudinally associated with poor word and nonword reading accuracy and fluency (Lyytinen et al., 2006; Thompson et al., 2015). An explanation of this relation is that letter knowledge supports learning letter-sound mappings necessary for accurate word reading. Specifically, letter knowledge in a given alphabetic system contributes to the transition toward the phoneme as a processing unit (Blaklock, 2004; Treiman et al., 1996) and activates the sensitivity toward specific grapheme-phoneme relations, facilitating the learning of grapheme-phoneme mappings (Huang et al., 2014).

However, little is known about the involvement of letter knowledge in accurately identifying words, that is, in discriminating a written word from other words in the lexicon. The results of several studies are compatible with the notion that letter knowledge facilitates visual encoding and identification of letters within words. For instance, letter knowledge in preschool children correlates with letter detection accuracy within letter strings (Bosse & Valdois, 2009), and detection accuracy predicts accurate word identification in orthographic choice tasks (Van Den Boer et al., 2015). It is, therefore, conceivable that mastering letter knowledge fosters the identification of letters within words, a key developmental step to ensure accurate word identification (Pritchard et al., 2018). The fact that
pre-readers identify written words once they know the names of at least some letters embedded in such words supports this view (e.g., identifying the word *MASK* when seeing the string *MSK*, Ehri & Wilce, 1985). This result aligns with the reported association between letter knowledge and accurate word discrimination in orthographic choice tasks in primary school children (Wang et al., 2014). These findings reveal that knowledge about single letters is a prerequisite for the accurate configuration of letters within words necessary for word identification.

Interestingly, empirical evidence suggests that the speed at which the child acquires letter knowledge can also predict later word reading accuracy and word identification. A study by Torppa et al. (2006), conducted in a very transparent orthography (Finnish), provides strong support for this view. In a follow-up study of a cohort from kindergarten to primary school, the authors categorized children according to their speed in learning letter sounds into delayed, typical, or precocious letter learners. They also analysed the predictive power of this categorization on later reading skills. The delayed letter learning condition was found to underpin worse reading fluency and comprehension in first grade, irrespective of the child’s risk of reading difficulties. More recently, Sunde et al. (2020) observed that Norwegian children whose teachers instruct on letters at a faster pace during the first school year – introducing three or more letters per week – show greater letter-sound knowledge and sight word identification skills at the end of the school year, compared with those children whose teachers instruct on letters at a slower pace – over a longer time frame. An explanatory argument for this effect is that early acquisition of letter knowledge ensures a greater number of encounters with those letters within written words. These findings sustain the notion that the speed by which a child acquires letter knowledge may indicate children’s potential for accurate reading and word identification.

An interpretive drawback of these studies is that they do not allow exploring the learning process from scratch. As children’s exposure to written language differs, employing real letters for evaluation makes it difficult to examine whether the effects are associated strictly with acquiring letter knowledge. One way to overcome this difficulty is using an artificial orthography, that is, a set of symbols unknown to all participants, who are trained to learn the symbols in isolation or unknown words constructed with these symbols. Subsequently, reading or recognition rates are evaluated. This paradigm makes it possible to examine the course of development of processes associated with letter and word learning by manipulating variables present in a real orthography (transparency, frequency), ruling out any possible effect due to prior experience.

