the british psychological society
promoting excellence in psychology

British Journal of Psychology (2021), 112, 52–91

www.wileyonlinelibrary.com

Registered report

Does orthographic processing emerge rapidly after learning a new script?

María Fernández-López[1], Ana Marcet[1] and Manuel Perea[1,2,3]*

[1]Universitat de València, Spain
[2]Basque Center on Brain, Cognition, and Language, Donostia, Spain
[3]Universidad Nebrija, Spain

Orthographic processing is characterized by location-invariant and location-specific processing (Grainger, 2018): (1) strings of letters are more vulnerable to transposition effects than the strings of symbols in same-different tasks (location-invariant processing); and (2) strings of letters, but not strings of symbols, show an initial position advantage in target-in-string identification tasks (location-specific processing). To examine the emergence of these two markers of orthographic processing, we conducted a same-different task and a target-in-string identification task with two unfamiliar scripts (pre-training experiments). Across six training sessions, participants learned to fluently read and write one of these scripts. The post-training experiments were parallel to the pre-training experiments. Results showed that the magnitude of the transposed-letter effect in the same-different task and the serial function in the target-in-string identification tasks were remarkably similar for the trained and untrained scripts. Thus, location-invariant and location-specific processing does not emerge rapidly after learning a new script; instead, they may require thorough experience with specific orthographic structures.

Reading is an acquired skill that involves some functional brain changes and requires, in alphabetic scripts, associating the letters that compose each word with their appropriate speech sounds. A common assumption in the literature is that, for a mature word recognition system, the process of identifying words comprises a series of stages that map the visual input onto abstract letter representations and, subsequently, onto whole-word representations (see Dehaene, Cohen, Sigman, & Vinckier, 2005; Grainger, 2008; but see Price & Devlin, 2011, for an alternative account). The processing of orthographic representations connects the low-level stages of visual processing to the higher-level linguistic processing of words. These orthographic representations contain information about the identity and order of each of the word's constituent letters, thus allowing readers to distinguish similarly spelled words like *cure* and *core,* which differ in the identity of just one letter, or words like *slat* and *salt*, which differ in the order of two of the letters (see Grainger, 2018, for review). Indeed, the question of how the brain encodes the identity and order of the letters that constitute each word is a central issue for all leading

*Correspondence should be addressed to Manuel Perea, Universitat de València, Av Blasco Ibañez, 21, 46010 Valencia, Spain (email: mperea@uv.es).

DOI:10.1111/bjop.12469

models of visual word recognition (e.g., Spatial Coding model: Davis, 2010; Overlap model: Gomez, Ratcliff, & Perea, 2008; Open Bigram model: Grainger & Van Heuven, 2004; Bayesian Reader model: Norris, Kinoshita, & van Casteren, 2010; SERIOL model: Whitney, 2001).

In the present experiments, we examined the emergence of two fundamental markers of orthographic processing after learning a new script: 1) location-invariant processing and 2) location-specific processing (see Grainger, 2018). For simplicity, the above-cited models focus on an extant perspective (i.e., they assume a fully developed word recognition system frozen in time) rather than on a developmental perspective. (We defer a discussion of the research focused on the development rather than the emergence of orthographic representations [e.g., Castles, Davis, Cavalot, & Forster, 2007; Grainger, Lété, Bertrand, Dufau, & Ziegler, 2012; Marinus, Kezilas, Kohnen, Robidoux, & Castles, 2018] until the Discussion section.) We first describe in some depth how location-invariant and location-specific processing differs between letters and other visual objects (e.g., symbols, unknown letters). Then, we describe how acquiring a new script may affect both phenomena. Finally, we offer a rationale for the two experiments proposed in the current paper.

Location-invariant processing refers to the mechanisms responsible for encoding the 'relative positions of a set of object identities' (Grainger, 2018, p. 345) (i.e., the encoding of the order of visual objects [letters] in a string composed of several objects [a word]). This has often been examined with the same-different matching task (see Krueger, 1978; Ratcliff, 1981, for early evidence), as it allows researchers to compare the processing of letters vs. the processing of other types of visual objects. In this task, participants have to decide if two strings of characters presented subsequently are the same or not (see Figure 1). The most studied phenomenon of location-invariant processing is the transposed-letter effect (henceforth, TL effect; see Grainger, 2018, for review). The TL effect refers to the insensitivity of readers to the position of letters compared to the identity of the same letters: 'no' responses to the transposed-letter pair FGJM-FJGM (the underline is to emphasize the manipulation) in a same-different matching task are slower and more error prone than the responses to the replacement-letter control FGJM-FPCM. These effects also occur with strings composed of symbols or unknown letters (e.g. , £§?@−£?§@ is slower and more error prone than £§?@−£#<@), which suggests that there is some positional noise in the representations of visual objects in a string (see Gomez *et al.*, 2008; Norris & Kinoshita, 2012). But the critical finding is that transposition effects are substantially larger for strings of letters than for strings of other visual objects (e.g., numbers, symbols, pseudoletters; Duñabeitia, Dimitropoulou, Grainger, Hernández, & Carreiras, 2012; Massol, Duñabeitia, Carreiras, & Grainger, 2013; see also García-Orza, Perea, & Muñoz, 2010; Muñoz, Perea, García-Orza, & Barber, 2012). To explain the greater transposition effect for strings of letters than for strings of other visual stimuli, Grainger (2018; see also Marcet, Perea, Baciero, & Gomez, 2019; Massol *et al.*, 2013) suggested that, on top of positional noise, there is an orthographic-specific mechanism used to encode location-invariant letter-in-word order. Critically, this orthographic-specific mechanism has been posited to emerge with literacy acquisition (Dandurand, Grainger, & Dufau, 2010; Duñabeitia, Lallier, Paz-Alonso, & Carreiras, 2015; Duñabeitia, Orihuela, & Carreiras, 2014). Therefore, when learning a new script, the emergence of location-invariant orthographic processes would produce an increase of letter transposition effects in a same-different task.

Location-specific processing of visual information refers to the parallel processing of the position of characters (e.g., letters) within one object (e.g., a word). This type of

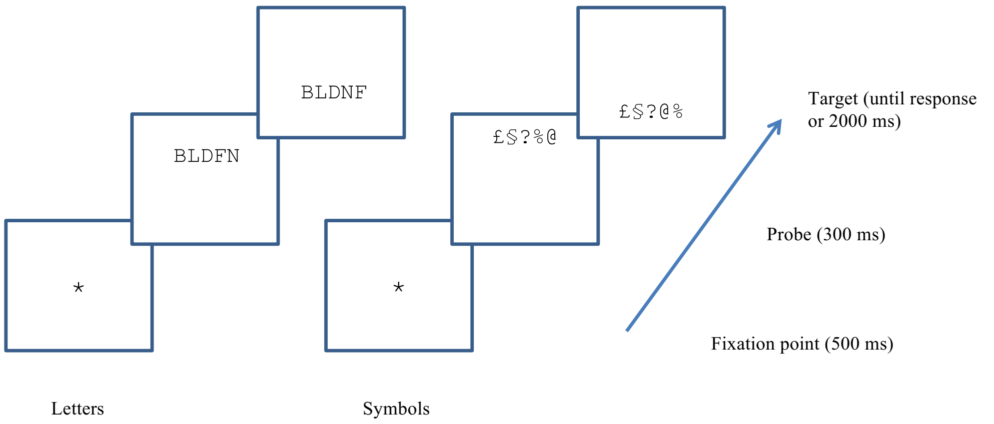

Figure 1. Depiction of the same-different task. [Colour figure can be viewed at wileyonlinelibrary.com]

processing has usually been examined with a post-cued partial report target-in-string identification task (henceforth, TSI task), based on the two-alternative-forced-choice task first introduced by Reicher (1969) and Wheeler (1970). In the typical set-up of the TSI task applied to location-specific processing (e.g., Tydgat & Grainger, 2009), a string of five characters (e.g., letters: FGJGM; symbols: £§?%@) is presented briefly while the participant is looking at the middle of the string. The string is subsequently followed by a pattern mask with a cue indicating one of the positions in the string (see Figure 2 for illustration). The participants' task is to choose, from the two alternatives, the one that matches the identity of the character at the cued location. When presented with strings of symbols, adult readers typically show a Λ-shape function (i.e., an advantage of the middle, fixated position) (Tydgat & Grainger, 2009; see also Chanceaux & Grainger, 2012; Chanceaux, Mathôt, & Grainger, 2014; Grainger, Tydgat, & Isselé, 2010; Scaltritti, Dufau, & Grainger, 2018; Vejnović & Zdravković, 2015). In contrast, for letter strings, adult readers typically show a W-shape serial position function of accuracy (see Tydgat & Grainger, 2009). That is, there is an advantage not only for the fixated, middle letter, but also of the exterior letters – primarily the initial letter. The dissociation in serial position function for strings of symbols vs. letters has also been obtained with developing readers (see Ziegler, Pech-Georgel, Dufau, & Grainger, 2010). To explain this pattern, Tydgat and Grainger (2009) proposed the Modified Receptive Field (MRF) theory. The idea is that, at the onset of learning-to-read, the status of letters changes from being independent visual objects to becoming parts of a higher-order object (i.e., the string). This is attained by adapting the mechanisms of visual object processing to the constraints of visual word processing (see Grainger, 2018; Grainger & Hannagan, 2014; see also Dehaene *et al.*, 2005, for neural correlates). Specifically, the MRF theory assumes that, with reading acquisition, location-specific letter detectors are developed and their receptive fields become reduced in size and elongated to the left – note that the initial letter is critical to translating orthographic representations into phonological representations (Grainger, Bertrand, Lété, Beyersmann, & Ziegler, 2016). Thus, the emergence of location-specific processing when learning a new script is expected to produce an initial position advantage in a TSI task.

The empirical data on the emergence of location-invariant and/or location-specific processing are very scarce (see Duñabeitia *et al.*, 2015, for an exception). Duñabeitia *et al.* (2015) examined the emergence of location-invariant processing in a longitudinal same-

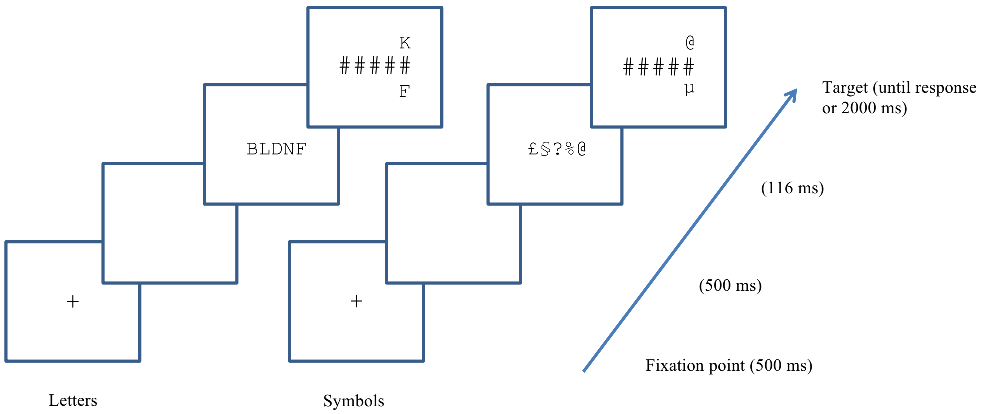

Figure 2. Depiction of the target-in-string identification task. [Colour figure can be viewed at wileyonlinelibrary.com]

different experiment with children from four years (i.e., pre-literate children) to six years (i.e., children who had acquired orthographic representations). In the accuracy data, Duñabeitia *et al*. (2015) found a significant TL effect for the older children, but not for the pre-literate children, and claimed that the 'the skills related to the processing of internal characters' identities and positions are inherently dependent on literacy' (p. 548). However, *d'* (i.e., a measure of sensitivity) did not differ from zero in the experiment with pre-literate children (i.e., children were performing at chance level), which raises questions about any interpretation of the data (see Perea, Jiménez, & Gomez, 2016, for discussion)[1].

The main goal of the present experiments was to overcome this gap by examining whether acquiring a new script affects location-invariant processing and location-specific processing by using a same-different task and a TSI task, respectively. We designed a laboratory analogue of children's reading acquisition in which adults were trained to read and write in a new, unfamiliar script. As Chetail (2017) indicated, the use of artificial scripts (i.e., sets of characters either from unknown scripts or newly devised) provides a unique opportunity to 'examine the developmental course of a given orthographic process which is stable in adults' (p. 103). Furthermore, recruiting adults as subjects allows us to control the participants' prior knowledge (i.e., we can make sure that participants are not familiarized with the characters; see Maurer, Blau, Yoncheva, & McCandliss, 2010; Taylor, Davis, & Rastle, 2017), and it also enables us to increase the number of conditions and trials of the experiments: Adult participants can carry out longer experimental sessions than children, and this allows us to ensure appropriate reliability. Additionally, comparing the results of pre-literate children and developing readers may lead to intricate interpretative issues, as the accuracy and latency data vary dramatically across groups (see Perea *et al*., 2016). Critically, the behavioural effects of learning an unfamiliar script in adults can be generalized to the effects elicited on children when learning their first language (see Taylor, Plunkett, & Nation, 2011, for discussion).

