

Shaping the precision of letter position coding by varying properties of a writing  
system

Clare Lally,<sup>1</sup> J.S.H. Taylor,<sup>2</sup> Chang Lee,<sup>3</sup> & Kathleen Rastle<sup>1</sup>

1. Department of Psychology, Royal Holloway, University of London
2. School of Life and Health Sciences, Aston University
3. Department of Psychology, Sogang University

RUNNING HEAD: SHAPING PRECISION OF POSITION CODING

Correspondence Address:

Kathleen Rastle  
Department of Psychology  
Royal Holloway, University of London  
Egham, Surrey TW20 0EX  
United Kingdom

Phone + 44 (0)1784 443 716  
Email: [Kathy.Rastle@rhul.ac.uk](mailto:Kathy.Rastle@rhul.ac.uk)

Abstract

There is substantial debate around the nature of letter position coding in reading. Research on a variety of Indo-European languages suggests uncertainty in position coding; for example, readers perceive transposed-letter stimuli (jugde) as similar to their base words (judge). However, these effects are not apparent for all languages. We developed a powerful new method to delineate how specific properties of a writing system shape the representation of letter position. Two groups of 24 adults learned to read novel words printed in artificial scripts. One group learned a dense orthography (i.e. with many anagrams) and one group learned a sparse orthography (i.e. no anagrams). Following four days of training, participants showed a larger transposed-letter effect in the sparse orthography than in the dense orthography. These results challenge existing models of orthographic processing in reading, and support the claim that orthographic representations are shaped by the nature of the writing system.

There is a broad consensus that printed words in alphabetic languages are recognized through the analysis of letters. Information about letter identity helps readers to distinguish words like SLAT and SPAT that differ by a single letter, while information about letter position permits readers to distinguish anagrams like SLAT and SALT that consist of the same letters in different positions. The nature of position coding in visual word recognition has become a point of major theoretical debate over the past decade (e.g. Davis, 2010; Grainger & Whitney, 2004; Gomez, Ratcliff & Perea, 2008).

Substantial evidence suggests that readers of Indo-European languages are tolerant of transposed letters in word identification (e.g., ‘jugde’ activates ‘judge’; Perea & Lupker, 2003). In standard visual lexical decision, nonwords that are transposed-letter anagrams of words (e.g. silimar) are harder to reject than nonwords that are not (e.g. sitinar; Andrews, 1996; Chambers, 1979; Lupker, Perea & Davis, 2008; Perea & Lupker, 2004). Similarly, masked priming studies show that recognition of a target word is speeded by prior presentation of a transposed-letter prime (e.g. sevrice-SERVICE), relative to a substitution prime (e.g. sedlice-SERVICE; Schoonbaert & Grainger, 2004). This *transposed-letter effect* extends to cases in which the transposition crosses a syllable boundary (e.g. caniso-CASINO versus caviro-CASINO; Perea & Lupker, 2003) and to more extreme modifications (e.g. snawdcih-SANDWICH versus skuvgpah-SANDWICH; Guerrera & Forster, 2008). These findings all suggest that there is a high degree of perceptual similarity between stimuli that comprise the same letters in different positions.

These results highlight a fundamental problem in word recognition. Clearly, we can distinguish *snawdcih* and *sandwich*, so letters must be coded for position.

However, this coding must comprise some degree of uncertainty or flexibility; otherwise, these stimuli would not be treated as perceptually similar. This insight has inspired a variety of competing theories that propose to solve this problem, including the SOLAR model (Davis, 2010), the Open Bigram model (Grainger & Whitney, 2004), the Noisy Channel model (Norris & Kinoshita, 2012), and the Overlap model (Gomez et al., 2008). Though these models have important differences, they all assert that letter position is coded in a way that leads to perceptual uncertainty. Further, uncertainty in letter position coding is argued to be a general property of the cognitive system (Perea & Carreiras, 2012), and caused by low-level visual (e.g. crowding, acuity; Grainger, Dufau & Ziegler, 2016) and neurobiological factors (e.g. noisy retinotopic firing, nature of the receptive field structure; e.g. Dehaene et al., 2005).

However, recent evidence suggests that letter position uncertainty does not extend to all writing systems. In a series of studies in Hebrew, Velan and Frost (2007, 2011) showed that word recognition is *not* facilitated by prior presentation of a transposed-letter prime relative to a substitution control. Frost (2012) argued that the reason for this can be traced to properties of the writing system. Specifically, Hebrew is very dense orthographically, with many anagrams. Hebrew readers must therefore develop precise orthographic position coding, as tolerance to disruptions of letter order would often result in accessing the meaning of the wrong word. Evidence for precise orthographic representations has also been provided in Korean (Lee & Taft, 2011; Rastle, Lally & Lee, 2019) – another language with a dense orthography, but which otherwise shares little similarity with Hebrew. Frost (2012) emphasized that reading is a learned skill, and that while this process will necessarily be constrained by low-level visual and neurobiological processes, flexibility will emerge only where it maximises the efficiency of word recognition. This conclusion is supported by

simulations showing that distributed-connectionist networks trained on artificial languages yield more flexible position coding for sparse orthographies compared to dense orthographies (Lerner, Armstrong & Frost, 2014).

Though Frost (2012) presents a compelling argument that orthographic density is a major constraint on letter-position coding, it is difficult to draw this conclusion definitively from cross-linguistic comparisons since there are substantial differences across languages over and above orthographic density. Hebrew is characterised by a non-concatenative morphological system comprising tri-consonantal roots, which modify properties of the verb such as person, gender and tense. Similarly, Korean is characterised by physically-demarcated syllable blocks with a rigid consonant-vowel-consonant structure. In addition, readers of these languages almost certainly differ in a myriad of ways (e.g. method of reading instruction, language and reading experience). In light of these differences, it is difficult to draw strong conclusions about the specific impact of density on the development of orthographic representations.

Our work brings a new dimension to this debate by using an innovative approach that has the potential to reveal how flexibility in position coding is influenced by specific properties of writing systems. We use a laboratory analogue of reading acquisition in which adults are trained on novel words in unfamiliar scripts (Taylor et al., 2011, 2017). This approach allows precise control over *what participants learn* and *how they learn* in a way that could never be achieved using natural language comparisons. We trained participants on novel words from artificial writing systems designed to be orthographically sparse or dense, but which otherwise were *identical* in factors relevant to word perception (e.g. syllable structure, morphological structure, positional letter frequency). We then used the transposed-letter effect to assess the precision of participants' emerging letter position coding. On

the basis of Frost (2012), we predicted that participants who had learned the orthographically dense writing system would show a smaller transposed-letter effect, indicating greater precision in letter position coding, than those trained on the sparse writing system.