**Contributions of Artificial Orthography Learning Paradigms**

In an influential study, Taylor et al. (2011) designed an artificial orthography learning paradigm in which adults were exposed to 36 words with CVC structures built with symbols that simulated English consonants (12) and vowels (6). In the first exposure phase, participants had to see whole pseudowords while listening to their pronunciations. In the second phase, they had to read aloud the same words. In the post-test orthographic-choice task, participants had to discriminate the previously presented structures (learned word vs. distractor). Participants showed both lexical (better identification of trained structures) and sublexical effects (poorer reading aloud of pseudowords with inconsistent pronunciations) typical in the word recognition literature. Using this paradigm, Schmalz et al. (2022) observed that the reading accuracy of artificial words was higher when adults were trained...
only on words with predictable pronunciations than when these words were mixed with inconsistent words with unpredictable pronunciations. This finding suggests that the ease of establishing the relationship between letters and sounds is key to promoting decoding skills. Samara and Caravolas (2017) compared identification rates of pseudowords constructed with real letters or symbols in two groups of adults with and without dyslexia. The training phase consisted of a visual presentation only (without pronunciation). The authors observed that both groups exhibited reasonably high identification rates in the orthographic-choice task, but dyslexic adults showed worse results with real letters. This outcome raises the possibility that having difficulty learning letter-sound mappings has a detrimental effect on reading and word identification skills. Rastle et al. (2021) compared two experimental conditions to test this issue. One group of adults received a pre-training session on single symbol-sound correspondences, while the other group did not. The training sessions included reading-focused (reading aloud; orthographic search) and meaning-focused (say the meaning; meaning judgement) tasks. Interestingly, participants who previously received the explicit symbol-sound instruction showed higher accuracy rates in several reading post-test tasks (i.e., reading aloud the trained words and new words) and in recognition memory tasks (i.e., identifying the trained words among a set of semantic and phonological distractors that were presented one by one). These findings suggest that learning to read and identify words is enhanced by prior knowledge of letter sounds in adult learners.

The process of learning to read using artificial orthographies has also been studied in children. A seminal study by Gombert and Peereman (2001) trained children aged 5 to 7 years to match three-symbol visual structures with their pronunciation. The post-test task consisted of matching the learned structure with two possible – correct or incorrect – pronunciations. As in the Taylor et al. (2011) study, children showed better identification rates when the structures were consistent than when they were inconsistent. In the same line, Arciuli and Simpson (2012) exposed children between 5 to 12 years to combinations of picture triplets visually while listening to pronunciations. Afterward, children had to read the presented triplets. Interestingly, the children who could accurately decode triplets were the ones who could better identify them in the post-test orthographic-choice task. These findings suggest that mastering letter-sound mapping abilities support children’s performance in word identification tasks, at least in orthographies following opaque spelling rules. Aravena et al. (2018) explored the specific impact of letter-sound knowledge on reading. They trained two groups of Dutch children aged 7 to 11 years to identify and name artificial (Hebrew) letters through a computer game. Learning was evaluated with two post-test tasks (letter-sound association and reading words made up with the artificial letters). Also, children’s reading and writing skills in Dutch were assessed. Accuracy in the artificial letter-sound association task correlated with the decoding accuracy of words made up of artificial letters, and it also predicted spelling accuracy in Dutch. This result was replicated by Law et al. (2018) in a sample of 8-year-old Dutch children, reinforcing the notion that letter knowledge is crucial for learning letter-sound mappings and applying them to read words successfully.

An important question is whether the speed at which children acquire letter knowledge modulates children’s reading and word recognition accuracy. According to previous longitudinal evidence, those children who are slower to acquire letter knowledge are more likely to experience reading difficulties later on (Torppa et al., 2006). One possibility is that early acquisition of letter knowledge ensures mastery of letter-sound associations and more opportunities for a successful reading experience which, in turn, facilitates the formation of
accurate word representations (Share, 2004). This hypothesis predicts a direct impact of letter knowledge on reading accuracy. Another possibility is that children who acquire letter knowledge faster have more opportunities to detect those letters within words and to grasp specific letter configurations within letter strings (Bosse & Valdois, 2009). Within this view, a direct impact of letter knowledge on word identification is expected. Therefore, we can expect children with greater ability to acquire letter knowledge – measured in the time required to gain that knowledge – might read and identify words more accurately than children with a lower ability to reach the same amount of letter knowledge. This study aims to test this issue empirically using an artificial orthography paradigm.