[1] *While the Perea* et al. *(2016) experiment with pre-literate children rules out an interpretation of TL effects as being fully dependent on literacy acquisition, it does not provide any insights as to the emergence of location-invariant processing. To test whether location-invariant processing is influenced by literacy in young children, one would need to run a retest – ideally with letters vs. symbols (or letters from a new alphabet) – after these children learn to read.*

In the current study, we employed a classic design with a pre-training phase and a post-training phase. The pre-training phase comprises two experiments: a same-different experiment on the TL effect (i.e., testing location-invariant processing; Experiment 1) and an experiment with a TSI task on the serial position function (i.e., testing location-specific processing). The pre-training experiments were conducted using eighteen consonant letters from an artificial monospaced font (BACS2serif; Vidal & Chetail, 2017). This font was used to create two different scripts: Script 1 (11 letters; two vowels and nine consonants) and Script 2 (11 letters; two vowels and nine consonants) (see Figure 3). One of the scripts was learned via print-to-sound training along five days, and the other was used as a control. An important issue is the choice of the appropriate control script. Keep in mind that the letters in the trained script would not only activate print-to-sound correspondences, but they would also be visually familiar. That is, after training, the pseudoletter ‘ ’ would not only correspond to a phoneme, but it also would become a familiar object. Thus, one could argue that any effects from the trained script in the post-training phase could be merely due to visual familiarity. To separate the effects of learning-to-read from the effects of visual familiarity, participants were familiarized with the visual form of the characters of the control script.

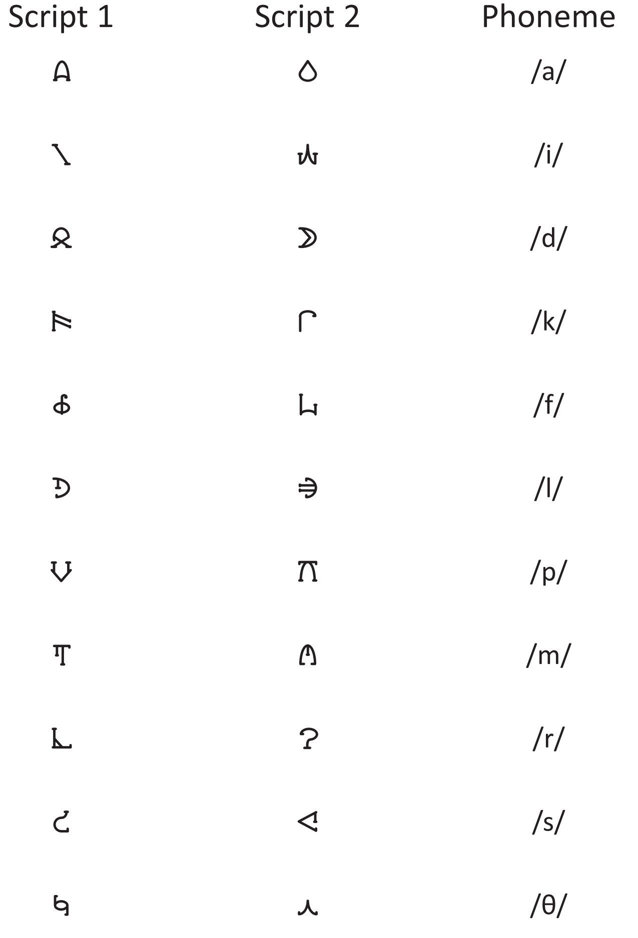

Figure 3. Association between the letters of the new scripts and their corresponding phonemes.

In the pre-training phase of Experiment 1, we conducted a same-different task using five-letter strings. In the pre-training phase of Experiment 2, we conducted a TSI task with five-letter strings. In both experiments, Script 1 and Script 2 were presented in separate blocks. Subsequently, each participant received training in one of the scripts (trained script) and was familiarized with the characters of the other script (control script). For the print-to-sound training, each individual learned the grapheme–phoneme associations of nine consonant letters and two vowel letters from one of the two scripts across six training sessions: Half of the participants learned the letters in Script 1 and the other half learned the letters in Script 2. Prior research has shown that readers can show some expertise in a new script quite rapidly. For instance, Chetail (2017) found that individuals acquired new regularities (e.g., letter and bigram frequency effects) after a relatively short amount of time, even in unfamiliar and complex scripts. Likewise, Brem *et al.* (2018) reported that two hours of training were enough for individuals to show some expertise for a novel script (e.g., an increase of the N1 amplitude). For the visual familiarization with the control script, participants were exposed to the eleven characters of the script, but without mentioning any orthographic or phonological information.

On day 1, participants first learned the grapheme–phoneme associations in the trained script. As in Spanish – their native language, all the grapheme–phoneme associations are transparent (e.g., the letter 'i' always corresponds to the phoneme/i/). Importantly, the print-to-sound training allows us to ensure that participants will learn the new script as a group of letters and not as mere symbols or shapes. As Chetail (2017) pointed out, 'a critical feature that distinguished letters from other symbols or shapes is that letters are used to transcribe speech according to a structure code' (p. 110). To consolidate the learning of the trained script over the next five sessions, which took place in a window of five working days, participants were asked to read aloud and write down series of items of increasing length, from 4-letter to 8-letter strings. Furthermore, the participants were familiarized with the visual forms of the characters of the control script. To that end, on each training session, each individual performed a character detection task and a character count task (see Figure 4 for depiction of each task; see also Chetail, 2017, for a similar strategy).

Finally, all participants had to pass a final test on the sixth day to show that they successfully acquired the experimental script. Once the individuals had passed this test (the criterion was set at 20 or more correct responses out of 24 in reading/writing within a time limit), they took part in the post-training experiments. What we should note here is that the print-to-sound training occurred in absence of semantics. Recent research has emphasized the role of sound-based strategies when learning-to-read a new script (e.g., see Brem *et al.*, 2018; Taylor *et al.*, 2017, for evidence with adults). This print-to-sound training also enabled us to isolate the orthographic processes from semantic processes, thus minimizing any influences from top-down processes.

The post-training experiments were parallel to the experiments from the pre-training phase. They were designed to test whether location-invariant and location-specific processing has emerged in the trained script – for comparison purposes, we also conducted a block with stimuli in an overlearned script (i.e., Roman alphabet). The predictions were clear. If location-specific orthographic processes emerge after literacy acquisition – as proposed by Dandurand *et al.* (2010) and Duñabeitia *et al.* (2014, 2015), we would expect a greater TL effect in a same-different task for the trained script in the post-training phase than in the pre-training phase. Keep in mind that, in the post-training phase, the TL effect for the newly learned script would have two constituents: a positional noise component – shared with the pre-training phase – and an orthographically based

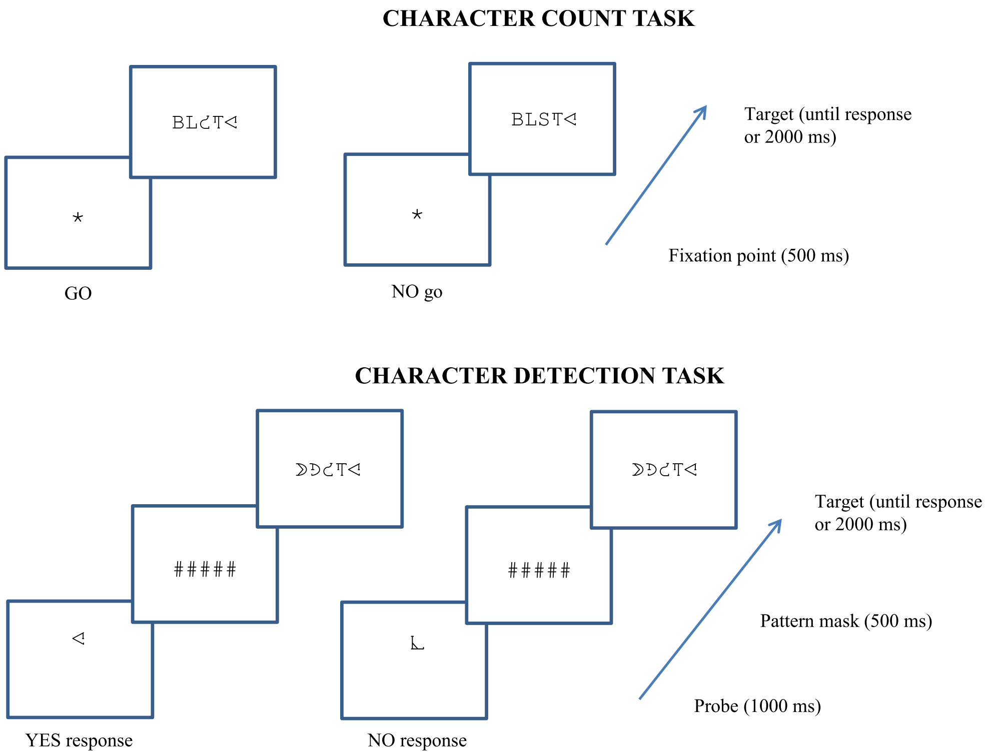

Figure 4. Depiction of the character count task (top panel) and the character detection task (low panel). [Colour figure can be viewed at wileyonlinelibrary.com]

component. For the control script, the TL effect should remain similar in magnitude in the pre- and post-training phases: the TL effect would be due to perceptual uncertainty. Alternatively, if TL effects were similar in magnitude in the pre- and post-training phase for the two scripts, this would reveal that the emergence of location-invariant processing does not emerge quickly after learning print-to-sound correspondences in a new script.

Second, if location-specific letter detectors are formed with literacy acquisition – as proposed by the MRF theory (Tydgat & Grainger, 2009), the trained script would elicit a first-letter advantage in the TSI task when measuring the serial position function in the post-training phase. This would be accompanied by an advantage of the fixed, middle position (i.e., a Λ-shape function) for the two scripts in the pre-training phase and for the control script in the post-training phase. Alternatively, if the pattern of data still shows a Λ-shape function for both the trained and untrained scripts in the post-training phase, this would suggest that print-to-sound training does not rapidly lead to the emergence of location-specific orthographic processing.

EXPERIMENT 1: LOCATION-INVARIANT PROCESSING

Method

Participants

The sample was composed of twenty-eight university students, all of them native speakers of Spanish with normal/corrected-to-normal vision and with no history of reading or hearing disorders. All of them signed an informed consent form before participating in the experiment. Participants received a small monetary compensation after the experiment. In consonance with the registered protocol, the final number of participants was determined via a sequential Bayes factor design maximal n (see Schönbrodt & Wagenmakers, 2018, for the advantages of this approach) starting with a sample size of 28 participants. To compute the Bayes factors for the critical interaction (i.e., the three-way interaction between Phase × Script × Probe-target relationship in the accuracy data) required for the sampling procedure, we obtained the Bayes factors in the by-subjects Bayesian ANOVA – note that all the stimuli were strings of random consonants (or consonants from artificial scripts) so generalization over participants was more important than generalization over random consonants. This Bayes factor was computed in JASP (Wagenmakers *et al.*, 2018) as the ratio of the model that contained the factor of interest (i.e., all the main effects, two-way interactions, and the three-way interaction) vs. the model that did not contain the effects of interest (i.e., all the main effects and the two-way interactions). This Bayes factor exceeded 6 (i.e., the criteria established in the pre-registered protocol) in favour of the null hypothesis ($BF_{10} = .081 \rightarrow BF_{01} > 12$), so sampling was stopped with $n = 28$.

Materials

We created 240 five-consonant string pairs (probe and target) in Script 1 (see Figure 3), 240 five-consonant string pairs in Script 2 (see Figure 3), and 240 five-consonant string pairs in Roman alphabet (using Courier New font; e.g., `STNGB`). The two artificial scripts stemmed from the same font: BACS2serif (Vidal & Chetail, 2017). The string pairs in Script 1 and Script 2 were presented in separate, counterbalanced blocks – the string pairs in the Roman script were presented at the end of the post-training phase. All character strings were composed of non-repeated letters. There were 120 ‘different’ pairs and 120 ‘same’ pairs for each character string type. For the ‘different’ pairs in each script, 60 pairs were created by transposing two adjacent letters (e.g., ᑕ丅ᒪᗤᐯ - ᑕᒪ丅ᗤᐯ; 1st-2nd, 2nd-3rd, 3rd-4th, 4th-5th), and 60 pairs were created by replacing two adjacent letters (e.g., ᑕ丅ᒪᗤᐯ - ᑕ⋝ᘓᗤᐯ; 1st-2nd, 2nd-3rd, 3rd-4th, 4th-5th). Thus, each block contained 240 pairs of character strings. In total, each participant was given 480 trials in the pre-training phase (240 string pairs in Script 1 and 240 string pairs in Script 2) and 720 trials in the post-training phase (240 string pairs in Script 1, 240 string pairs in Script 2, and 240 string pairs in Roman script). The proportion of transpositions/replacements was the same in all possible locations. To counterbalance the probe-target pairs, we created two lists for each script (see Massol *et al.*, 2013, for a similar procedure). For the practice phase, we also created eight five-consonant string pairs for each block (eight pairs in the Script 1 block, eight pairs in the Script 2 block, and eight pairs in the Roman alphabet block).