### **Method**

#### Participants.

Forty-eight monolingual English speakers completed the experiment at Royal Holloway University of London, in exchange for £60. All participants were aged 18-25 years old and had no history of language or reading difficulties. Participants were assigned to one of the two writing systems.

#### Stimuli.

**Trained items.** Two artificial writing systems were constructed, each comprising 24 pseudowords printed in an unfamiliar script. In both writing systems, each novel word consisted of five letters and two syllables, and had a CVCVC structure. These novel words were constructed from 17 letters (12 consonants, 5 vowels), and the spelling-to-sound relationship in both languages was consistent, i.e. each letter had one sound. Critically, both the overall frequency and positional frequency of individual symbols was equated across writing systems, with consonants appearing 6 times and vowels appearing 8-10 times in the trained novel words. However, one writing system was sparse (i.e. no anagrams) while the other writing system was dense (i.e. each word was an anagram of another word in the orthography, created by switching the initial and final consonant or by switching the initial and middle consonant). Figure 1 presents an example of the trained stimuli from the

sparse and dense writing systems and their pronunciations. A full list of stimuli can be found in the Open Science Framework storage for this project.

--Insert Figure 1 about here --

Test items. In addition to the trained items, test tasks (conducted on the fifth day) required development of five sets of 24 untrained novel words for each writing system. Untrained words all comprised the same CVCVC structure as trained words, and each set was group-wise matched to trained words on letter frequency. The first four sets of untrained words were created for the visual lexical decision test task. The first set comprised novel words that transposed the second and third consonants of a trained word (TL-C), while the second set comprised novel words that replaced the second and third consonants of a trained word with different consonants from the alphabet (RL-C). The third set comprised novel words that transposed the first and second vowels of a trained word (TL-V), while the fourth set comprised novel words that replaced the first and second vowels with different vowels from the alphabet (RL-V).<sup>1</sup> The fifth set of untrained words was used to assess generalisation performance in reading aloud.

In designing the stimuli, we took great care to make sure that the similarity between test stimuli and trained stimuli was equivalent across sparse and dense orthographies. We used the Match Calculator (Davis, 1999) to assess the degree of similarity between trained and test items on a number of different input coding schemes. Each comparison generated a number between 0 and 1, where 0 indicated

---

<sup>1</sup> We used non-adjacent transpositions of consonants and vowels for the reason that in our alphabets, the symbols associated with consonants and vowels only occur in certain positions (e.g. vowel symbols do not occur in the third position). We had no predictions about consonant-vowel status on the transposed-letter effect, and note that this comparison in any case is confounded with position of disrupted letters.

total dissimilarity and 1 indicated a perfect match. It is evident from the average match scores provided in Table 1 that there were no differences in trained–test item similarity across the two orthographies. This tight control was essential so that any differences in lexical decision performance could be attributed to orthographic density, rather than low-level differences in discrimination difficulty across sparse and dense orthographies as a result of higher orthographic overlap with trained and untrained items.

--Insert Table 1 about here --

Procedure.

Each participant was trained on the novel words from one writing system over four days and tested on the fifth day. The correct response was given as feedback on each trial for training tasks; no feedback was given on test tasks.

During Day 1, participants completed three tasks, with each task comprising three runs. The first task was phonic training. For two runs, participants were exposed to individual letters and their sounds and asked to repeat each sound aloud. In the third run, participants were presented with the letter and had to produce the sound. The second task was reading aloud; participants saw each novel word and were asked to read it aloud. The third task was orthographic search; participants heard a novel word and selected its visual form from a grid of all 24 novel words. During Days 2-4, participants completed three blocks of training each day. Blocks consisted of three repetitions of reading aloud and one repetition of orthographic search. Training on each day took approximately 75 minutes.

On the fifth day, participants completed four test tasks in a fixed order. These included tasks similar to the reading aloud and orthographic search tasks practiced in



training; however, each stimulus was presented once per task, and participants received no feedback on the correct response. In addition, participants completed visual lexical decision and generalisation. In the lexical decision task, participants were presented with letter strings and asked to decide whether they were words that they had learned.<sup>2</sup> The letter strings included trained words, and the four sets of untrained novel words (TL-C, RL-C, TL-V, RL-V). Trained words were repeated four times, so that ‘yes’ and ‘no’ responses were balanced. Trained items were included as fillers in order to provide a correct ‘yes’ response, and also to measure participants’ overall recognition of trained items. Untrained items were included to measure the transposed letter effect (shown by the difference in performance for TL and RL foils), reflecting the degree of position uncertainty in each orthography. In the generalisation task, participants were asked to read the fifth set of untrained novel words aloud. This allowed us to assess the extent to which participants had extracted underlying spelling–sound regularities from training on the novel words.

## Results

Data from one participant were removed from all analyses due to poor learning of the trained items (63% correct on reading aloud test; 49% correct on ‘yes’ response in lexical decision test). Data were analysed using analyses of variance (ANOVA) on accuracy and response times (RTs), although we note that previous studies in which adults have learned to read in an artificial script have typically focused only on accuracy (e.g. Taylor et al., 2011, 2017). Spoken responses were

---

<sup>2</sup> We chose to investigate transposed letter phenomena using standard lexical decision rather than masked priming because we judged that this would be more suitable for use with an artificial orthography training paradigm. Though there is ample evidence that participants can discriminate trained from untrained stimuli in such paradigms (e.g. Taylor et al., 2017), we are unaware of any evidence suggesting that trained items would yield masked repetition priming effects.

hand-marked for accuracy and RT by a research assistant naïve to the purpose of the study using CheckVocal software (Protopapas, 2007). Analyses were conducted on by-subject (F1) and by-item (F2) means. Results were interpreted as significant when effects held across both F1 and F2 analyses. Data and analysis scripts are available in the OSF storage for this project.

Training Data (Days 1-4).

*Phonic Training (Day 1).* The analysis of phonic training data considered performance in the third run of phonic training, and included Orthography (sparse vs dense) as a factor. The analysis of accuracy data revealed no difference between sparse ( $M = 0.47$ ,  $SE = 0.04$ ) and dense ( $M = 0.42$ ,  $SE = 0.04$ ) writing systems,  $F_1(1, 45) = 0.62$ ,  $p = 0.43$ ;  $F_2(1, 32) = 0.71$ ,  $p = 0.40$ . Similarly, there was no difference in RTs between sparse ( $M = 2177$ ms), and dense ( $M = 1957$ ms) writing systems,  $F_1(1, 45) = 1.37$ ,  $p = 0.25$ ;  $F_2(1, 32) = 1.76$ ,  $p = 0.19$ . These data provide confidence that there were no initial differences between the language groups on ability to learn the artificial alphabets.

*Reading Aloud (Days 1-4).* The analysis of reading aloud training data considered Orthography (sparse vs dense) and Day as factors. Figure 2 provides a visual representation of the data.