The Present Study

The main goal of the present experiment is to examine whether children’s ability to learn new letters modulates the decoding and identification accuracy of words made up of these letters. To that aim, we designed an artificial letter learning task in which children were trained to name and copy nine new, unfamiliar symbols. Afterward, children were exposed to words made up with the trained symbols while listening to word pronunciations. Artificial words were created simulating the characteristics of Spanish, a highly transparent orthography with two exceptions: context-dependent syllables (letter c is pronounced /k/ when followed by vowels a, o, u, but is pronounced /θ/ when followed by vowels e, i) and phoneme-to-grapheme inconsistent syllables (letters b and v share the same pronunciation /b/). Then, children had to complete two post-test tasks: reading (i.e., read aloud the trained words) and orthographic choice (i.e., identifying the trained word when presented with a form-similarity distractor). Based on children’s speed in learning the sounds of at least seven symbols, we created a letter learning threshold that reflected children’s learning pace (i.e., at least 80% correct). The best learners could name 7 out of 9 symbols in the 1st attempt (i.e., Threshold 1), and the worst learners reached this threshold in the 4th attempt (i.e., Threshold 4). The predictions are straightforward. If the child’s letter learning ability paves the way for the acquisition of decoding and word identification skills, children who learn the letters in the initial thresholds will be able to read and identify artificial words with greater accuracy than children who learn the letters at later thresholds. This outcome would reveal that artificial orthography paradigms can help measure childrens’ learning potential. An additional exploratory question is whether the presence of context-dependent and inconsistent syllables within artificial words modulates the findings in the post-test tasks.

Methods

Participants

A total of 30 primary school children in 3rd grade (M age = 8.5 years, SD = 0.45, range = 7.4–8.9; 19 females) participated in this experiment with the informed consent of their parents. Sample size was based on previous artificial learning research conducted with children of the same age (N = 28, 36 trials, Nigro et al., 2015) and in which participants showed good item-based learning in one single session. This sample was selected from a school in an urban area of the Basque Country, Spain. Children met the following inclusion criteria: (a) be enrolled in 3rd grade; (b) absence of neuropsychiatric disorder or sensory problems; (c) no history of special education services or reading or language
therapy. The method of reading instruction used in the school was the phonics method. All participants were bilingual Spanish-Basque, Spanish being L1 and schooling language.

Basque is also a transparent language in which all letter-sound correspondences overlap with Spanish, with a few exceptions (letters c, v, ñ and y do not exist in Basque).

Materials and Design

Control Measures

Nonverbal Intelligence. Children’s nonverbal reasoning ability was assessed using the Kaufman Brief Intelligence Test matrix task (K-BIT, Kaufman & Kaufman, 1990) to discard individual differences due to IQ. Internal consistency estimates for this subtest range from 0.74 to 0.93. Typified measures showed that all children were within the normal range (M = 107.27; SD = 10.73, range = 89–138).

Phonological Memory. Based on the classic paradigm, it was measured using the non-word repetition task (Hulme & Tordoff, 1989). This consists of auditory-presented sets of syllabic nonwords, which the child is asked to repeat in the same order. Four sets of six nonwords each were presented matched on syllable frequency (mean per set = 12.1) and complexity. The first set consisted of two-syllable chains, and each set incorporated an additional up to a total of five. The total span was obtained by summing the number of correctly remembered items (M = 12.7, SD = 4.7; range = 5–21).

Artificial Orthography Learning Measures

Letter Training Phase. The stimuli consisted of 9 artificial letters drawn from the Armenian and Georgian alphabets: two vowels (sounds /a/ and /i/), one consonant that could lead to context-dependent syllables within words (linked to sound /θ/ but pronounced /k/ when followed by a), two consonants that could lead to phoneme-to-grapheme inconsistent syllables within words (both linked to the same sound /b/) and four regular consonants (/f/ /m/ /t/ and /s/ sounds). In this phase, children were exposed to letter images and their sounds (note that the letter name is equivalent to its sound in Spanish due to its transparency). This phase aimed to detect how easy it was for the child to learn the shape of the letter and its corresponding sound.

Word Training Phase. This phase allowed the child to realize that specific associations between letters and sounds could vary depending on the letter combination in the syllable. For this purpose, 24 two-syllable words with a CV.CV structure were constructed. These words were divided into three sets, eight words that included a context-dependent syllable (ՔՋ, ՔՋ, ՔՋ, ՔՋ, ՔՋ, ՔՋ, ՔՋ -kami, kati, tshi, shi, shi, shi, shi -mati, fima, safi, sifa, tami, mita, tisi, tisa-) and eight words that included an inconsistent syllable (ՔՋ, ՔՋ, ՔՋ, ՔՋ, ՔՋ, ՔՋ, ՔՋ -fab, savi, tiba, miva, bita, bafi, vima, vasi). The position of the key letter was counterbalanced across the first and second syllable positions. The items employed in the artificial orthography are presented in Table 1.
Reading Post-Test Task. The same items employed for the word training phase were presented in the reading task in random order. Each child saw each item individually and was instructed to read it aloud.