For the learning-to-read sessions, we created a template in a standard presentation software with the graphemes of the new script and the associated phoneme (see Figure 3), which were recorded by a female voice and digitalized at a sampling rate of 44.1 kHz. For each script, we created 18 items and 18 utterances of 4 characters and 5 characters, respectively, 30 items and 30 utterances of 6 characters, 66 items and 66 utterances of 7 characters, and 120 items and 120 utterances of 8 characters. The items and utterances were grouped separately by lists of 12 items that were presented with standard presentation software.

To have the participants familiarized with the characters of the control script, we employed a character detection task and a character count task (see Figure 4). For the character detection task, we created a total of 144 pairs of items in each script (i.e., probe and target). The probe was always a single character and the target consisted of a string of artificial characters with different length (18 items of four and five characters, respectively, 30 items of 6 characters, 66 items of 7 characters and 120 items of 8 characters). For the character count task, we employed, for each script, 18 items of four and 5 characters, respectively, 30 items of 6 characters, 66 items of 7 characters and 120 items of 8 characters. Each string could be composed by artificial characters or by artificial and Roman characters.

Procedure

Pre-training-test. Participants were tested either individually or in groups of two in a quiet room. DMDX software (Forster & Forster, 2003) was used to display the sequence of stimuli and to register the timing/accuracy of the responses. Response times were measured from target onset until the participant's response. On each trial, a fixation point (*) was displayed for 500 ms in the centre of a computer screen. Next, the fixation point was replaced by a probe, which was presented for 300 ms and positioned 3 mm above the centre of the screen. Then, the target item appeared one line 3 mm below the centre of the screen. The target remained on the screen until the response or 2,000 ms had passed. All stimuli were presented in a monospaced font (15 pt BACS2serif for Scripts 1 and 2; 15 pt Courier New for the Roman letters) in black on a white background. Participants were told that they would be presented with two strings of consonants and that they would have to press the 'yes' key if they were the same, and they were asked to press the 'no' button if they were different (see Figure 1). Participants were instructed to make this decision as quickly and as accurately as possible. Eight practice trials preceded the 240 experimental trials in each block. Participants did not receive feedback during the experiment. The session lasted for 18–22 min.

Training. Participants were trained individually in the presence of the experimenter along a window of six working days in a quiet room (see Figure 5). Half of the participants learned the grapheme–phoneme associations in Script 1 and the other half learned the grapheme–phoneme associations in Script 2. There were two blocks in each session: For the trained script, participants received print-to-sound training (i.e., grapheme–phoneme association) and, for the control script, they participated in tasks that entailed visual familiarization with the control characters – the order of each block was counterbalanced across sessions.

On the first day, after the pre-training experiments, participants learned the association between the spoken forms of each grapheme in one of the unknown scripts (i.e., the experimental script: Script 1 or Script 2; see Figure 3) and they also familiarized with the visual form of the control script (i.e., the script not used for the grapheme–phoneme association). For the print-to-sound training, the characters were presented on a computer screen with their corresponding sound (participants could click on the character with the mouse and listen to its associated phoneme). Participants were also asked to hand-copy the new letters on a sheet of paper and they were given as much time as they needed to learn these associations.

For the visual familiarization part, the characters of the control script were presented one by one on a computer screen in absence of any orthographic or phonological

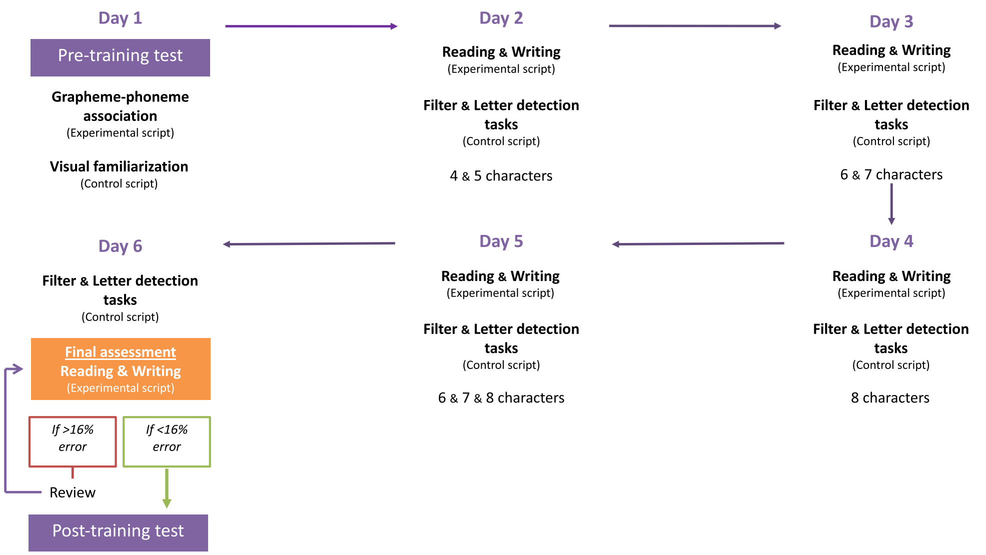

Figure 5. Schematic depiction of the training sessions. [Colour figure can be viewed at wileyonlinelibrary.com]

information. Participants had to hand-copy the control characters on a sheet of paper, without a time deadline. Then, all the characters of the control script were presented and participants took as much time as they wanted to familiarize with them (see Chetail, 2017, for a similar procedure). The experimenter checked and corrected (when necessary) the writing of the letters and characters to minimize the differences in handwriting quality[2].

On the second day, participants were presented with items of four and five characters; on the third day, they were presented with items of six and seven characters; on the fourth day, they practised with items of eight characters and, on the fifth day, participants were presented with items of six, seven and eight characters (see Figure 5). For the print-to-sound training, the general procedure was as follows. Participants had to read aloud and write down 36 items without time deadline. The items were presented on the computer screen, divided into alternating blocks (reading and writing) of 12 items. For reading aloud, a list of 12 items was presented on the screen (e.g., 'ᛟᗅ⊧⧵') and participants were asked to read the items one by one (e.g.,/daki/). During this task, the experimenter provided feedback after each item (i.e., correct/incorrect response). If the participant made a mistake, the experimenter encouraged her/him to read it again on her/his own. If the participant could not figure out the correct response, the experimenter indicated it, remarking the grapheme–phoneme correspondences. For the writing blocks, a list of 12 'loudspeaker' signs (🔊) were presented on the screen. Participants were asked to press the 🔊 sign, listen to the pronunciation (e.g.,/daki/), and then write down the corresponding graphemes in a sheet of paper (e.g., 'ᛟᗅ⊧⧵'). As blocks of 12 items were presented simultaneously, participants were able to see and listen to each item as many times as needed. The experimenter provided feedback after each item (i.e., correct/

[2] *Exact handwritten copies of the character were not required, as neither are exact copies of the letters when children learn to write. It was enough if the handwritten copy of the character approximated to the original to be identified and distinguishable from the other characters.*

incorrect response). If the participant made a mistake, the experimenter asked her/him to listen the item again. If the participant could not figure out the mistake on her/his own, the experimenter told her/him the correct response remarking the grapheme–phoneme correspondences.

For each block in both tasks (i.e., reading aloud and write down), the experimenter annotated the mistakes (if any), the order of the tasks, performing times, and other comments (e.g., the most repeated errors) in an assessment form. Moreover, after each block of 12 items, the experimenter provided general feedback of performance (i.e., correct responses, type of errors, and timing). Importantly, on the first sessions (days 2 and 3), the learning goal was to correctly establish the grapheme–phoneme correspondences. Thus, the experimenter focused mainly on the errors made by the participants. Then, on sessions 4 and 5, when the participants hardly made any mistakes, the experimenter encouraged them to read and write down as fast as possible.

For the visual familiarization block with the control script, participants completed a character detection task and a character count task in each training session. The items had the same length as the items of its corresponding print-to-sound training session. For both tasks, DMDX software (Forster & Forster, 2003) was used to display the sequence of stimuli and to register the timing/accuracy of the responses. On each trial of the character detection task (see Figure 4), a probe was presented for 1,000 ms one line 3 mm above the centre of the screen. The probe was subsequently replaced by a pattern mask with same length as the subsequent target ('#####') on the centre of the screen for 500 ms. Then, the target appeared and remained on the screen until response or 2,000 ms had passed. All stimuli were presented in a monospaced font (15 pt BACS2serif) in black on a white background. Participants were told that they would be presented with a character and then with a string of characters (both in the control script) and they would have to press the 'yes' key if the probe appeared in the subsequent string, and they were asked to press the 'no' button if the single character did not appear in the string.

On each trial of the character count task (see Figure 4), a fixation point (*) was displayed for 500 ms on the centre of a computer screen. Next, the fixation point was replaced by the target (i.e., a character string). The target remained on the screen until the response or 2,000 ms had passed. All stimuli were presented in a monospaced font (15 pt BACS2serif and 15 pt Courier New) in black on a white background. Participants were asked to press the 'yes' key only when the item presented was composed of 3 or more characters of the control script – keep in mind that the target items consisted of only control script characters (BACS2serif) or a mixture of characters of the control and Roman scripts. 30% of the items consisted of only one or two control script characters and Roman characters (i.e., participants should not press the 'yes' key). Participants received feedback on the general accuracy after each task.

Finally, on the sixth day, before conducting the post-training experiments, participants received a final test with 24 items of 8 characters (12 for reading aloud and 12 for writing) presented in the same format as in the training. They had to do the test in less than 1 min and 30 s, and 3 min and 30 s, respectively[3]. Those participants who passed the assessment with at least 84% of accuracy (20 out of 24 correct responses) took part in the post-training experiments[4].

---

[3] *The time limit was set by averaging the reaction times of two pilot participants and adding 30 s more – keep in mind that the pilot participants were members of the laboratory and they were highly motivated.*

[4] *A minimum of 70% accuracy was required in the visual control tasks. All participants met this criterion.*

Post-training test. The post-training test was the same as in Experiment 1a, with a final additional block with Roman letters. The post-training tests lasted for approximately 25–30 min.

Results

All participants were able to write and read the newly learned script fluently and passed the final training test at the first attempt (see Appendix B, for performance of participants along the training sessions). In accordance with the pre-registered protocol, one participant was replaced because of an overall accuracy level below .60 in the replacement-letter condition.

Confirmatory analysis

The dependent variables were response time and accuracy. Error and extremely short responses (less than 250 ms: 0 responses) were omitted from the latency analyses – there were no responses longer than the 2-s deadline (i.e., they were automatically categorized as errors). The mean RTs for the correct responses and the accuracy in each experimental condition are presented in Table 1 (see also Figures 6 and 7). We performed the statistical inference not only using (generalized) linear mixed-effects models, but we also computed Bayes factors. Table 2 presents a summary of the main points of the experiment (i.e., research question, key comparisons, predictions, statistical analyses, main findings, and conclusions).

Different trials

In the inferential analyses, we focused on 'different' trials, as these are the ones with the TL manipulation. The main research question in the experiment was whether location-invariant processing – as measured by the TL effect – emerges in the trained but not in the untrained script in the post-training phase. To test this hypothesis, we employed (generalized) linear mixed-effects (LME) models in R (R Core Team, 2019) using the lme4 1.1-21 package (Bates, Maechler, Bolker, & Walker, 2019) and the BayesFactor 0.9.12-4.2 package (Morey & Rouder, 2018) with three fixed factors: Phase (pre- vs. post-training), Script (trained vs. untrained), and Probe-target relationship (transposed, replaced). Regarding the LME analyses, because of the normality assumption required, the raw RTs were inverse-transformed (−1000/RT). The most complex fitted model that converged

Table 1. Mean correct response times (in ms) and accuracy (in brackets) in the different conditions of Experiment 1

	Untrained Script			Trained Script		
	Different			Different		
	Transposed	Replaced	Same	Transposed	Replaced	Same
Pre-training	624 (.550)	602 (.714)	550 (.908)	639 (.551)	624 (.733)	581 (.868)
Post-training	585 (.603)	566 (.783)	540 (.888)	600 (.603)	572 (.795)	538 (.910)

Note. For the Roman script, the correct response times and accuracy (in brackets) were 590 ms (.529) for transposed pairs, 574 ms (.812) for replaced pairs, and 511 ms (.931) for same pairs.