For accuracy, there was a main effect of Day,  $F_1(3, 135) = 212.53$ ,  $p < .001$ ;  $F_2(3, 138) = 1531.50$ ,  $p < .001$ , with performance becoming more accurate over time. Although Figure 2 suggests slightly higher accuracy for the dense group, neither the effect of Orthography,  $F_1(1, 45) = 3.17$ ,  $p = .08$ ;  $F_2(1, 46) = 16.29$ ,  $p < .001$ , nor the interaction between Day and Orthography,  $F_1(3, 135) = 0.98$ ,  $p = .41$ ;  $F_2(3, 138) = 6.14$ ,  $p < .001$ , was reliable across by-subject and by-item analyses.

For RTs, there was a main effect of Day,  $F_1(3, 135) = 150.42, p < .001$ ;  $F_2(3, 138) = 881.99, p < .001$ , with faster responses emerging over time. The RT data showed no effect of Orthography,  $F_1(1, 45) = 0.24, p = .63$ ;  $F_2(1, 46) = 3.53, p = .07$ , and no interaction between Day and Orthography,  $F_1(1, 135) = 0.28, p = .84$ ;  $F_2(1, 138) = 5.09, p < .01$ .

-- Insert Figure 2 about here --

*Orthographic Search (Days 1-4)*. The analysis of orthographic search training data considered Orthography (sparse vs dense) and Day as factors. Figure 3 provides a visual representation of the data. The accuracy analysis revealed an effect of Day,  $F_1(3, 133) = 16.62, p < .001$ ;  $F_2(3, 138) = 105.26, p < .001$ , as accuracy increased over time. Although Figure 3 again suggests slightly higher accuracy for the dense group, there was no effect of Orthography,  $F_1(1, 43) = 0.54, p = .47$ ;  $F_2(1, 46) = 6.95, p < .05$ , and no interaction between Day and Orthography,  $F_1(3, 133) = 0.49, p = .69$ ;  $F_2(1, 138) = 3.40, p < .05$ , that was reliable across by-subject and by-item analyses.

For RTs, there was a main effect of Day,  $F_1(3, 133) = 100.50, p < .001$ ;  $F_2(3, 138) = 391.89, p < .001$ , as latencies decreased over time. The RT data showed no effect of Orthography,  $F_1(1, 43) = 0.65, p = 0.43$ ;  $F_2(1, 46) = 1.16, p = 0.29$ , and no interaction between Day and Orthography,  $F_1(3, 133) = 0.77, p = 0.51$ ;  $F_2(3, 138) = 2.12, p = 0.10$ .

--Insert Figure 3 about here --

Overall, training data suggest that trained words were learned to a high degree of accuracy, with no reliable differences across sparse and dense orthographies.<sup>3</sup>

---

<sup>3</sup> We note that there was some indication from by-item analyses that the dense orthography may have been easier to learn for some of the participants (although

Testing Data (Day 5).

*Reading aloud.* The analysis of reading aloud test data included Orthography (sparse vs dense) and Lexical Status (trained vs untrained) as factors. Figure 4 provides a visual representation of the data.

The analysis of accuracy revealed a significant effect of Lexical Status,  $F_1(1, 45) = 139.97, p < .001$ ;  $F_2(1, 92) = 450.67, p < .001$ , with trained items read aloud more accurately than untrained items. There was also a significant effect of Orthography,  $F_1(1, 45) = 12.17, p < .01$ ;  $F_2(1, 92) = 50.96, p < .001$ , with higher accuracy in the dense orthography. However, these main effects were qualified by an interaction,  $F_1(1, 45) = 11.63, p < .01$ ;  $F_2(1, 92) = 37.02, p < .001$ . This interaction revealed that whilst performance on trained items did not differ as a function of Orthography,  $F_1(1, 45) = 0.75, p = .39$ ;  $F_2(1, 46) = 1.54, p = .22$ , performance on untrained items was more accurate for the dense than the sparse orthography,  $F_1(1, 45) = 13.19, p < .001$ ;  $F_2(1, 46) = 53.32, p < .001$ .

The analysis of RT revealed an effect of Lexical Status,  $F_1(1, 44) = 160.38, p < .001$ ,  $F_2(1, 92) = 331.67, p < .001$ , with longer latencies for untrained than trained items. However, there was no effect of Orthography,  $F_1(1, 44) = 0.03, p = .87$ ;  $F_2(1, 92) = 0.001, p = 0.98$ , and no interaction between these factors,  $F_1(1, 44) = 0.01, p = .94$ ;  $F_2(1, 92) = 0.20, p = 0.66$ .

--Insert Figure 4 about here--

---

performance converged by the end of training). This may suggest that there are meaningful individual differences in how these types of writing systems are learned. Future higher-powered studies may wish to investigate this possibility.

*Orthographic search.* The analysis of orthographic search test data included Orthography (sparse vs dense) as a factor. The analysis of accuracy revealed no significant difference between sparse ( $M = 0.97$ ,  $SE = 0.07$ ) and dense ( $M = 0.98$ ,  $SE = 0.05$ ) orthographies,  $F_1(1, 45) = 0.17$ ,  $p = .69$ ;  $F_2(1, 46) = 0.47$ ,  $p = 0.50$ ). Similarly, the analysis of RT revealed no significant difference between sparse ( $M=6851$ ) and dense ( $M=7131$ ) orthographies,  $F_1(1, 45) = 0.28$ ,  $p = .60$ ;  $F_2(1, 46) = 0.61$ ,  $p = 0.44$ .

*Lexical decision.* Analysis of the ‘YES’ response included Orthography (sparse vs dense) as a factor. The analysis of accuracy revealed no difference in recognition of targets learned in sparse ( $M = 0.95$ ,  $SE = 0.01$ ) and dense ( $M = 0.94$ ,  $SE = 0.02$ ) orthographies,  $F_1(1, 45) = 0.21$ ,  $p = 0.65$ ;  $F_2(1, 46) = 1.07$ ,  $p = 0.31$ . Similarly, the analysis of RT revealed no difference in the speed with which targets learned in sparse ( $M = 3502\text{ms}$ ) and dense ( $M = 3635\text{ms}$ ) orthographies were accepted,  $F_1(1, 45) = 0.14$ ,  $p = 0.72$ ;  $F_2(1, 46) = 0.83$ ,  $p = 0.37$ .