Orthographic-Choice Post-Test Task. To evaluate whether children were able to recognize the trained words, two stimuli were presented at the same time. One stimulus was the target, and the other was a distractor with a similar form. This was equal to the target except for one letter, based on the evidence that form-similarity pairs generate interference due to orthographic overlap (Bowers et al., 2005). Two blocks of items were created. One block consisted of the target word and a distractor resulting from the substitution of a vowel in each set: context-dependent (e.g., cami-cimi), regular (e.g., mati-miti) and inconsistent (e.g., bita-bata). In the other block, the distractor resulted from the substitution of a consonant in the context-dependent (e.g., cami-sami), consistent (e.g., mati-fati) and inconsistent set (e.g., bita-vita). The mean log bigram frequency was controlled among the experimental and distractor items (1.96 vs. 1.95, respectively). All tests were implemented in a computer using DMDX to guarantee the homogeneous presentation of stimuli. Distractor location – right or left – was counterbalanced across trials.

Procedure

The study was conducted under the guidelines of the ethical committee of the University of the Basque Country UPV/EHU, project approval reference M10_2017_158.

All tests were administered individually. The first session consisted of individual application of cognitive and reading tests. The artificial orthography learning procedure was carried out in the second session as a video game.

Letter Training Phase. In the first phase, an image was presented with some child explorers and a group of aliens explaining the participant’s mission, which consisted of learning the alien alphabet. Each participant was asked to observe and listen to the sound of each letter. Subsequently, each letter was presented on the screen individually, along with the audio of its sound for 1000 ms for each item. A fixation point during 400 ms anticipated the appearance of each letter. After listening to the letter sounds, the experimenter handed over a template with the original letters for the child to copy each letter by hand and asked her to pronounce the corresponding sound. During learning, feedback on misnamed letters was provided. Depending on the error rate, an optimal learning threshold was established, estimated at seven well-named letters out of 9 (a correct naming of 78% of the total items to be learned). Those children who named 7 of the nine letters well on the
first attempt (learning threshold 1) did not repeat the phase. The rest repeated the letter reading phase until they correctly named seven letters. This way, four letter learning thresholds were established: four was the maximum threshold needed to achieve the seven-letter criterion that determined the letter learning pace.

**Word Training Phase.** After the letter learning phase, children were presented with an image of the same child explorers, who explained that in this phase, the participant should learn whole words in the alien language. Each target word was presented individually with its corresponding pronunciation for 2000 ms. Before each item, the computer screen was blank for 1000 ms, and a fixation point for 400 ms anticipated the appearance of the next word. The word set was presented twice, and words were randomized within each block.

**Word Reading Task.** In the post-test phase, the experimenter presented a card with a randomized list of words, and each child had to read each word aloud. Each well-decoded word was scored with one point. A zero was assigned to each badly decoded word regardless of the type of error (substitution or use of a sound that did not correspond to the letter in its position). Accuracy was the dependent variable.

**Orthographic-Choice Task.** The task consisted of 48 trials distributed in two blocks. The 24 trained items were presented twice, once with the consonant distractor and once with the vowel distractor. In each trial, the child was given two items, one on the right and one on the left of the centre of the screen and was asked to identify as fast as possible which one was the familiar – trained – item. Accuracy rates reflect the discrimination accuracy of the familiar item relative to the distractor. Trials were randomized across blocks, and the distractor position – left or right – was counterbalanced between conditions.

**Data Analysis**

The data in the letter learning phase were used to establish a learning threshold of 1 to 4. This letter learning threshold was employed as a learning pace measure. Specifically, five children learned letter sounds on the first attempt, nine subjects on the second, 11 on the third, and five on the fourth (17%, 30%, 37% and 17% learning at threshold 1, 2, 3 and 4, respectively).