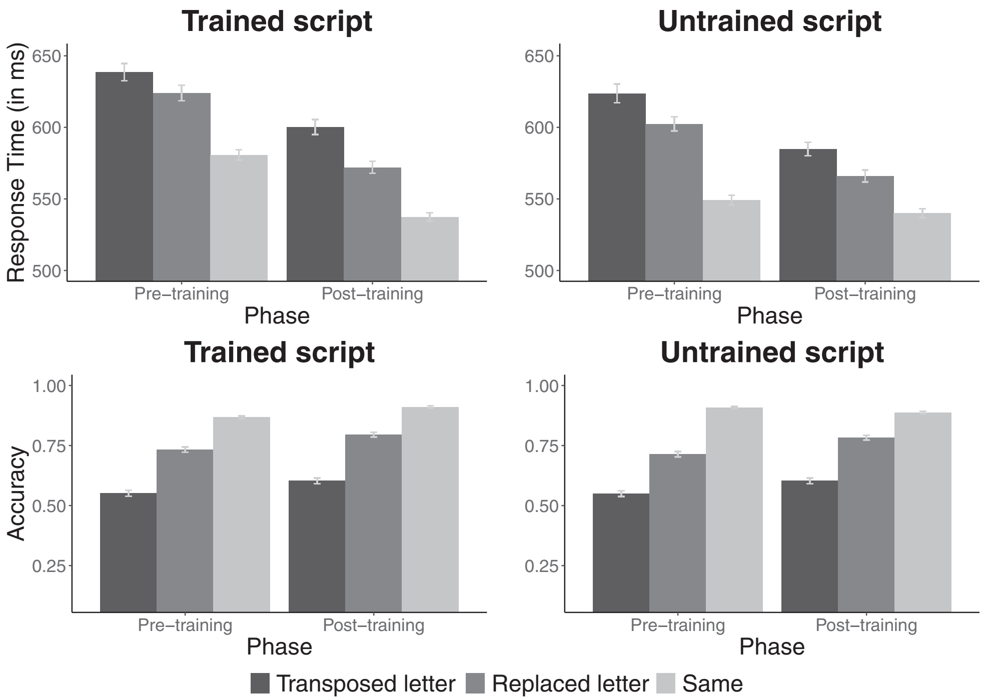

Figure 6. Mean reaction times (top panel), accuracies (bottom panel), and standard errors in the trained and untrained scripts of Experiment 1.

was: −1000/RT ~ script*condition*phase + (1 + script+phase|subject) + (1|item). For the generalized linear mixed analyses of the accuracy data, responses were coded as binary values (1 = correct, 0 = incorrect) and we used the glmer function in the lme4 package (family = binomial). The most complex fitted model that converged was as follows: accuracy ~ script*condition*phase + (1 + script|subject) + (1|item). (In Appendix A, we report the [non-pre-registered] analyses using Bayesian linear mixed-effects models with the maximal random structure – the results were essentially the same as those reported here.) To compute the Bayes factors on the latency data, we used lmBF function from the BayesFactor package with the default Cauchy distribution (centred around 0 and with a width parameter $\delta = 0.707$) (see Rouder, Speckman, Sun, Morey, & Iverson, 2009; Wagenmakers *et al*., 2017, for discussion). To compute the Bayes factors on the accuracy data, we calculated the Bayes factors from Bayesian analyses of variance (ANOVAs) with the aggregated data by participants – note that items were strings of letters in an artificial script. For the computation of the Bayes factors for each effect, we followed the same logic as in prior research (see Leinenger, Myslín, Rayner, & Levy, 2017; Staub & Goddard, 2019, for illustration). For the numerator, we compared the maximal model that included the effect of interest vs. a null model that does not assume any fixed effects or interactions. For the denominator, we compared the maximal model after excluding the effect of interest vs. a null model that does not assume any fixed effects or interactions. The ratio between these two Bayes factors was the Bayes Factor of the effect.

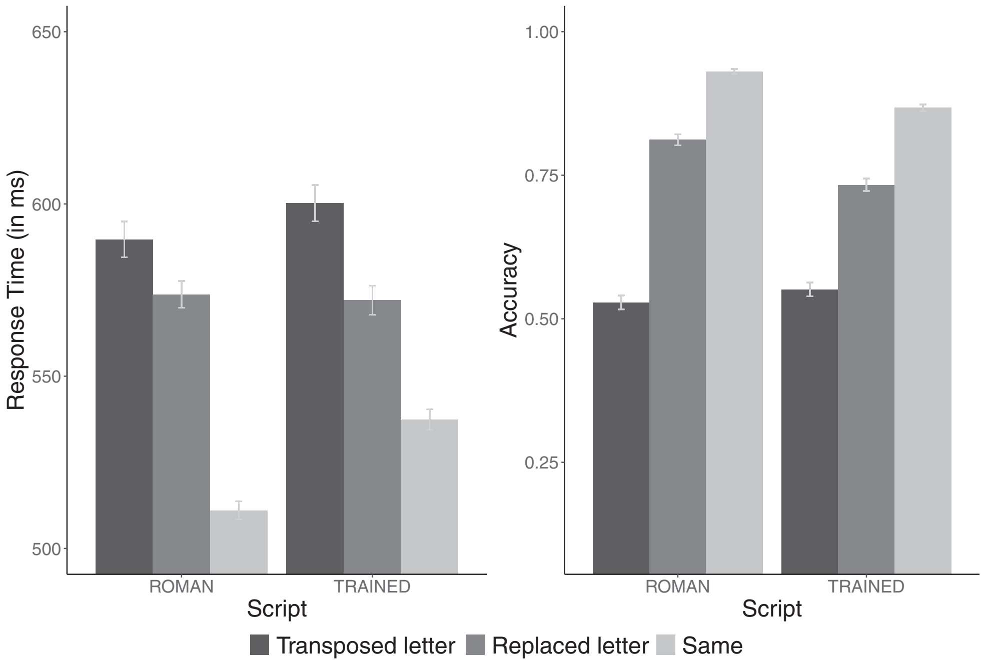

Figure 7. Mean reaction times (left panel), accuracies (right panel), and standard errors in the trained (post-training phase) and Roman scripts of Experiment 1.

Latency analyses. Responses were, on average, 21 ms faster for the RL condition than for the TL condition (i.e., overall TL effect; 591 vs. 612 ms; $b = .08$; $t = 6.18$ $p < .001$; $BF_{10} = 6.69e + 05$) and participants responded, on average, 41 ms faster in the post-training tests than in the pre-training tests (581 vs. 622 ms; $b = .13$; $t = 4.55$ $p < .001$; $BF_{10} = 20.34$). We found no overall differences in response times between the trained and untrained script ($b = -.01$; $t = -.60$, $p = .55$; $BF_{10} = 1.48$). The interaction between Phase and Script barely reached the significance level in the frequentist analyses ($b = -.04$; $t = -2.43$, $p = .02$), but the Bayes factors indicated anecdotal evidence towards a null effect ($BF_{10} = 0.55$). None of the other interactions approached significance (all $ts < .78$, $ps > .59$; all $BF_{10} < .35$) (see Figure 6).

Accuracy analyses. Accuracy was higher for the RL condition than for the TL condition (.756 vs. .577; $b = -1.05$; $z = -12.73$, $p < .001$; $BF_{10} = 1.071e + 39$), and participants were more accurate in the post-training phase than in the pre-training phase (.696 vs. .637; $b = -.38$; $z = -4.44$, $p <.001$; $BF_{10} = 415911$). Neither the effect of Script ($b = -.08$; $z = -.83$, $p = .41$; $BF_{10} = .21$) nor any of the interactions approached significance (all $zs < 1.19$, $ps > .23$; all $BFs_{10} < .26$) (see Figure 6).

Exploratory analyses

As indicated in the pre-registration, we also compared the TL effect of the trained script (post-training phase) and the TL effect in an overlearned script (i.e., the Roman script). The two fixed factors in the analyses were Script (Trained [post-training] vs. Roman) and Probe-target relationship (transposed, replaced) – the inferential analyses were parallel to

Table 2. Summary of Experiment 1 (Location-invariant processing; same-different task)

Question	Key comparisons	Predictions				Statistical analyses	Results (main findings)	Conclusions	
Does location - invariant processing emerge rapidly after learning-to-read an artificial script?	Script (Trained vs. Untrained) × Phase (Pre- vs. Post-training) × Condition (Transposed vs. Replaced letter)	Greater TL effect for the trained script in the post-training phase than in the pre-training phase. Similar TL effect for the control script in both phases.	Confirmatory (Different trials)	✓	(G)LME	−1000/RT ~ script*condition*phase + (1 + script+phase|subject) + (1|item)	Condition Phase Phase*Script	The magnitude of the transposed-letter effect was similar for the trained script and for the untrained script.	Location-invariant processing does not emerge rapidly after learning-to-read an artificial script
						accuracy ~ script*condition*phase + (1 + script|subject) + (1|item)	Condition Phase		
				✓	BF	RT ~ subject * script * condition * phase	Condition Phase		
						Accuracy: Bayesian ANOVA with script, condition, and phase as factors	Condition Phase		
				✗	BRMS	1000/RT ~ script * condition * phase + (1 + script * condition * phase|subject) + (1 + condition * phase| item)	Condition Phase		
						accuracy ~ script * condition * phase + (1 + script * condition * phase|subject) + (1 + condition * phase| item)	Condition Phase		
	Script (Trained [post-training] vs. Roman) × Condition (Transposed vs. Replaced letter)	Greater TL effect for the Roman script than for the trained script.	Exploratory	✓	(G)LME	−1000/RT ~ script*condition + (1 + script|subject) + (1|item)	Condition	The magnitude of the transposed-letter effect was greater in the Roman script than in the newly learned script in the accuracy analyses	
						accuracy ~ script*condition + (1 + script|subject) + (1|item)	Condition Condition*Script		
				✓	BF	RT ~ subject * script * condition * phase	Condition		
						Accuracy: Bayesian ANOVA with script and condition as factors	Condition Condition*Script		
				✗	BRMS	−1000/RT ~ script*condition + (1 + script*condition |subject) + (1 + script*condition|item)	Condition		
						−1000/RT ~ script*condition + (1 + script*condition |subject) + (1 + script*condition|item)	Condition Condition*Script		

Continued

Table 2. (*Continued*)

Question	Key comparisons	Predictions	Statistical analyses				Results (main findings)	Conclusions	
	Script (Trained vs. Untrained) × Phase (Pre- vs. Post-training)		Same trials	✗	LME	−1000/RT ~ script*phase + (1 + phase|subject) + (1|item)	Phase Script*Phase	♁'Same' responses were substantially faster and more accurate in the post-training phase than in the pre-training phase only for the trained script.	Learning-to-read in the new script helped encoding the letter strings emergence of rudimentary orthographic representations.
						accuracy ~ script*phase + (1 + phase|subject) + (1|item)	Phase Script Script*Phase		
				✗	BRMS	−1000/RT ~ script*phase + (1 + script*phase|subject) + (1 + script*phase|item)	Phase Script*Phase		
						accuracy ~ script*phase + (1 + script*phase|subject) + (1 + script*phase|item)	Phase Script Script*Phase		

Note. ✔ [pre-registered analyses]. ✗ [non-pre-registered analyses]; number of participants was determined via sequential Bayes factor design obtaining BF in the by-subjects ANOVA ($BF_{10} = .081 \rightarrow BF_{01} > 12$).

those described above. The most complex fitted model that converged was: Dependent_Variable [−1000/RT or accuracy] ~ script*condition + (1 + script|subject) + (1|item) (see Figure 7).

Latency analyses. Participants responded, on average, 22 ms faster in the RL condition than in the TL condition (573 vs. 595; $b = .077$ $t = 6.00, p < .001$; $BF_{10} = 129.80$). There were no overall differences between the trained script and the Roman script ($b = .02$ $t = .38$, $p = .71$; $BF_{10} = 0.53$). The interaction between probe-target relationship and script was not significant ($b = -.01$; $t = -.41, p = .68$; $BF_{10} = 0.54$).

Accuracy analyses. Participants were more accurate in the RL condition than in the TL condition (.804 vs. .566; $b = -1.12$; $z = -13.06$, $p < .001$; $BF_{10} = 3.308e + 22$), whereas the effect of script was not significant ($b = .22; z = 1.14, p = .25$; $BF_{10} = 0.734$). We found a significant interaction between the two factors ($b = -.58$; $z = -4.67$, $p < .001$; $BF_{10} = 7.758$), which reflects that the TL effect was greater in the Roman script than in the trained script (.283 vs. .192).

Same trials

While not indicated in the pre-registered protocol, the examination of 'same' responses in the pre- and post-test phases for the trained and untrained scripts may shed some light on the role of orthographic-phonological training when processing letter strings. To analyse 'same' responses, we employed (generalized) linear mixed-effects models on the latency and accuracy data. The two fixed effects were Script (trained vs. untrained) and Phase (pre-training, post-training). The most complex model that converged in the (generalized) linear mixed-effects models was: Dependent_Variable [−1000/RT or accuracy] ~ script * phase + (1 + phase | subject) + (1 | item). These analyses were complemented with Bayesian linear mixed-effects models using the maximal random factor structure (see Appendix A).

Latency analyses. Responses were, on average, 26 ms faster in the post-training phase than in the pre-training phase (539 vs. 565 ms; $b = .14; t = 2.93, p = .01$), whereas there were no signs of an effect of script ($b = .01$; $t = -.51$, $p = .61$). More important, the interaction between Script and Phase was significant ($b = -.12$; $t = -8.09$, $p < .001$). This reflected that responses were faster in the post-training than in the pre-training phase for the trained script (41 ms; 537 vs. 581 ms), but not for the untrained script (9 ms; 540 vs. 549 ms) (see Figure 6).