Analysis of the ‘NO’ response included Orthography and TL status (TL vs RL) as factors. Figure 5 provides a visual representation of the data. The analysis of accuracy revealed an impact of TL status, with lower accuracy in rejecting TL foils than RL foils,  $F_1(1, 45) = 50.29$ ,  $p < .001$ ;  $F_2(1, 46) = 31.29$ ,  $p < .001$ . There was also a main effect of Orthography,  $F_1(1, 45) = 5.63$ ,  $p < .05$ ;  $F_2(1, 46) = 58.72$ ,  $p < .001$ , with accuracy in the dense orthography higher than in the sparse orthography. Critically, however, these main effects were qualified by a significant interaction,  $F_1(1, 45) = 8.33$ ,  $p < .01$ ;  $F_2(1, 46) = 5.09$ ,  $p < .05$ , which indicated a larger TL effect in the sparse orthography than in the dense orthography.

The analysis of RT revealed no effect of TL status,  $F_1(1, 45) = 8.11$ ,  $p < .01$ ;  $F_2(1, 46) = 3.32$ ,  $p = .07$ , no effect of Orthography,  $F_1(1, 45) = 0.001$ ,  $p = 0.98$ ;  $F_2(1,$

46) = 0.001,  $p = 0.99$ , and no interaction between TL status and Orthography,  $F_1(1, 45) = 0.41$ ,  $p = 0.53$ ;  $F_2(1, 46) = 0.09$ ,  $p = 0.77$ .

-- Insert Figure 5 about here --

## Discussion

Substantial research suggests that letter position is represented flexibly in skilled reading (e.g. Perea & Lupker, 2004; Schoonbaert & Grainger, 2004). However, recent research in Hebrew (Frost, 2012; Velan & Frost, 2011) and Korean (Rastle et al., 2019) suggests that this may not be a universal property of reading, but rather may depend on the orthographic density of a writing system. We sought to investigate the impact of orthographic density on the emergence of letter position coding using an artificial language learning paradigm. Over four days, participants learned to read novel words printed in an artificial orthography that was sparse (no anagrams) or dense (many anagrams). On the fifth day, they were tested in a variety of ways for their knowledge of the artificial orthographies. We assessed the precision of letter position coding through a lexical decision task, in which participants were required to accept trained words but to reject transposed-letter and replaced-letter foils. We took the size of the transposed-letter effect on rejection decisions as an index of flexibility in position coding (e.g. Andrews, 1996), and expected this to be larger in the sparse orthography than in the dense orthography.

Results revealed the predicted difference in the size of the transposed-letter effect on rejection decisions across sparse and dense orthographies. Though participants across the two writing systems learned trained words to the same high degree of accuracy, the underpinning orthographic representations clearly differed.

Critically, participants who learned the sparse orthography were more likely to accept the transposed-letter foils as trained words (relative to a replaced-letter control) than participants who learned the dense orthography. This result indicates that participants' emerging orthographic representations were more precisely coded for letter position when they learned to read the dense orthography than the sparse orthography. We note that these findings arose on accuracy rather than RT. It is not surprising that findings should be confined to accuracy given the low level of experience that participants had with the novel alphabets. Indeed, the fact that 'no' decisions in the lexical decision task hovered around 5000 ms suggests that reading of these alphabets was not fully automatized. The critical point is that there is no evidence of a speed-for-accuracy trade-off that would undermine the result on accuracy. If anything, the RT data go in the same direction as the accuracy data (i.e. larger transposed-letter effect in the sparse orthography).

These results are consistent with previous cross-linguistic studies demonstrating reductions in transposed-letter effects in orthographically-dense scripts such as Hebrew (Velan & Frost, 2007, 2011) and Korean (Lee & Taft, 2011; Rastle et al., 2019). However, our findings are particularly powerful because the impact of orthographic density on letter position coding cannot be attributed to other confounding language characteristics or to variations in participant groups across languages. These results support Frost's (2012) claim that the flexibility of letter position coding in reading arises as a consequence of the statistical structure of a writing system. However, a deeper question relates to how theories of reading acquisition might account for the impact of orthographic density on flexibility of letter position coding.

Several theories of reading acquisition highlight the linguistic environment as a key factor in forming optimal word representations. The amalgamation theory (Ehri & Wilce, 1980) and the lexical tuning hypothesis (Castles et al, 2001) both propose that readers develop more precise representations of words with a high neighbourhood density, due to the increased risk of confusability. This prediction has been supported in masked-priming studies showing that words in dense neighbourhoods show reduced substituted letter priming and transposed letter priming than words in sparser neighbourhoods (Castles et al., 2007; Forster et al., 1987; Kinoshita, Castles & Davis, 2009; Perea & Rosa, 2000). Our work suggests that the proposals of these theories regarding flexible tuning within a language might also be invoked to understand cross-linguistic differences. Readers of dense orthographies may require more precise tuning of word representations than readers of sparse orthographies, resulting in lower tolerance to transpositions.

Similarly, while our findings are inconsistent with the proposal that letter position flexibility arises solely as a result of low-level visual or neurobiological phenomena, we can envisage ways in which these theories might accommodate an influence of orthographic density. For example, the local combination detector model (Dehaene et al, 2005) proposes that detector sizes are larger for writing systems in which the reader is reliant on larger orthographic units (e.g. languages with low grapheme-phoneme transparency). This proposal offers a potential way forward for thinking about the impact of orthographic density on position flexibility, as in dense orthographies the reader may need to consider positional information from a larger window of letters in the word in order to reliably differentiate between anagrams.

However, we believe that the full range of the results observed are most compatible with the dual-pathway model of Grainger and Ziegler (2011). This model



proposes that skilled readers use coarse- and fine-grained codes in parallel in order to decode written words. The coarse-grained route identifies letter combinations in the absence of precise positional information to provide a fast-track to semantic information. In contrast, the fine-grained route is more sensitive to the precise ordering of letters. The precision of orthographic information along the fine-grained pathway permits mapping onto phonological information as well as chunking of frequently-occurring contiguous letter combinations, such as morphemes. It seems plausible that during reading acquisition, learned representations of words are tuned to reflect an optimal balance of coarse-grained and fine-grained processing. If so, then readers of dense orthographies may be less able to utilise coarse-grained information, as the lack of position specificity would be inefficient for identifying words with many orthographic neighbours. Rather, they would need to develop greater reliance on the fine-grained pathway. In contrast, readers of sparse orthographies with few orthographic neighbours would have more weight assigned to less precise representations as there is a much lower chance of identifying a transposed-letter neighbour in error. The reliance on less precise representations in orthographies with fewer orthographic neighbours would result in larger transposed-letter effects in sparse orthographies, as observed in the current work.

This account suggests that reading acquisition is characterised by a process of learning the degree of precision that is required for efficient word recognition. The optimal degree of precision may vary locally across different types of words, and may vary cross-linguistically based on orthographic density, as the present results suggest. This interpretation is supported by research suggesting that the magnitude of the transposed-letter effect *increases* through the period of reading acquisition (Ziegler et al., 2014 in French; Colombo, Sulpizio & Peressotti, 2017 in Italian). This evidence

stands in contrast to the predictions of the lexical quality hypothesis (Perfetti, 2007), stating that the process of reading acquisition is characterised by increased fine-tuning of representations (i.e. greater precision) through the accumulation of print experience.