The accuracy data for the two post-test tasks were modelled with the Bernoulli family function (1 = correct, 0 = incorrect) using Bayesian linear mixed-effect models (brms package in R, Bürkner, 2019) – this procedure allows us to fits the maximal random-effect structure models. The predictors were Threshold (1, 2, 3 and 4; a zero-centered quantitative variable) and the fixed factor Word Type (context-dependent, regular, and inconsistent, where ‘regular’ was the reference level for the contrasts). The model was accuracy ~ threshold * type + (1 + typelsubject) + (1|item). For the fits of each model, we employed four chains, each with 5000 iterations (1000 warm-ups). The chains converged successfully (\( R^\hat{\text{S}} = 1.00 \)). The output indicates each fixed effect’s estimate, standard error and 95% credible interval (CrI). Those estimates whose 95% CrI did not overlap with zero are interpreted as evidence of an effect.
Results

None of the children showed floor or ceiling effects in the artificial reading and identification tasks. The descriptive data of the measures in the artificial orthography tasks are presented in Table 2.

Examination of Letter Learning Threshold on Word Reading

The analysis of reading accuracy showed a decreasing linear effect for Threshold ($b = -0.70$, $SE = 0.34$, 95% CrI [−1.39, −0.05]): word reading accuracy diminished slightly as the letter learning threshold increased. Notably, there were no trends of an effect of Word type (context-dependent vs. regular: $b = -0.11$, $SE = 0.38$, 95% CrI [−0.85, 0.65]; inconsistent vs. regular: $b = -0.00$, $SE = 0.38$, 95% CrI [−0.74, 0.77]) or an interaction with Threshold (see Figure 1).

Examination of Letter Learning Threshold on Orthographic-Choice

The analyses on identification accuracy revealed a linear trend for Threshold ($b = -0.69$, $SE = 0.29$, 95% CrI [−1.28, −0.15]): Word identification accuracy diminished slightly as the letter learning threshold increased (see Figure 2). The lack of interaction between Threshold and Type of Word (context-dependent vs. regular: $b = -0.32$, $SE = 0.43$, 95% CrI [−1.22, 0.50]; inconsistent vs. regular: $b = -0.40$, $SE = 0.41$, 95% CrI [−1.28, 0.33]) reflected that accuracy decreased progressively across thresholds in a similar way for words containing regular context-dependent and inconsistent syllables.

An interpretive question is whether children’s learning outcomes were merely associated with better cognitive skills. We created three separate models for Threshold, reading and word identification accuracy as dependent variables to examine this issue. The predictors were IQ and phonological memory. None of the models showed any clear signs of an impact of these predictors: 3% of the variance in letter threshold ($\beta = -.16$, $p = .40$; $\beta = -.11$, $p = .54$ for IQ and memory, respectively), 0.4% in reading ($\beta = .04$, $p = .94$; $\beta = .06$, $p = .74$) and 4% in word identification rates ($\beta = .02$, $p = .91$; $\beta = .19$, $p = .30$). These results provide evidence in favour of letter-sound and word learning skills being involved in children’s learning rates in the artificial orthography paradigm rather than in more general cognitive skills.

Discussion

This paper examined whether the speed at which children acquire letter knowledge modulates word decoding and identification skills. To that aim, children learned an artificial orthography that simulated the properties of Spanish (i.e., orthographic transparency and the existence of context-dependent and inconsistent symbol-sound correspondences in the syllable). We included a letter learning phase in which children learned each symbol’s shape and its corresponding sound and a whole-word exposure phase in which children were exposed to symbol arrays while listening to their pronunciations. This phase was aimed at noticing context-dependent or inconsistent syllables within words. The number of repetitions required to reach a minimum threshold of letter knowledge (7 letters out of 9) was set as a measure of speed (ranging from Threshold 1 when children learned the letters on the first
Table 2. Descriptive measures of artificial orthography post-test tasks over participants.