Accuracy analyses. Participants were more accurate in the post-training than in the pre-training phase (i.e., main effect of phase; $b = -.48$; $z = -2.97, p = .003$) and with the untrained than with the trained script (i.e., main effect of script; $b = -.27$; $z = -3.17$, $p = .001$) – note that the effect of script was .009 and was not corroborated by the Bayesian linear mixed-effects analyses (see Table A3). More important, mimicking the latency analyses, we found an interaction between the two factors ($b = .72$; $z = 6.16$, $p < .001$): Participants were more accurate in the post-training test than in the pre-training test for the trained script (.910 vs. .868, respectively), but not for the untrained script (.888 in the post-training vs. .908 in the pre-training test) (see Figure 6).

Discussion

As usual, we found a substantial transposed-letter effect for 'different' responses in the same-different task: Participants' responses were faster and more accurate for replacement-letter pairs than for transposed-letter pairs (see Krueger, 1978; Ratcliff, 1981, for early evidence; see also Duñabeitia *et al.*, 2012; Massol *et al.*, 2013; Perea *et al.*, 2016). More important for the purposes of the experiment, the magnitude of the transposed-letter effect was similar for the trained script and for the (visual) control script in both latency and accuracy data. We did find that the responses to 'different' trials were, on average, faster and more accurate in the post-training phase than in the pre-training phase. However, this occurred similarly in both the trained and visual control scripts; hence, it could have been due to the participants' being more visually familiar with the new letters. In addition, the size of the transposed-letter effect was greater in the Roman script than in the newly learned script in the accuracy analyses (28.3% vs. 19.2% of errors, respectively) – note that previous studies on the transposed-letter effect also showed significant effects on accuracy, but not on response latencies (e.g., Massol *et al.*, 2013; Perea *et al.*, 2016; Perea & Lupker, 2004)[5]. This is consistent with the idea of orthographic location-invariant mechanisms being at work in the Roman script, but not in the newly learned script (see Duñabeitia *et al.*, 2012; Massol, *et al.*, 2013; see also García-Orza *et al.*, 2010; Muñoz *et al.*, 2012, for greater transposed-letter effect for letters than for other visual objects [symbols, digits, false fonts]).

The above results may offer the impression that training a new script did not create any stable orthographic representations. However, this interpretation is difficult to reconcile with the fact that 'same' responses were substantially faster and more accurate in the post-training phase than in the pre-training phase for the trained script (538 vs. 581 ms; .910 vs. .868), but not for the untrained script (540 vs. 549 ms; .888 vs. .908). This finding strongly suggests that learning-to-read in the new script helped encoding the letter strings, thus reflecting the emergence of rudimentary orthographic representations.

In sum, the current same-different experiment favours the view that location-invariant processing, as measured by the transposed-letter effect, does not emerge rapidly after learning-to-read in a new script. We defer a more detailed discussion of this issue in the General Discussion.

EXPERIMENT 2: LOCATION-SPECIFIC PROCESSING

Method

Participants

They were the same as in Experiment 1. To compute the Bayes factors for the critical interaction (i.e., the three-way interaction between Phase $\times$ Script $\times$ Position in the

---

[5] *Furthermore, for the Roman script, we found that size of the transposed-letter effect was considerably smaller for external than for internal transpositions: 14.5% vs. 42.1%, respectively (e.g., see Gomez* et al., *2008, for a similar pattern). In contrast, for the trained script, the size of the transposed-letter effect was only slightly lower for external transpositions than for internal transposition (17.0% vs. 21.4%, respectively). This again suggests that the transposed-letter effect in the Roman script and the trained script reflects different underlying processes.*

accuracy data) required for the sampling procedure, we obtained the Bayes factors in the by-subjects Bayesian ANOVA. This Bayes factor exceeded 6 (i.e., the criteria established in the pre-registered protocol), $BF_{10} = .05 \rightarrow BF_{01} = 20$, so sampling was stopped with $n = 28$.

Materials

Based on the design and procedure used by Tydgat and Grainger (2009), we created a set of 180 five-consonant strings in two unfamiliar scripts: 90 in Script 1 and 90 in Script 2, and – for the post-training phase – a set of 90 five-consonant strings in Roman alphabet (e.g., STNGB). None of the letter strings contained repeated characters.

We designed three different blocks (one for script: Script 1, Script 2, and Roman script) with 90 experimental and 9 practice trials each one. The order of the artificial script blocks was counterbalanced between subjects. We manipulated the target position in the array (1st, 2nd, 3rd, 4th, and 5th position). As in the Tydgat and Grainger (2009) experiments, each of the target characters was presented 2 times at each of the five target positions (once above and once below the backward mask), and 40 times at a non-target position (i.e., each target character played as alternative at each of the five positions). Importantly, the incorrect alternative was never presented in the stimulus array. We created two lists for each of the artificial scripts, manipulating the orientation of the target character (i.e., in List 1, the target was presented above the array, whereas in List 2 the same target was presented below the array). These two lists were presented to all participants, one for the pre-training test and the other for the post-training test. The sub-experiment with the Roman script was presented at the end of the post-training test.

The learning sessions materials were the same as in Experiment 1.

Procedure

Pre-training test. Participants were tested individually or in groups of two in a quiet room. DMDX software (Forster & Forster, 2003) was used to display the sequence of stimuli and to record the timing and accuracy of the responses. Each trial began with a fixation point (i.e., '+') that stayed on the screen for 500 ms and was followed by a 500-ms with the blank screen. Then, a string of five letters was presented for 116 ms (see Scaltritti *et al.*, 2018, for the same set-up). The array of characters was followed by a backward mask ('#####') accompanied by two characters, above and below the mask, respectively, at one of the five possible array positions (i.e., characters position as a post-cue) (see Figure 2). The stimuli were displayed on the screen until the participant responded or 2 seconds had passed. All stimuli were presented in black on a white background. We employed a monospaced font for the two scripts. Participants were asked to decide which of the two characters was present in the corresponding position of the preceding array. They were required to press the 'up arrow' key on the keyboard for the character above and the 'down arrow' key for the character below the array. They were explicitly instructed to fixate at the centre of the array and make the decision as quickly and as accurately as possible. The two scripts were presented in separated blocks, counterbalanced by subjects. Nine practice trials preceded the 90 experimental trials in each of the experimental conditions (90 trials in Script 1 and 90 trials in Script 2). Participants did not receive feedback during the experiment. There was a short break between blocks. The session lasted for around 20–25 min.

Training. It was the same as in Experiment 1 (see Figure 5).

Post-training-test. It was the same as in the pre-training test, except for the addition of a third block with stimuli in the Roman script. The session lasted for around 30 min.

Results

As indicated in the pre-registration, those participants with less than .60 of accuracy in the middle position in the pre- and post-training experiments were replaced – this occurred with four participants. Mean accuracies (and standard errors) for all target types and target positions are presented in Figure 8. The three fixed factors were Phase (pre- vs. post-training), Script (trained vs. control), and Position (1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 5^{th}). By-subjects and by-items classical and Bayesian ANOVAs were performed on the accuracy data. The computation of the Bayes factors was parallel to that described for accuracy in Experiment 1. In Appendix A, we report supplementary [non-pre-registered] analyses using Bayesian linear mixed-effects models (see Table 3 for a summary of the main points of Experiment 2).

Confirmatory analyses

The ANOVAs showed that accuracy was a function of serial position, $F_1(4, 108) = 67.87$, $p < .001$, $BF_{10} = 2.497e + 66$; $F_2(4,175) = 106.45$, $p < .001$, $BF_{10} = 9.445e + 36$. Accuracy levels were higher in the central, third position (.778) than in the first, second, fourth, and fifth positions (with mean accuracy levels of .535, .502, .510, and .507, respectively). Furthermore, the overall accuracy levels were virtually the same for the pre- and post-training phase, $F_1(1, 27) = 1.18$, $p = .29$. $BF_{10} = .10$; $F_2 < .001$, $BF_{10} = .09$, and for the trained and control scripts (both Fs < 1; both $BFs_{10} < .11$) (see Figure 6).

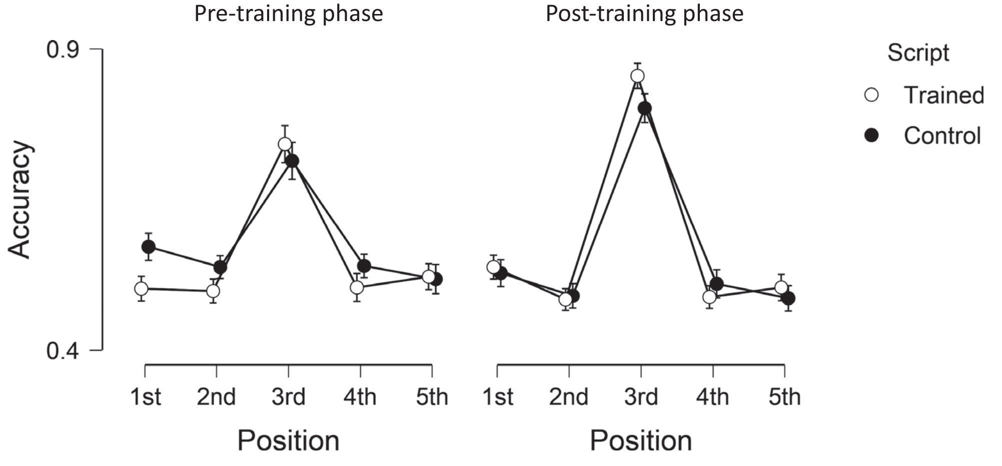

Figure 8. Mean accuracies and standard errors in the trained and untrained scripts of Experiment 2.

Table 3. Summary of Experiment 2 (location-specific processing; target-in-string identification task)

Question	Key comparisons	Predictions	Statistical analyses			Results (main effects)	Conclusions		
Does location-specific processing emerge rapidly after learning-to-read an artificial script?	Script (Trained vs. Untrained) × Phase (Pre- vs. Post-training) × Position (1st, 2nd, 3rd, 4th, 5th)	Greater accuracy for first-letter position in the trained script in the post-training phase than in the pre-training phase – this would be accompanied by the advantage of the middle position.	Confirmatory	✓	Accuracy: Frequentist ANOVA with Script, Phase, and Position as factors	Position (3rd) Position (3rd)*Phase	Advantage of the middle position Accuracy on third position greater in the post-training phase than in the pre-training phase	Advantage of the middle, fixated position over the other positions for the trained and untrained scripts in the pre- and post-training phases	Location-specific does not emerge rapidly after learning-to-read a new script.
				✓	Accuracy: Bayesian ANOVA with Script, Phase, and Position as factors	Position (3rd) Position (3rd)*Phase			
				✗	BRMS accuracy ~ script * position * phase + (1 + script * position * phase | subject) + (1 + position * phase | item)	Position (3rd) Position (3rd)*Phase			
	Script (Trained [post-training] vs. Roman) × Position (1st, 2nd, 3rd, 4th, 5th)	Greater first-letter advantage for the Roman script than for the trained script (post-training phase)	Exploratory	✓	Accuracy: Frequentist ANOVA with Script and Position as factors	Position (3rd)	Advantage of the middle position No differences between the Roman and Trained script		Attentional capture to the centre of the string
				✓	Accuracy: Bayesian ANOVA with Script and Position as factors	Position (3rd)			
				✗	BRMS accuracy ~ script * position + (1 + script * position | subject) + (1 + position | item)	Position (3rd)			

Note. ✓[pre-registered analyses]. ✗ [non-pre-registered analyses]; number of participants was determined via sequential Bayes factor design obtaining BF in the by-subjects ANOVA ($BF_{10} = .05 \rightarrow BF_{01} = 20$)

The serial position function differed in the pre-training and post-training phase (Position × Phase interaction; $F_1(4, 108) = 8.65$, $p < .001$, $BF_{10} = 241.21$; $F_2(4, 175) = 6.22$, $p < .001$, $BF_{10} = 53.42$): This occurred because accuracy in the third position was higher in the post-training phase than in the pre-training phase (.828 vs .728, respectively). Neither the interaction between Position and Script ($F_1(4, 108) = 2.65$, $p = .04$, $BF_{10} = .23$; $F_2 < 1$, $BF_{10} = .01$) nor the interaction between Script and Phase ($F_1(1, 108) = 3.22$, $p = .08$, $BF_{10} = .50$; $F_2 < 1$, $BF_{10} = .12$) were significant. Finally, there were no signs of a Phase × Script × Position interaction (both $F_s < 1$; $BFs_{10} < .09$).

Exploratory analyses

We also compared the serial position function of the trained script (in the post-training phase) and an overlearned script (i.e., Roman script) (see Figure 9). The two fixed factors in the ANOVAs were Script (Trained [post-training] vs. Roman) and Position (1st, 2nd, 3rd, 4th, 5th). The analyses were parallel to those described above.

Accuracy was a function of serial position, $F_1(4, 108) = 154.29$, $p < .001$, $BF_{10} = 4.104e + 48$; $F_2(4, 260) = 60.82$, $p < .001$, $BF_{10} = 7.918e + 23$. Participants were substantially more accurate on the third position (.841) than on the other letter positions (.540, .504, .525, and .527, in the first, second, fourth, and fifth positions, respectively). In addition, we did not find any clear signs of a difference in the overall accuracy levels in the trained script and the Roman script (.574 vs. .601; $F_1(1, 27) = 3.14$, $p = .09$; $BF_{10} = .17$; $F_2 < 1$; $BF_{10} = .78$). Finally, as can be seen in Figure 9, the serial position functions of the Roman and trained scripts were remarkably similar and the interaction between the two factors was not significant, $F_1(4, 108) = 2.28$, $p = .09$; $BF_{10} = .09$; $F_2(4, 260) = 2.05$, $p = .09$, $BF_{10} = .31$).