One problem with this account based on the dual-pathway model (Grainger & Ziegler, 2011) is that seems to allow too many degrees of freedom. That is, one might argue that the model allows the researcher to explain any number of effects simply by suggesting *posthoc* that coarse-grained or fine-grained processing dominated. The account would be more persuasive if we had additional, independent evidence that participants in our dense orthography condition were more reliant on fine-grained processing. Remarkably, data from the generalisation test task provides this independent evidence. Results indicated that the trained words were learned to the same high degree of accuracy across writing systems. Yet, when participants were asked to read aloud untrained words using the same symbols, participants who learned the dense script showed a substantial advantage. This suggests that participants who learned the dense orthography developed more componential representations, reflecting greater fine-grained letter-to-sound knowledge, than participants who learned the sparse orthography. Once again, the evidence indicates that the nature of the writing system impacted on how the words were learned.

The introduction of the artificial orthography training paradigm has allowed us to study the unique impact of orthographic density on the acquisition of orthographic representations. Due to associations between orthographic density and other factors in existing writing systems, this type of highly-controlled study is only possible in a simulated environment. However, there are clearly limitations of these paradigms, introduced largely due to constraints on what participants are able to learn over a

reasonable time period. Further, we have simplified our vocabularies in many ways to facilitate the learning task (e.g. use of a strict syllabic structure for all items), and to ensure perfect matching across orthographies. These simplifications may have had unintended consequences. For example, while participants across orthographies differed substantially in their treatment of *untrained* items in the lexical decision and reading aloud tasks, we observed no differences across orthographies in the speed or accuracy with which *trained* items were processed. We believe that the data from untrained items indicates that the writing systems were learned in different ways, but we would not like to speculate that writing system has no bearing on the speed or accuracy with which words are processed once learned. It may be that the null effect of orthography on the processing of trained items reflected the very tight, artificial matching across orthographies, or that our tasks were insufficiently sensitive to detect effects on trained items (see also Footnote 3). These arguments suggest that while artificial language studies of this nature form an important part of the evidence base, they must be interpreted as complementary to studies of existing languages and writing systems.

Overall, our results provide a strong demonstration of the impact of the orthographic density of a writing system on the precision of orthographic representations. Using an artificial language approach, we varied orthographic density across two artificial writing systems, while controlling all other stimulus and participant factors that confound this comparison in studies using natural languages. Our results challenge existing cognitive and neurobiological models of position coding in reading, and support the argument put forward by Frost (2012) that orthographic representations are shaped by the statistical structure of the writing system one learns to read (see also Lerner et al., 2014). We look forward to using this

method to delineate how the complex associations between orthographic, phonological and semantic information across the world's writing systems shape the acquisition of the reading skill.

### References

- Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language*, *35*(6), 775-800. doi: 10.1006/jmla.1996.0040.
- Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*(3), 165-182. doi: 10.1016/j.jecp.2007.01.006.
- Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning & Verbal Behavior*, *18*(2), 225-241. doi: 10.1016/S0022-5371(79)90136-1.
- Colombo, L., Sulpizio, S., & Peressotti, F. (2017). Serial mechanism in transposed letters effects: A developmental study. *Journal of Experimental Child Psychology*, *161*, 46-62. doi: 10.1016/j.jecp.2017.04.002.
- Davis, C. J. (1999). The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition (Doctoral dissertation). University of New South Wales, Sydney, Australia.
- Davis, C.J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713-758. doi: 10.1037/a0019738.
- Dehaene, S., Cohen, L., Sigman, M. & Vinickier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, *9*(7), 335-341. doi: 10.1016/j.tics.2005.05.004.

- Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, *1*(4), 371-385. doi: 10.1017/S0142716400009802.
- Frost, R. (2012). A universal approach to modelling visual word recognition and reading: Not only possible, but also inevitable. *Behavioural and Brain Sciences*, *35*(5), 310-329. doi: 10.1017/S0140525X12000635.
- Gomez, P., Ratcliff, R. & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*(3), 577-600. doi: 10.1037/a0012667.
- Grainger, J., Dufau, S. & Ziegler, J.C. (2016). A Vision of Reading. *Trends in Cognitive Sciences*, *20*(3), 171-179. doi: 10.1016/j.tics.2015.12.008.
- Grainger, J. & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences*, *8*(2), 58-59. 10.1016/j.tics.2003.11.006.
- Grainger, J., & Ziegler, J. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, *2*, 54. doi: 10.3389/fpsyg.2011.00054.
- Guerrera, C. & Forster, K. (2008) Masked form priming with extreme transposition. *Language and Cognitive Processes*, *23*(1), 117-142. doi: 10.1080/01690960701579722.
- Kinoshita, S., Castles, A., & Davis, C. (2009). The role of neighbourhood density in transposed-letter priming. *Language and Cognitive Processes*, *24*(4), 506-526. doi: 10.1080/01690960802289213.

- Lee, C. H., Kwon, Y., Kim, K., & Rastle, K. (2015). Syllable transposition effects in Korean word recognition. *Journal of Psycholinguistic Research*, *44*(3), 309-315. doi: 10.1007/s10936-015-9353-7.
- Lee & Taft, (2011). Subsyllabic structure reflected in letter confusability effects in Korean word recognition. *Psychonomic Bulletin & Review*, *18*(1), 129-134. doi: 10.3758/s13423-010-0028-y.
- Lerner, I., Armstrong, B.C. & Frost, R. (2014). What can we learn from learning models about sensitivity to letter-order in visual word recognition? *Journal of Memory and Language*, *77*, 40-58. doi: 10.1016/j.jml.2014.09.002
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review*, *1*(4), 476–490. doi: 10.3758/BF03210951.
- Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed-letter effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes*, *23*(1), 93-116. doi: 10.1080/01690960701579714.
- Norris, D. & Kinoshita, S. (2012). Reading through a noisy channel: why there's nothing special about the perception of orthography. *Psychological Review*, *119*(3), 517-545. doi: 10.1037/a0028450.
- Perea, M., & Carreiras, M. (2012). Perceptual uncertainty is a property of the cognitive system. *Behavioral and Brain Sciences*, *35*(5), 298-299. doi: 10.1017/S0140525X12000118.