<table>
<thead>
<tr>
<th>Distractor</th>
<th>Con-dep</th>
<th>Regular</th>
<th>Incons</th>
<th>Con-dep</th>
<th>Regular</th>
<th>Incons</th>
<th>Con-dep</th>
<th>Regular</th>
<th>Incons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Threshold 1</td>
<td>0.95 (6.8)</td>
<td>0.96 (5.5)</td>
<td>0.95 (6.8)</td>
<td>0.97 (5.5)</td>
<td>0.97 (5.5)</td>
<td>0.95 (11.1)</td>
<td>0.87 (20.3)</td>
<td>0.99 (0.1)</td>
<td>0.95 (11.1)</td>
</tr>
<tr>
<td>Threshold 2</td>
<td>0.87 (16.5)</td>
<td>0.87 (17.7)</td>
<td>0.92 (10.8)</td>
<td>0.93 (6.3)</td>
<td>0.91 (9.1)</td>
<td>0.98 (4.2)</td>
<td>0.96 (9.1)</td>
<td>0.94 (6.3)</td>
<td>0.92 (4.2)</td>
</tr>
<tr>
<td>Threshold 3</td>
<td>0.87 (12.5)</td>
<td>0.86 (13.1)</td>
<td>0.87 (15.8)</td>
<td>0.95 (8.4)</td>
<td>0.89 (10.9)</td>
<td>0.96 (5.1)</td>
<td>0.96 (5.1)</td>
<td>0.96 (8.4)</td>
<td>0.89 (13.1)</td>
</tr>
<tr>
<td>Threshold 4</td>
<td>0.62 (35.5)</td>
<td>0.70 (31.9)</td>
<td>0.60 (41.1)</td>
<td>0.81 (24.3)</td>
<td>0.80 (10.4)</td>
<td>0.75 (22.3)</td>
<td>0.76 (16.2)</td>
<td>0.80 (16.2)</td>
<td>0.78 (16.2)</td>
</tr>
</tbody>
</table>
Figure 1. Possible parameter values in the posterior distributions of the estimates in the Bayesian linear mixed-effects models of accuracy in the reading task (Threshold \(\times\) Type of word). The dot line corresponds to an effect of zero. The red area corresponds to the tail out of the 95% credible interval. The blue, green, orange and purple areas correspond to the 50%, 75%, 89%, and 95% credible intervals, respectively. The upper right part of the figure reflects children’s reading accuracy rate (y axis) in each letter learning Threshold (x axis) per Word Type.

Figure 2. Possible parameter values in the posterior distributions of the estimates in the Bayesian linear mixed-effects models of accuracy in the orthographic-choice task (Threshold \(\times\) Type of word). The dot line corresponds to an effect of zero. The red area corresponds to the tail out of the 95% credible interval. The blue, green, orange, and purple areas correspond to the 50%, 75%, 89%, and 95% credible intervals, respectively. The upper right part of the figure reflects children’s word identification accuracy rate (y axis) in each letter learning Threshold (x axis) per Word Type.
LETTER KNOWLEDGE IN ARTIFICIAL ORTHOGRAPHY

attempt, to Threshold 4, the maximum number of attempts to acquire the minimum knowledge). The two post-test tasks assessed reading accuracy and the identification of previously presented artificial words. Overall, the data showed that (i) participants were able to read and identify the trained artificial words with high accuracy, regardless of whether they entailed transparent, context-dependent, or inconsistent syllables, (ii) the speed at which children learned the artificial letters modulated the decoding rates and automatic identification accuracy of the trained words. This pattern of results supports the hypothesis that children’s speed of acquiring letter knowledge impacts their ability to properly apply letter-sound mappings and identify words accurately.

Letter Knowledge Predicts Reading Aloud

The first aim of the study was to examine whether the ability to learn artificial letters modulated the accuracy of artificial word reading. Our results replicate and extend previous evidence with adults in an artificial orthographic paradigm (Rastle et al., 2021) by showing that the child’s speed of acquiring letter knowledge (i.e., knowledge about symbol shape and sound) is a powerful predictor of the child’s potential to read words. Children could decode the trained words by applying letter-sound mapping rules with high accuracy. Critically, those who experienced greater difficulty acquiring letter knowledge showed worse reading accuracy rates. This finding is consistent with the longitudinal relationship observed between early letter knowledge and decoding in primary school, both in transparent (de la Calle et al., 2018; Leppänen et al., 2008) and in opaque orthographies (Caravolas et al., 2013; Muter et al., 2004), suggesting that letter knowledge is a crucial foundational skill for reading in all alphabetic orthographies. Our results make an additional contribution to this evidence by demonstrating that the speed at which children acquire letter knowledge predicts reading accuracy. In other words, learning letter sounds faster facilitates the successful application of letter-sound associations, which may cascade into reading fluency (see Torppa et al., 2006).