Discussion

The current experiment, using a target-in-string identification task, showed an advantage of the middle, fixated position over the other positions for the trained and control scripts

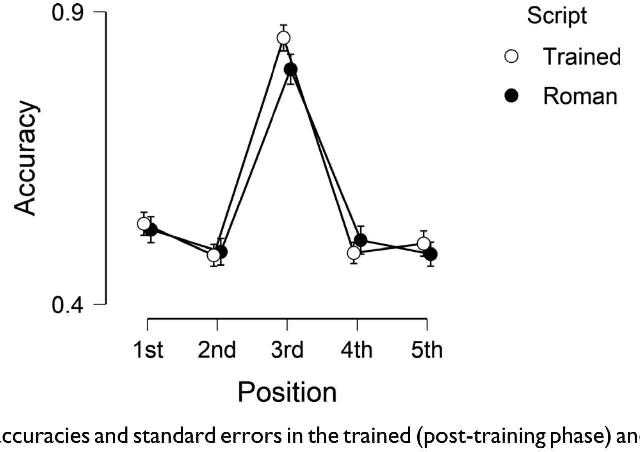

Figure 9. Mean accuracies and standard errors in the trained (post-training phase) and Roman scripts of Experiment 2.

not only in the pre-training phase, but also in the post-training phase. We did find a small numerical advantage of the initial position over the second position (see Figure 8), but this difference was similar for the trained and untrained script in the pre-/post-training phases – this pattern was also corroborated in the analyses with Bayesian linear mixed-effects models (see Appendix A). We also found that accuracy was higher in the post-training than in the pre-training session – this occurred mainly in the central, fixated letter. This effect was similar for the trained script and for the visual control script; thereby, it can be parsimoniously explained in terms of visual familiarity. Taken together, these findings strongly suggest that location-specific letter detectors do not emerge rapidly after learning-to-read and write in an artificial script.

Finally, in the exploratory analyses, we failed to find a stronger initial-letter advantage in the Roman script when compared to the trained script. We prefer to keep cautious about this latter finding. First, in the instructions, we emphasized that participants should be looking at the fixation point at the beginning of each trial. Second, because it was an exploratory analysis, participants always performed the task with Roman letters in the last block. As a result, the attentional capture at the middle position that occurred in initial blocks with the artificial scripts could have been dragged into the last block. A more comprehensive explanation is presented in the General Discussion section below.

GENERAL DISCUSSION

We designed two experiments to track the emergence of early orthographic processes in the first stages of learning-to-read through the examination of two markers of orthographic processing (Grainger, 2018): location-invariant processing (Experiment 1) and location-specific processing (Experiment 2). To that aim, we employed a design with a pre-training phase and a post-training phase in which adults were trained in reading and writing nonsense words in an artificial script along six sessions. Participants successfully mastered the trained script after the learning-to-read training (see Appendix B). All of them passed the final assessment in the prescribed time with no errors in the reading aloud and writing down tasks. To control for visual familiarity, participants were also familiarized with the visual form of the characters of the control script during the training sessions.

The emergence of location-invariant processing

The first research question was whether readers show some location-invariant processing for the newly learned script on top of the position uncertainty that may affect all visual objects in a string. As stated in the Introduction, location-invariant processing has been proposed to emerge with literacy acquisition (Dandurand *et al.*, 2010; Duñabeitia *et al.*, 2014, 2015).

To our knowledge, the only published study that directly examined this issue was conducted by Duñabeitia *et al.* (2015). They employed a longitudinal design using a same-different experiment in which a group of children was tested in their antepenultimate pre-school year (i.e., pre-literate children; mean age = 4.24 years), in their last pre-school year (i.e., pre-literate children; mean age = 5.21 years), and in the first year of primary school (i.e., they had already acquired literacy skills; mean age = 6.32 years). They used four-letter strings in which, for 'different' trials, they had transposed-letter and replaced-letter pairs. Duñabeitia *et al.* (2015) only found a transposed-letter effect when the children were in first grade (42.9% vs. 30.6% of errors, for transposed vs. replaced-letter pairs) and concluded that 'position uncertainty emerges as a consequence of literacy training' (p. 549). However, the null effects obtained with the children in their pre-school years faced interpretative difficulties because these children did not adequately perform the task (d' was close to zero; see Perea *et al.*, 2016). Indeed, Perea *et al.* (2016) found robust transposed-letter effects in pre-literate children with a simplified version of the same-different task. Nevertheless, they did not run a retest when these children learned to read – note that first graders could have been at ceiling with this simplified task, thus making the comparison uninterpretable.

To test the emergence of location-invariant processing, while avoiding the interpretive difficulties of comparing performance of pre-school vs. school children, we conducted a same-different matching task with adult participants using transposed-letter vs. replaced-letter pairs as 'different' trials before and after learning-to-read in a new script (Experiment 1). To control for mere visual familiarity, we included an untrained script that was also presented during training. Results showed transposed-letter effects of similar size for the trained and untrained scripts in both the pre- and post-training phases. Furthermore, Bayesian analyses offered substantial evidence in favour of a null interaction between training, phase, and script. Thus, learning-to-read and write in a new script does not lead to the rapid emergence of location-invariant processing.

Importantly, we did find some training benefit in the trained script for 'same' responses in both response times and accuracy, which suggests the emergence of an early and basic visual specialization for letter strings. However, the newly acquired expertise in the new script was not sufficient to induce location-invariant processing. Indeed, the transposed-letter effect was greater for the Roman script than for the trained script (i.e., 28.3% vs. 19.2% of errors, respectively). This is the typical pattern when comparing strings of letters vs. strings of other visual objects (e.g., symbols, unknown letters)[6]. This pattern can be parsimoniously explained in terms of an orthographically specific location-invariant component in the Roman script over and above the location uncertainty common to all visual objects (see Massol *et al.*, 2013).

Our findings can also shed some light on the early developmental trajectory of the letter-specific position coding when learning-to-read. The absence of the emergence of location-invariant processing in the very early stages of learning-to-read can be accommodated by the dual-route model of orthographic development proposed by Grainger and Ziegler (2011; see also Grainger *et al.*, 2012; Ziegler, Bertrand, Lété, & Grainger, 2014). This model assumes that, in the first stages of reading acquisition, the processing of letters in a word is serial, thereby letter position coding is very strict (i.e., fine-grained orthographic coding). It is only when readers have more extensive reading experience that a more parallel processing of letters is developed, thus speeding the mapping of letters onto orthographic representations and producing greater transposed-letter effects (i.e., coarse-grained orthographic coding; see Grainger *et al.*, 2012). In the context of the current experiment, participants acquired some basic orthographic skills, as revealed by better performance for 'same' responses in the post-training phase. However, this expertise did not suffice for a coarse-grained processing to emerge. Indeed, in a lexical decision experiment that compared the error rates to pseudohomophone and

---

[6] *Furthermore, for the Roman script, we found that size of the transposed-letter effect was considerably smaller for external than for internal transpositions: 14.5% vs. 42.1%, respectively (e.g., see Gomez* et al., *2008, for a similar pattern). In contrast, for the trained script, the size of the transposed-letter effect was only slightly lower for external transpositions than for internal transposition (17.0% vs. 21.4%, respectively). This again suggests that the transposed-letter effect in the Roman script and the trained script reflects different underlying processes.*

orthographic controls, Grainger *et al.* (2012) found effects greater than 30% in Grade 1 and Grade 2 children – these effects were smaller with older children (i.e., the effects were 20% in Grade 3, 21% in Grade 4 and 16% in Grade 5). That is, beginning readers use phonological recoding (i.e., a fine-grained orthographic coding) rather than the coarse-grained coding responsible for location-invariant processing. Thus, the greater transposed-letter effect in the Roman (overlearned) script than in the newly learned script obtained in the current experiments suggests that the emergence of location-invariant processing requires a more complete establishment of a written orthographic code (see Grainger *et al.*, 2012; Grainger & Ziegler, 2011; Ziegler *et al.*, 2014)[7].

The emergence of location-specific processing

The second research question was whether location-specific processing emerges rapidly for the newly learned script using a target-in-string identification task. The dissociation in the accuracy serial position functions of letters (W-shape function) and symbols (Λ-shape function) is this task is assumed to be to an adaptation of the mechanisms of visual object processing to cope with visual word processing (i.e., location-specific letter detectors creation). Importantly, Dandurand *et al.* (2010; Grainger *et al.*, 2016) hypothesized that the conversion of the mechanisms of simple visual object processing into location-specific detectors occurs with reading acquisition.

Results in the target-in-string identification task (Experiment 2) showed a clear advantage of the middle position for both the trained and control scripts in all scenarios. More critically, we found no signs of an interaction in accuracy between training, phase, and position, as shown in the Bayesian analyses. We also found a small advantage of the initial-letter position over the other letter positions (see Figure 6; see also Appendix A), but this difference was not modulated by training (i.e., the difference was approximately constant in the pre- and post-training phases and in the trained and untrained scripts). Finally, the overall accuracy in the post-training phase was greater than in the pre-training phase for both, the learned and the control script, but this occurred essentially for the middle, fixated, letter. This latter finding can be parsimoniously explained in terms of better performance due to increased visual familiarity rather than on location-specific processing.

To our knowledge, unlike for location-invariant processing, no study has directly examined the emergence of location-specific processing – neither with children nor with adults. Nevertheless, for comparison purposes, it may be relevant to briefly discuss the studies that examined the developmental trajectory of the first-letter advantage in the very early stages of learning-to-read. Grainger *et al.* (2016) showed a small increase in the initial-letter advantage across school grades (from 1st to 4th) (see also Schubert, Badcock, & Kohnen, 2017, for a similar pattern of results). Importantly, the accuracy in the first and second position for 1st- to 3rd-grade children was very similar, around 55% and 60% in the Grainger *et al.*'s (2016) experiment. The lack of a sizeable first position advantage in the initial grades in developing readers is in consonance with the results of Experiment 2. Taken together, these findings suggest that location-specific processing does not emerge

---

[7] *An alternative account of orthographic development is Castles* et al.*'s (2007) lexical tuning model. The model assumes that acquiring more and more words in the lexicon involves an increasingly dense neighbourhood of orthographically similar words. To efficiently identify these words, the orthographic representations become increasingly fine-tuned – this includes more precise positional representation of the visual input. Our experiments, however, were not designed to test the development of the orthographic lexicon (i.e., participants were trained to read and write pseudowords).*

in the first stages of learning-to-read; instead, there appears to be a long route for this mechanism to emerge and develop.

Nevertheless, the lack of the sizeable first-letter advantage in the (overlearned) Roman script suggests that some caution when interpreting the findings of Experiment 2. In the experimental set-up, participants always received the blocks with the artificial scripts (i.e., the main blocks for the purposes of the experiment) before the final block with the Roman script, and furthermore, instructions stressed that they should fixate at the centre position at the beginning of each trial. Thus, a parsimonious explanation of the lack of a substantial first-letter advantage in the Roman script is that, to cope with the highly demanding blocks with the artificial scripts, participants' attention was focused on the middle position and this strategy was dragged into the Roman block. Thus, one might argue that the settings of Roman block were not optimal to capture a W-serial position function in the Roman block. Future research should examine to what degree the accuracy function in target-in-string identification tasks is modulated by task context and instructions (e.g., see Winskel *et al.*, 2014, for evidence of different accuracy functions depending on the nature of the writing system).

On the emergence of orthographic processing when learning-to-read a new script

Recent research with adult readers has shown that orthographic processes can emerge rapidly after learning a script during a relatively short amount of time (e.g., Chetail, 2017; Lally, Taylor, Lee & Rastle, 2020; Taylor *et al.*, 2011). For instance, in the Taylor *et al.*'s (2011) experiments, adults learned 36 words in an artificial script during 30–45 min. In the post-training phase, participants had to discriminate between trained and untrained items (i.e., an analogue to lexical decision). Results showed that participants could successfully discriminate trained from untrained items and, more important, the response times to the trained items were sensitive to vowel frequency. In a subsequent generalization phase, participants were asked to read aloud a series of new (untrained) items. Results showed an effect of both vowel frequency and consistency. All and all, the Taylor *et al.* (2011) experiments suggest that participants can quickly and efficiently extract sub-word spelling–sound regularities in a new script (see Chetail, 2017, for a similar pattern of results regarding letter and bigram frequency).

More recently, Lally *et al.* (2020) conducted an experiment in which participants learned 24 five-letter pseudowords either in a sparse or in a dense artificial orthography, using a between-subject design, during a four-day training. (The 24 pseudowords in the dense orthography included 12 anagram pairs, whereas none of the 24 pseudowords sparse included anagrams.) When tested in an old–new recognition task (i.e., they had to discriminate between trained and untrained items), participants made fewer false positives for untrained items created by transposing two letters in the dense orthography than in the sparse orthography. Therefore, the findings reported by Lally *et al.* (2020) are a demonstration that the properties of the writing systems may modulate how letter order is encoded in a newly learned script.