- Perea, M. & Lupker, S.J. (2003). Does judge activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition*, *31*(6), 829-841. doi: 10.3758/BF03196438.
- Perea, M. & Lupker, S.J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*(2), 231–246. doi.org/10.1080/01690960701579714.
- Perea, M., & Rosa, E. (2000). Repetition and form priming interact with neighborhood density at a brief stimulus onset asynchrony. *Psychonomic Bulletin & Review*, *7*(4), 668-677. doi: 10.3758/BF03213005.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383. doi: 10.1080/10888430701530730.
- Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859-862.
- Rastle, K., Lally, C., & Lee, C. H. (2019). No flexibility in letter position coding in Korean. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(4), 458-473. doi: 10.1037/xhp0000617.
- Schoonbaert, S. & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*(3), 333-367. doi: 10.1080/01690960344000198.
- Taylor, J.S.H., Plunkett, K. & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: an artificial orthography



paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 60-76. doi: 10.1037/a0020126.

Taylor, J.S.H., Davis, M.H. & Rastle, K. (2017). Comparing and validating methods of reading instruction using behavioural and neural findings in an artificial orthography. *Journal of Experimental Psychology: General*, 146(6), 826-858. doi: 10.1037/xge0000301

Velan, H. & Frost, R. (2007). Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review*, 14(5), 913-918. doi: 10.3758/BF03194121.

Velan, H. & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition*, 118(2), 141-156. doi: 10.1016/j.cognition.2010.11.013.

Ziegler, J. C., Bertrand, D., Lété, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology*, 50(4), 1026-1036. doi: 10.1037/a0035187.

**Author Note**

This work was supported by Global Research Network program through the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2017S1A2A2035130). The materials, data, and analysis scripts for this project can be found on <https://osf.io/g74vp/>.

**Table 1**

Match Calculator (Davis, 1999) statistics displaying mean orthographic overlap between trained and untrained items.

<b>Orthography</b>	<b>Absolute</b>	<b>SOLAR (Spatial Coding)</b>	<b>Overlap Open Bigram</b>	<b>SERIAL Open Bigram</b>	<b>Binary Open Bigram</b>
<i>Dense</i>	0.13	0.19	0.04	0.04	0.04
<i>Sparse</i>	0.13	0.19	0.04	0.03	0.04