It is essential to highlight that even when our learning procedure included isolated letter-sound training, children could quickly learn inconsistent and context-dependent letter-sound mappings only after two exposures to whole words by inferring that the learned sound was not always mapped to the learned letter in the syllable. One explanation is that explicit training in letter knowledge encouraged the identification of letters in syllables and their association with specific sounds. This explanation is consistent with Treiman and Kessler’s account (2003), who argue that letter knowledge allows categorizing the letter as a visual image and associating it with a verbal unit, thus helping the child infer the relationship between graphemes and phonemes when exposed to words. Additionally, children were trained in a transparent orthography, which could have facilitated learning exception letter-sound mapping rules during the word exposure phase (Seymour et al., 2003).

One limitation of the present study is that phonological awareness was not included as a control variable (Torppa et al., 2006, 2016). Although our data rule out the hypothesis that children’s IQ and phonological memory underlie the observed learning rates, the possibility that children’s phonological awareness could influence children’s ability to apply letter-sound mappings when reading artificial words cannot be discarded. It has to be noted, however, that the studies conducted in real orthographies report a predictive value of letter knowledge on word reading that goes beyond that of phonological awareness (e.g., Catts et al., 2015; de la Calle et al., 2018). Additionally, intervention studies have
revealed that instruction in letter knowledge benefits the acquisition of basic reading skills irrespective of whether phonological awareness is included in the instruction (Piasta & Wagner, 2010).

**Letter Learning Pace Predicts Word Identification**

Another important goal of the study was to examine whether the speed at which letter knowledge was acquired could modulate the identification accuracy of the trained words. Our results support this hypothesis: children’s letter learning threshold modulated their identification accuracy in the orthographic-choice task. Indeed, the letter learning threshold was the factor that differentiated those children with a better word identification accuracy (97% precision) and those with the worst identification accuracy (75%). Specifically, as children acquired letter knowledge at later Thresholds, their identification accuracy rates decreased. Remarkably, the impact of learning the letters earlier was beneficial for identifying all words, either regular or containing context-dependent and inconsistent syllables. One potential explanation is that children who learned the letters faster paid more attention to letter combinations within words, so all trained items were equally easy to discriminate afterward.

Considering that there was no previous familiarity with the trained words and that these were presented only twice (number of presentations to ensure whole-word learning, Nation et al., 2007), our results underline the critical role of letter knowledge in word decoding and identification skills. Systematic symbol-sound training – note that our instruction also included writing the symbols – promoted both letter-sound learning and word identification skills, probably by facilitating the encoding of the exact artificial letter combinations (Bossé & Valdois, 2009). Consistent with this view, McCandliss et al. (2003) reported that children who are less proficient in using letter-sound mappings tend to rely on global visual strategies, which can hinder letter-encoding processes within words. It is plausible that such children need additional letter knowledge training and word-level letter encoding strategies to ensure correct word memorization.

The finding that letter learning abilities can serve as early support for the consolidation of words in the lexicon is consistent with the principles of the multiple path model (Grainger et al., 2012), according to which letter knowledge can optimize the process of letter identification and location within words. The end result is the effective coding of specific letter combinations and the automatic processing of complete word representations. Our findings are also compatible with connectionist and dual-route computational models that assume that any training focusing on letters and their corresponding names favours the generalization of sublexical regularities (Harm et al., 2003) and boosts access to consolidated word representations.

At this point, it is important to underscore that children in this study showed very high reading and word identification rates in the artificial language. This could be due to the striking differences between our paradigm and those employed in previous artificial orthography studies. Firstly, symbol-sound correspondences matched grapheme-phoneme correspondences in the participants’ native orthography. That is, each symbol had a one-to-one mapping to a letter name that was already available in their phonological repertoire. Additionally, the trained words respected the Spanish letter-sound rules, including how the context affected the letter c. This similarity between the artificial and Spanish orthography could explain our high reading and identification rates. Secondly, how letters were trained
could have facilitated learning the trained words. Clearly, writing the letters, together with
the naming and corrective feedback – as done in the present study – favoured the adequate
internalization of the concrete form of the letter and thus strengthened its association with
the corresponding name (James, 2017). This hypothesis is in line with the findings of
Aravena et al. (2013). They observed that the learning rate was higher in an explicit letter
training condition (observation, naming and feedback) than in an implicit learning condi-
tion (through a trial-and-error video game). These reasons and the limited number of words
in our experiment might have facilitated children’s attention to how letters were arranged in
words, resulting in high decoding and word identification rates.