In the current experiments, participants were able to read and write in the trained script with some fluency. Notably, in line with above-cited studies with artificial script training, we found some letter-specific processing as a consequence of learning-to-read: responses to 'same' trials in the same-different task were faster and more accurate in the post-trained phase for the trained script, but not for the untrained script. As Krueger (1978) indicated, fast and accurate responses for 'same' trials imply that participants require an exhaustive processing of the letter string. Thus, this pattern suggests that early literacy induced some specialization for letter strings that made the identification of same pairs more effortless. However, we found no signs reflecting the emergence of location-invariant and location-specific processing in the newly learned script.

In addition, we found an improvement in performance in the post-training phase for both the learned script and the visual control script that can be parsimoniously explained in terms of visual familiarity. Indeed, previous event-related potentials (ERP) experiments have shown that the N1 component (i.e., a component related to familiar written language processing) emerges right-lateralized in preschoolers at the start of reading training (e.g., Maurer *et al.*, 2006). Crucially, this early emergence of the N1 effect has been associated with letter knowledge and it also likely reflects visual familiarity with print (Maurer, Brem, Bucher, & Brandeis, 2005). The development of the characteristic left-lateralization of the component is assumed to occur via the automatization of orthographic-phonological mappings established during learning-to-read (Maurer & McCandliss, 2007; McCandliss & Noble, 2003; Maurer *et al.*, 2006; Posner & McCandliss, 2000; see also Maurer *et al.*, 2010, for evidence with adults learning an artificial script; Brem et al., 2005, for the same pattern with visual training of symbols). Thus, the increase in performance in the post-training phase, coupled with the absence of differences in location-invariant and location-specific processing between the two phases, favours the idea that these effects are associated with visual expertise with the novel scripts (i.e., a reading-related perceptual expertise; see Maurer *et al.*, 2010, for similar claims).

What we should also note is that, although both the same/different matching task and target-in-string identification task have been widely used to demonstrated orthographic effects (e.g., see Duñabeitia *et al.*, 2012; Massol *et al.*, 2013; Perea *et al.*, 2016; Scaltriti & Balota, 2013; Tydgat & Grainger, 2009), they can be performed on the basis of visual representations of the stimuli – this is the reason why these tasks can be used in both pre-training and post-training phases. As a result, some effects obtained from newly learned scripts may reflect a mixture of increased visual familiarity to the new letters together with some incipient orthographic representations (i.e., a specific to letters visual expertise, see Maurer, Brandeis, & McCandliss, 2005). Instead, skilled readers, who have already automatized the orthographic-phonological mappings, would perform these tasks not only on the basis of visual familiarity, but also on the activation of the orthographic representations of the stimuli. In other words, the patterns observed in beginner readers may rely on visual familiarization with the learned scripts, whereas the effects of skilled readers may represent an interaction between early visual processing at the letter level and feedback from orthographic representations (see Marinus *et al.*, 2018).

Taken together, our findings suggest that in order to boost the automatization of the orthographical–phonological mappings and a more parallel coarse-coding processing, a much longer learning-to-read period may be required. We now discuss several options for further research. The first option would be to run a large-scale longitudinal experiment with pre-literate children – for the sake of the argument, we assume that the experimental tasks would allow a meaningful comparison across age (see Perea *et al.*, 2016, for discussion). The experimental design would include three scripts: Roman letters, digits (i.e., a familiar visual object), and a control artificial script. This would allow us to examine not only the emergence of location-invariant and location-specific processing in a natural setting (i.e., children learning-to-read), but also how these orthographic markers vary as a consequence of literacy acquisition. Furthermore, this design would also allow examining the variations due to orthographic processing (i.e., specific to letters) vs. visual familiarity (numbers vs. artificial letters). A second option would be to train adult participants for a long period of time in an ecological setting – instead of the ecological limitations of learning an artificial script. The most realistic design would be run the experiment with adults who are starting to learn a new language that uses an unfamiliar alphabetic script (e.g., Georgian, Armenian). In either scenario, it would be desirable to complement the behavioural tasks with the recording of brain activity during training, as this may help to disentangle the effects due to orthographical–phonological decoding from the effects due to visual training (e.g., see Maurer, Brem, *et al.*, 2005, 2006; Pleisch *et al.*, 2019, for evidence of print sensitivity in the N1 ERP amplitude; see Pleisch *et al.*, 2019, for evidence of changes in the activation of crucial orthographic processing brain regions [ventral occipitotemporal cortex and left fusiform gyrus] in the first steps of reading acquisition).

To sum up, we conducted two experiments that examined whether two markers of orthographic processing (location-invariant and location-specific processing) arise rapidly after learning-to-read and write a new script. Notably, examining when these effects emerge is essential to help interpret the subsequent developmental trajectory of orthographic effects. While participants were able to read and write with some fluency in the new script and showed some rudimentary orthographic processing, we found no evidence favouring the hypotheses that location-invariance and location-specific processing emerge quickly after learning-to-read. Instead, the emergence of these two markers of orthographic processing may take much more time, probably via the automatization of orthographic-phonological mappings.

Conflicts of interest

The authors declare no conflict of interest.

Funding

This study was supported by the Spanish Ministry of Science, Innovation, and Universities (PRE2018-083922, PSI2017-86210-P) and by the Department of Innovation, Universities, Science and Digital Society of the Valencian Government (GV/2020/074)

Author contribution

María Fernández-López (Conceptualization; Methodology; Writing – original draft) Ana Marcet (Conceptualization; Methodology; Writing – original draft) Manuel Perea, PhD (Conceptualization; Funding acquisition; Methodology; Writing – original draft)

Data Availabiity Statement

All the training materials, data and script files are available on the following OSF site: https://osf.io/um6rw/?view_only=7d4754bbb5f445adb5e34530162ba552

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2019). lme4: Linear Mixed-Effects Models using 'Eigen' and S4 (1.1-20) [software]. Retrieved from https://CRAN.R-project.org/package=lme4.

Brem, S., Hunkeler, E., Mächler, M., Kronschnabel, J., Karipidis, I. I., Pleisch, G., & Brandeis, D. (2018). Increasing expertise to a novel script modulates the visual N1 ERP in healthy adults. *International Journal of Behavioral Development*, *42*, 333–341. https://doi.org/10.1177/0165025417727871

Brem, S., Lang-Dullenkopf, A., Maurer, U., Halder, P., Bucher, K., & Brandeis, D. (2005). Neurophysiological signs of rapidly emerging visual expertise for symbol strings. *Neuroreport*, *16*, 45–48. https://doi.org/10.1097/00001756-200501190-00011

Bürkner, P. C. (2016). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.

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*, 165–182. https://doi.org/10.1016/j.jecp.2007.01.006

Chanceaux, M., & Grainger, J. (2012). Serial position effects in the identification of letters, digits, symbols, and shapes in peripheral vision. *Acta Psychologica*, *141*, 149–158. https://doi.org/10.1016/j.actpsy.2012.08.001

Chanceaux, M., Mathôt, S., & Grainger, J. (2014). Effects of number, complexity, and familiarity of flankers on crowded letter identification. *Journal of Vision*, *14*, 1–17. https://doi.org/10.1167/14.6.7

Chetail, F. (2017). What do we do with what we learn? Statistical learning of orthographic regularities impacts written word processing. *Cognition*, *163*, 103–120. https://doi.org/10.1016/j.cognition.2017.02.015

R Core Team (2019). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.

Dandurand, F., Grainger, J., & Dufau, S. (2010). Learning location-invariant orthographic representations for printed words. *Connection Science*, *22*, 25–42. https://doi.org/10.1080/09540090903085768

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*, 713–758. https://doi.org/10.1037/a0019738

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, *9*, 335–341. https://doi.org/10.1016/j.tics.2005.05.004

Duñabeitia, J. A., Dimitropoulou, M., Grainger, J., Hernández, J. A., & Carreiras, M. (2012). Differential sensitivity of letters, numbers, and symbols to character transpositions. *Journal of Cognitive Neuroscience*, *24*, 1610–1624. https://doi.org/10.1162/jocn_a_00180

Duñabeitia, J. A., Lallier, M., Paz-Alonso, P. M., & Carreiras, M. (2015). The impact of literacy on position uncertainty. *Psychological Science*, *26*, 548–550. https://doi.org/10.1177/0956797615569890

Duñabeitia, J. A., Orihuela, K., & Carreiras, M. (2014). Orthographic coding in illiterates and literates. *Psychological Science*, *25*, 1275–1280. https://doi.org/10.1177/0956797614531026

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods*, *35*, 116–124. https://doi.org/10.3758/BF03195503

García-Orza, J., Perea, M., & Muñoz, S. (2010). Are transposition effects specific to letters? *The Quarterly Journal of Experimental Psychology*, *63*, 1603–1618. https://doi.org/10.1080/17470210903474278

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*, 577–600. https://doi.org/10.1037/a0012667

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, *23*, 1–35. https://doi.org/10.1080/01690960701578013

Grainger, J. (2018). Orthographic processing: A 'mid-level' vision of reading: The 44th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology*, *71*, 335–359. https://doi.org/10.1080/17470218.2017.1314515

Grainger, J., Bertrand, D., Lété, B., Beyersmann, E., & Ziegler, J. C. (2016). A developmental investigation of the first-letter advantage. *Journal of Experimental Child Psychology*, *152*, 161–172. https://doi.org/10.1016/j.jecp.2016.07.016

Grainger, J., & Hannagan, T. (2014). What is special about orthographic processing? *Written Language & Literacy*, *17*, 225–252. https://doi.org/10.1075/wll.17.2.03gra

Grainger, J., Lété, B., Bertrand, D. D. S., & Ziegler, J. (2012). Evidence for multiple routes in learning to read. *Cognition*, *123*, 280–292. https://doi.org/10.1016/j.cognition.2012.01.003

Grainger, J., Tydgat, I., & Isselé, J. (2010). Crowding affects letters and symbols differently. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 673–688. https://doi.org/10.1037/a0016888

Grainger, J., & van Heuven, W. J. B. (2004). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 1–23). Hauppauge, NY: Nova Science Publishers.

Grainger, J., & Ziegler, J. (2011). A dual-route approach to orthographic processing. *Frontiers in psychology*, *2*, 54. https://doi.org/10.3389/fpsyg.2011.00054

Krueger, L. E. (1978). A theory of perceptual matching. *Psychological Review*, *85*, 278–304. https://doi.org/10.1037//0033-295X.85.4.278

Lally, C., Taylor, J. S. H., Lee, C. H., & Rastle, K. (2020). Shaping the precision of letter position coding by varying properties of a writing system. *Language, Cognition and Neuroscience*, *35*, 374–382. https://doi.org/10.1080/23273798.2019.1663222

Leinenger, M., Myslín, M., Rayner, K., & Levy, R. (2017). Do resource constraints affect lexical processing? Evidence from eye movements. *Journal of Memory and Language*, *93*, 82–103. https://doi.org/10.1016/j.jml.2016.09.002

Marcet, A., Perea, M., Baciero, A., & Gomez, P. (2019). Can letter position encoding be modified by visual perceptual elements? *Quarterly Journal of Experimental Psychology*, *72*, 1344–1353. https://doi.org/10.1177/1747021818789876

Marinus, E., Kezilas, Y., Kohnen, S., Robidoux, S., & Castles, A. (2018). Who are the noisiest neighbors in the hood? Using error analyses to study the acquisition of letter-position processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1384–1396. https://doi.org/10.1037/xlm0000524

Massol, S., Duñabeitia, J. A., Carreiras, M., & Grainger, J. (2013). Evidence for letter-specific position coding mechanisms. *PLoS One*, *8*, e68460. https://doi.org/10.1371/journal.pone.0068460

Maurer, U., Blau, V. C., Yoncheva, Y. N., & McCandliss, B. D. (2010). Development of visual expertise for reading: rapid emergence of visual familiarity for an artificial script. *Developmental Neuropsychology*, *35*, 404–422. https://doi.org/10.1080/87565641.2010.480916

Maurer, U., Brandeis, D., & McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions*, *1*, 13. https://doi.org/10.1186/1744-9081-1-13

Maurer, U., Brem, S., Bucher, K., & Brandeis, D. (2005). Emerging neurophysiological specialization for letter strings. *Journal of Cognitive Neuroscience*, *17*, 1532–1552. https://doi.org/10.1162/089892905774597218

Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., . . . Brandeis, D. (2006). Coarse neural tuning for print peaks when children learn to read. *NeuroImage*, *33*, 749–758. https://doi.org/10.1016/j.neuroimage.2006.06.025

Maurer, U., & McCandliss, B. D. (2007). The development of visual expertise for words: the contribution of electrophysiology. In E. L. Grigorenko & A. J. Naples (Eds.), *Single-word reading: cognitive, behavioral and biological perspectives* (pp. 43–64). Mahwah, NJ: Lawrence Erlbaum Associates.

McCandliss, B. D., & Noble, K. G. (2003). The development of reading impairment: a cognitive neuroscience model. *Mental retardation and developmental disabilities research reviews*, *9*, 196–205. https://doi.org/10.1002/mrdd.10080

Morey, R. D., & Rouder, J. N. (2018). Package 'BayesFactor'. R Package version 0.9.12-4.2. Retrieved from https://richarddmorey.github.io/BayesFactor/.