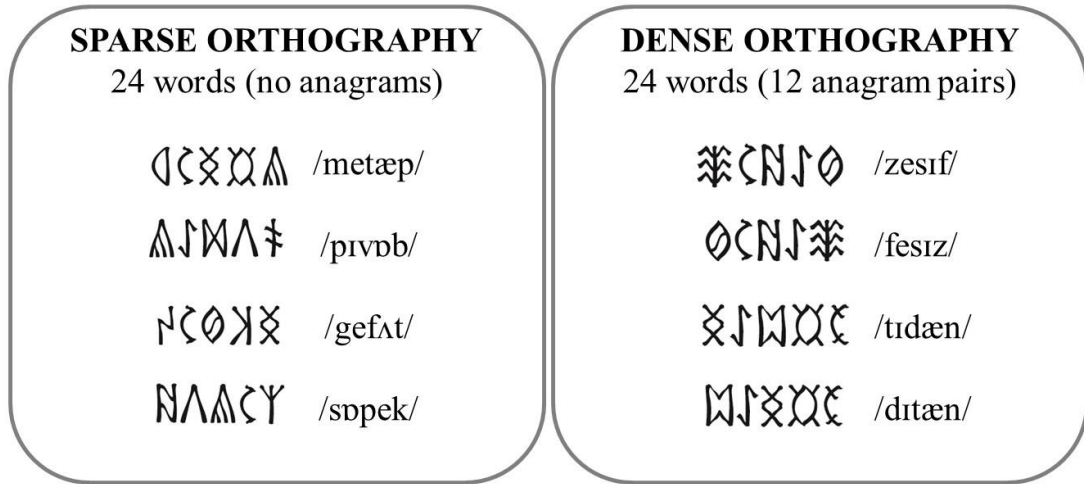
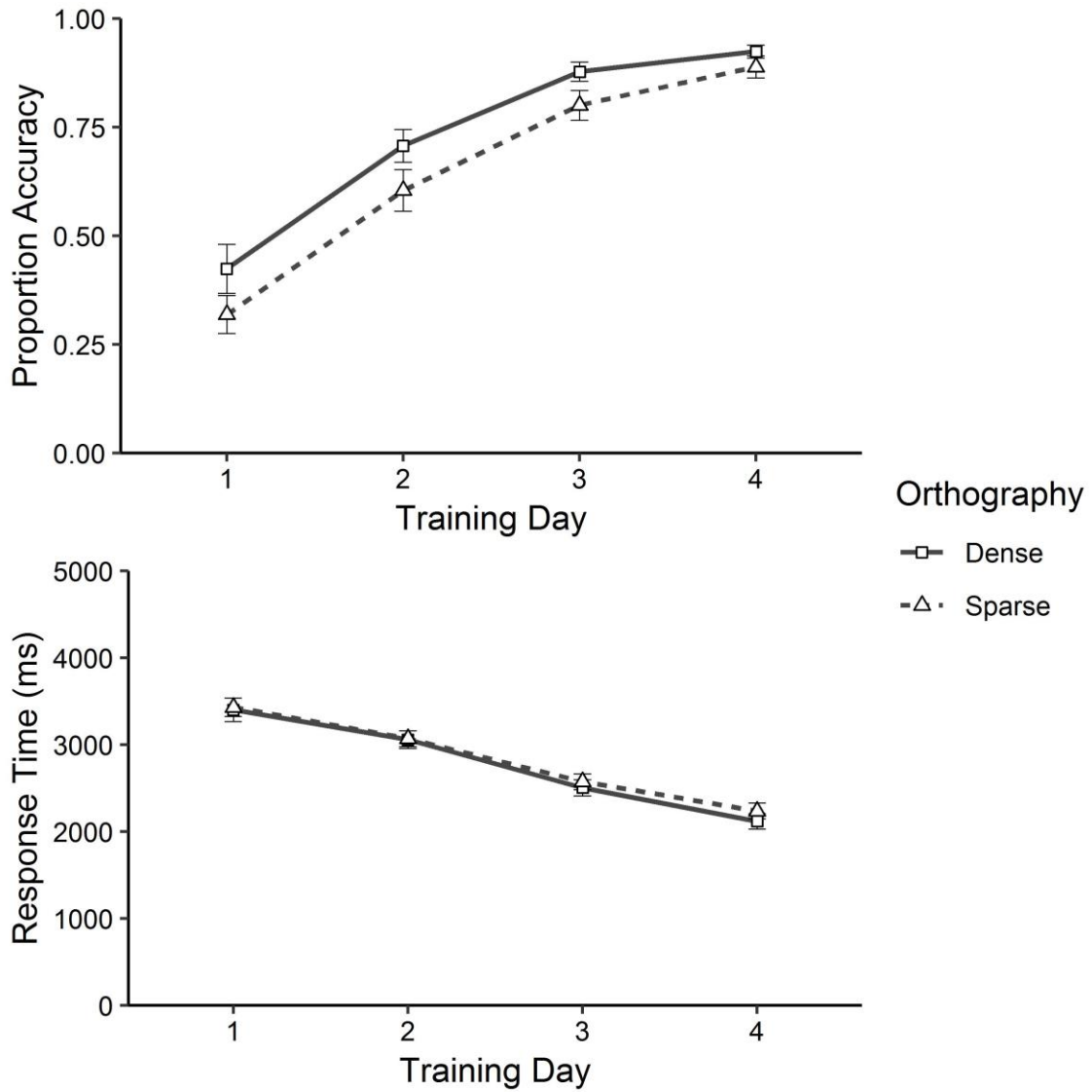
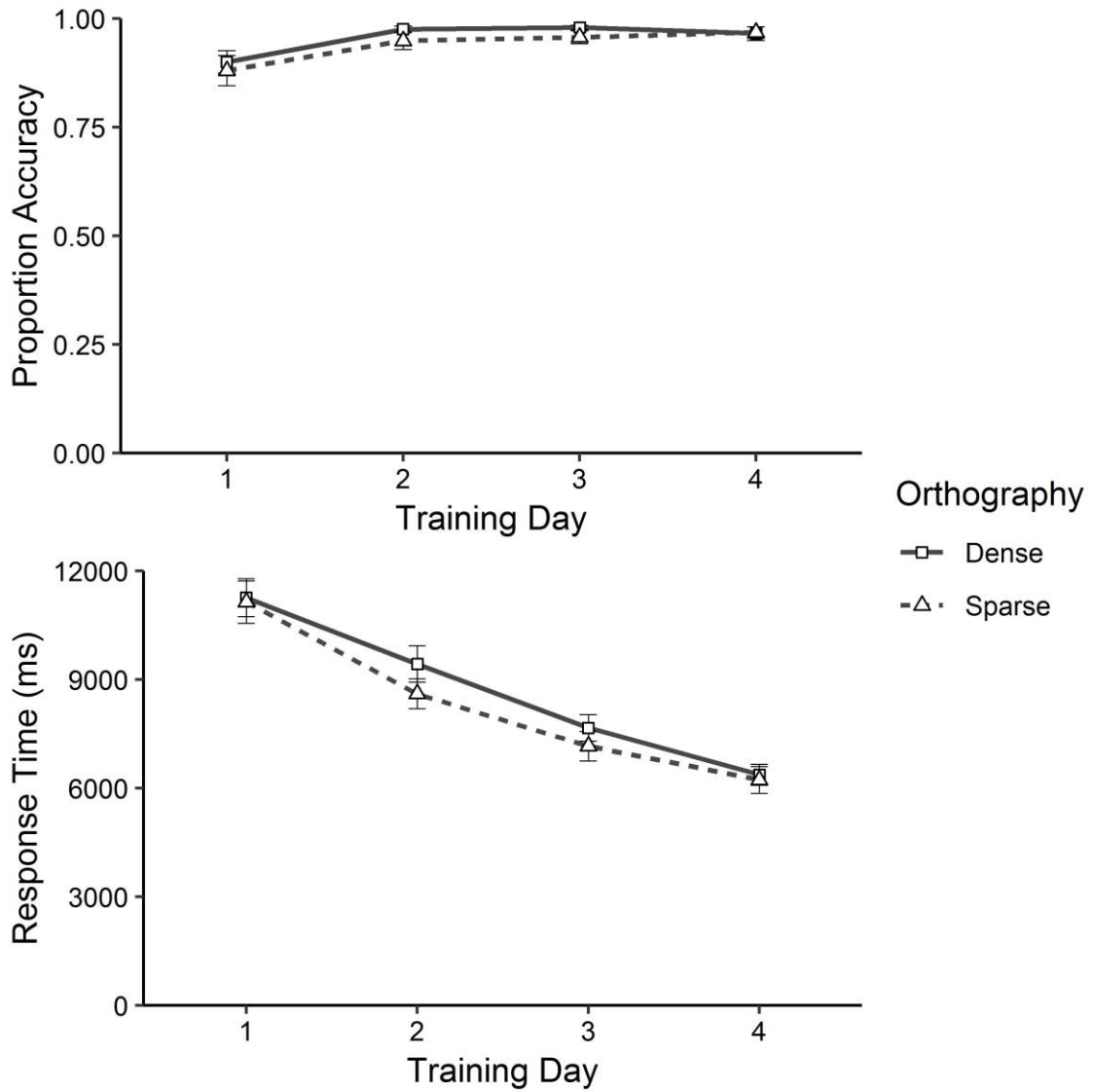


Figure 1. Examples of stimuli in dense and sparse orthographies.



*Figure 2.* Mean accuracy and response times for each day of the reading aloud training task. Error bars display one standard error from the mean, calculated for between-subjects designs. Data are averaged across three repetitions of the task on Day 1 and nine repetitions of the task on Days 2-4.



*Figure 3.* Mean accuracy and response times for each day of the orthographic search training task. Error bars display one standard error from the mean, calculated for between-subjects designs. The data for each day are averaged across three repetitions.