In sum, by employing an artificial orthography paradigm, this study demonstrated that
children who acquired letter knowledge more rapidly could read and identify artificial
words with greater accuracy. Our data sustain the claim that letter knowledge provides
the child with a piece of foundational knowledge for the development of reading skills. Ex-
ploring letter knowledge acquisition ability at the beginning of the reading stage may be
helpful for clinicians or educators to detect children’s potential literacy difficulties in time.
To that end, an artificial orthography paradigm, such as the one employed in the current
study, can be a very convenient tool.

Acknowledgements

The authors acknowledge the invaluable help of all the children and families that willingly
have taken part on this study. We confirm that the study has been conducted under the
guidelines of the ethical committee of the UPV/EHU project approval reference
M10_2017_158 and has been partially 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 Number CIAICO/2021/172 from the Valencian
Government. Correspondence concerning this article should be addressed to Joana Acha.
[Correction added on 8 August 2023, after first online publication: The funding details
have been updated in this version]

Data Availability Statement

The materials, the trial level data, the scripts, and the outputs are available at the OSF link:
https://osf.io/5h3q7/?view_only=aaad3ef90ff477bbf849a132ed4f359

References

https://doi.org/10.1016/j.jecp.2013.03.009

(6), 552–564. https://doi.org/10.1177/0022219417715407


LETTER KNOWLEDGE IN ARTIFICIAL ORTHOGRAPHY


Joana Acha is a psychologist, expert on cognitive neuroscience and child development. She obtained the Extraordinary Doctorate Award and the SEPEX Research Award in 2009. She is currently a associate professor and researcher in child cognitive development at the Department of Basic Psychological Processes and their Development at the University of the Basque Country UPV/EHU. She has collaborated with various laboratories investigating the mechanisms involved in language and reading in adults (Department of Psychological and Brain Sciences, University of Massachusetts Amherst) and in bilingual and monolingual children with and without learning difficulties (La Mente Bilingüe Laboratory, UPV/EHU; Basque Center of Cognition Brain and Language). She is an ANEP project reviewer since 2006 and a member of the Frontiers in Language Sciences editorial review team since 2011. Her current line of research is the study of the mechanisms involved in the development of language and reading with three aims: (1) detect the influence of cognitive mechanisms of early development such as attention or executive functions in the developmental trajectory of language and reading and related disorders such as SLI or dyslexia; (2) compare different learning methods – handwriting, repetition and visual recognition – to optimize the process of linguistic and reading development; and (3) identify those variables of the family context that can promote or hinder such development.

Nuria Rodriguez is an Educational Psychologist working for the public educational system. She is also a PhD student at the University of the Basque Country UPV/EHU. She is linking her educational practice on detection and intervention of developmental disorders and her research work.

Manuel Perea obtained his bachelor degrees in both Philosophy and Psychology in 1989 at the University of Valencia (“National Award for best Academic Record”). After his dissertation (“Excellent Cum Laude”, 1993), he completed his formation abroad in top labs of word recognition and reading: Ken Paap (4 months: NMSU, 1994), Ken Forster (3 months: University of Arizona, 1995), Sandy Pollatsek and Keith Rayner (1 year: UMass, 1996–1997), Steve Lupker (2 months: University of Western Ontario, 1999), Roger Ratcliff (1 year: Northwestern University, 2001–2002), and John Gabrieli (6 months: MIT). His research focuses mainly on the experimental study of reading processes and how they are influenced by orthographic, phonological, and semantic aspects. Likewise, he has paid close attention to special populations (reading in braille and reading in deaf people), as well as to the study of bilingualism.

Received 26 April 2022; revised version received 13 May 2023.

Address for correspondence: 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