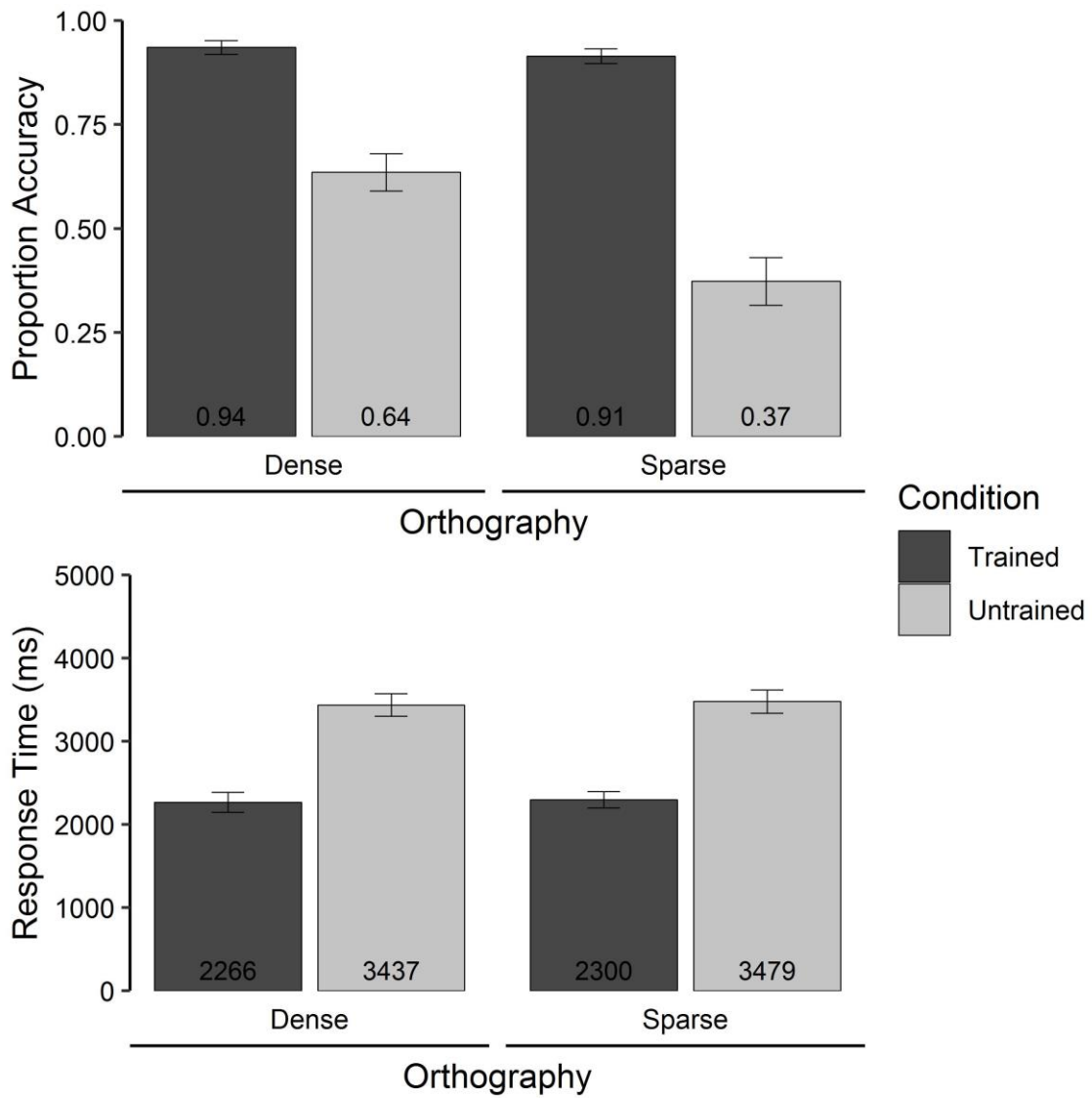


Figure 4. Mean accuracy and response times for reading aloud trained and untrained stimuli on Day 5. Error bars display one standard error from the mean, calculated for between-subjects designs.

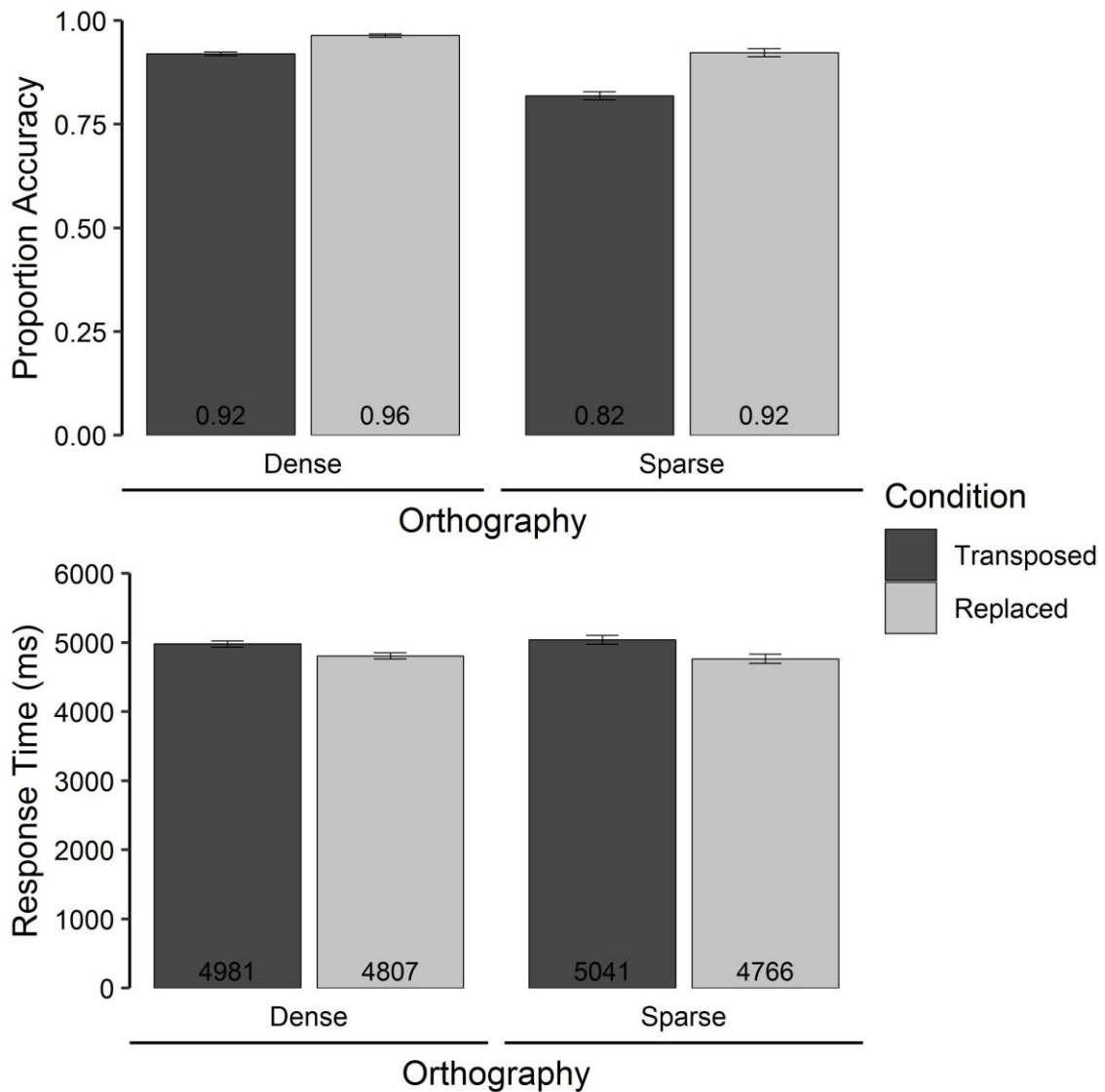


Figure 5. Mean accuracy and response times for the visual lexical decision test task on Day 5. Error bars display one standard error from the mean, calculated for within-subject designs (Loftus & Masson, 1994). Error bars display within-subject variability because the comparison of interest is the size of the transposed-letter effect within each orthography.