Muñoz, S., Perea, M., García-Orza, J., & Barber, H. A. (2012). Electrophysiological signatures of masked transposition priming in a same-different task: Evidence with strings of letters vs. pseudoletters. *Neuroscience Letters*, *515*, 71–76. https://doi.org/10.1016/j.neulet.2012.03.021

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review*, *119*, 517–545. https://doi.org/10.1037/a0028450

Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, *62*, 254–271. https://doi.org/10.1016/j.jml.2009.11.002

Perea, M., Jiménez, M., & Gomez, P. (2016). Does location uncertainty in letter position coding emerge because of literacy training? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 996–1001. https://doi.org/10.1037/xlm0000208

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*, 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Pleisch, G., Karipidis, I. I., Brauchli, C., Röthlisberger, M., Hofstetter, C., Stämpfli, P., . . . Brem, S. (2019). Emerging neural specialization of the ventral occipitotemporal cortex to characters through phonological association learning in preschool children. *NeuroImage*, *189*, 813–831. https://doi.org/10.1016/j.neuroimage.2019.01.046

Posner, M., & McCandliss,B. D. (2000). Brain circuitry during reading. In R. Klein & P. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 305-337). Cambridge: MIT Press.

Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*, 246–253. https://doi.org/10.1016/j.tics.2011.04.001

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*, 552–572. https://doi.org/10.1037/0033-295X.88.6.552

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 275–280. https://doi.org/10.1037/h0027768

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237. https://doi.org/10.3758/PBR.16.2.225

Scaltritti, M., & Balota, D. A. (2013). Are all letters really processed equally and in parallel? Further evidence of a robust first letter advantage. *Acta Psychologica*, *144*, 397–410. https://doi.org/10.1016/j.actpsy.2013.07.018

Scaltritti, M., Dufau, S., & Grainger, J. (2018). Stimulus orientation and the first-letter advantage. *Acta Psychologica*, *183*, 37–42. https://doi.org/10.1016/j.actpsy.2017.12.009

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*, 128–142. https://doi.org/10.3758/s13423-017-1230-y

Schubert, T., Badcock, N., & Kohnen, S. (2017). Development of children's identity and position processing for letter, digit, and symbol strings: A cross-sectional study of the primary school years. *Journal of Experimental Child Psychology*, *162*, 163–180. https://doi.org/10.1016/j.jecp.2017.05.008

Staub, A., & Goddard, K. (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 110–127. https://doi.org/10.1037/xlm0000561

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*, 826–858. https://doi.org/10.1037/xge0000301

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*, 60–76. https://doi.org/10.1037/a0020126

Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 480–498. https://doi.org/10.1037/a0013027

Vejnović, D., & Zdravković, S. (2015). Side flankers produce less crowding, but only for letters. *Cognition*, *143*, 217–227. https://doi.org/10.1016/j.cognition.2015.07.003

Vidal, C., & Chetail, F. (2017). BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior Research Methods*, *49*, 2093–2112. https://doi.org/10.3758/s13428-016-0844-8

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., . . . Meerhoff, F. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*, 58–76. https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., . . . Morey, R. D. (2017). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*, 35–57. https://doi.org/10.3758/s13423-017-1343-3

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*, 59–85. https://doi.org/10.1016/0010-0285(70)90005-8

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin and Review*, *8*, 221–243. https://doi.org/10.3758/BF03196158

Winskel, H., Perea, M., & Peart, E. (2014). Testing the flexibility of the modified receptive field (MRF) theory: Evidence from an unspaced orthography (Thai). *Acta Psychologica*, *150*, 55–60. https://doi.org/10.1016/j.actpsy.2014.04.008

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*, 1026–1036. https://doi.org/10.1037/a0035187

Ziegler, J. C., Pech-Georgel, C., Dufau, S., & Grainger, J. (2010). Rapid processing of letters, digits and symbols: what purely visual-attentional deficit in developmental dyslexia? *Developmental Science*, *13*, F8–F14. https://doi.org/10.1111/j.1467-7687.2010.00983.x

Received 26 September 2019; revised version received 18 June 2020

Appendix A:

Supplementary analyses with Bayesian linear mixed-effects models

Experiment 1: location-invariant processing

For the sake of completeness, we also examined the latency and accuracy data of the confirmatory analyses using Bayesian linear mixed-effects models using the brms package in R (Bürkner, 2016). An advantage of this procedure – via Stan – is that it allows us to fit the models using the maximal random-effect structure (see Bates, Mächler, Bolker & Walker, 2015, for arguments in favour of maximal models).

Different trials

Trained script vs. Untrained script. The fitted model was: Dependent_Variable [i.e., −1000/RT or accuracy] ~ script * condition * phase + (1 + script * condition * phase | subject) + (1 + condition * phase | item). More complex random-effects terms resulted in model non-convergence. Furthermore, these models offer the Bayesian 95% credible intervals for each parameter based on the posterior distributions. For the latency data, we employed the same response time transformation as in LME analyses (i.e., −1000/RT; family = gaussian), whereas for the accuracy data, we used the Bernoulli distribution (family = bernoulli) – this is the parallel to family = binomial in GLME models. We

Table A1. Parameter estimates in latency and accuracy supplemental analyses of Experiment 1 ('different' trials: trained vs. untrained script)

Latency data	Estimate	SE	95% Credible Interval
Intercept	−1.84	0.06	[−1.95, −1.73]
Script	−0.01	0.02	[−0.05, 0.02]
Condition	**0.08**	**0.01**	**[0.05, 0.10]**
Phase	**0.13**	**0.04**	**[0.06, 0.21]**
Script × Condition	−0.01	0.02	[−0.05, 0.03]
Script × Phase	−0.05	0.04	[−0.12, 0.03]
Condition × Phase	−0.01	0.02	[−0.05, 0.03]
Script × Condition × Phase	0.01	0.03	[−0.04, 0.06]
Accuracy data			
Intercept	1.56	0.15	[1.27, 1.87]
Script	−0.07	0.10	[−0.28, 0.13]
Condition	**−1.06**	**0.11**	**[−1.28, −0.84]**
Phase	**−0.35**	**0.12**	**[−0.60, −0.10]**
Script × Condition	0.07	0.12	[−0.17, 0.31]
Script × Phase	−0.04	0.13	[−0.29, 0.21]
Condition × Phase	0.12	0.13	[−0.13, 0.37]
Script × Condition × Phase	0.03	0.17	[−0.29, 0.36]

Note. Those effects with 95% credible intervals beyond 0 are in bold.

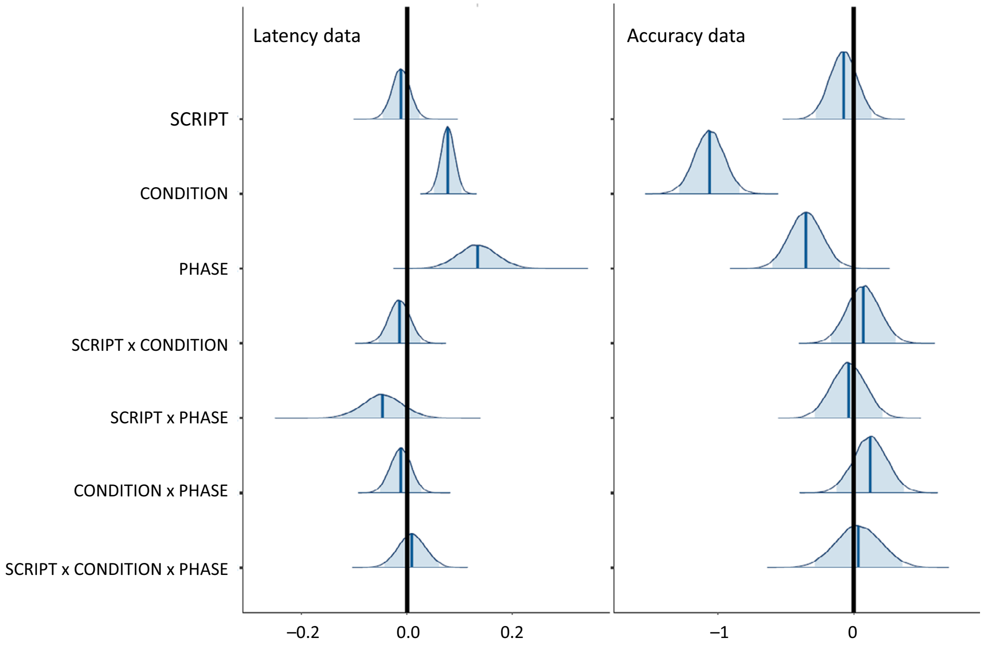

Figure A1. Posterior effects estimates from the Bayesian linear mixed models for 'different' trials in Experiment 1 (Trained vs. Untrained script) (left panel: latency analysis; right panel: accuracy analysis). The thick black line corresponds to an effect of zero, the dark grey line corresponds to the estimates, and the shaded area corresponds to the 95% credible interval. [Colour figure can be viewed at wileyonlinelibrary.com]

Table A2. Parameter estimates in latency and accuracy supplemental exploratory analyses of Experiment 1 ('different' trials: Roman vs. trained script)

Latency data	Estimate	SE	95% Credible Interval
Intercept	−1.83	0.04	[−1.91, −1.74]
Condition	0.07	0.02	[0.04, 0.11]
Script	−0.02	0.05	[−0.13, 0.09]
Condition × Script	0.00	0.02	[−0.04, 0.05]
Accuracy data			
Intercept	1.86	0.17	[1.55, 2.20]
Condition	−1.66	0.14	[−1.92, −1.39]
Script	−0.26	0.21	[−0.69, 0.16]
Condition × Script	0.56	0.17	[0.23, 0.90]

Note. Those effects with 95% credible intervals beyond 0 are in bold.

employed 4 chains, each with 10,000 iterations after a warm-up of 1000 iterations. The maximal random-effect structure models converged successfully: the values of Rhat were 1.00 for all parameters.

As can be seen from the estimates and 95% credible intervals presented in Table A1, we found robust evidence of an effect of Condition (i.e., probe-target relationship) and Phase in both latency and accuracy analyses, thus corroborating the pre-registered

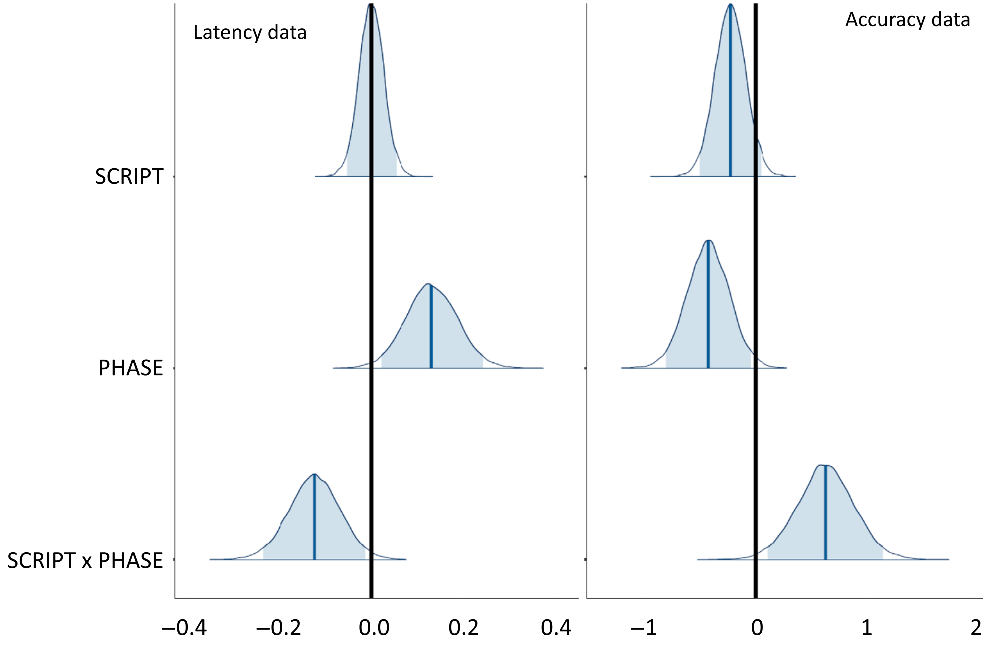

Figure A2. Posterior effects estimates from the Bayesian linear mixed models in 'different' trials in Experiment 1 (Roman vs. Trained script) (left panel: latency analysis; right panel: accuracy analysis). The thick black line corresponds to an effect of zero, the dark grey line corresponds to the estimate, and the shaded area corresponds to the 95% credible interval. [Colour figure can be viewed at wileyonlinelibrary.com] analyses. Similarly, we did not find any evidence of a three-way interaction between Script, Condition, and Phase (see Figure A1, for the posterior distributions).

Roman script vs. trained script

The fixed factors were Script (Trained [post-training] vs. Roman) and Condition (transposed, replaced). We followed the same procedure as above, with 5,000 iterations (Rhat = 1.00 in all cases). Table A2 shows the estimates and 95% credible intervals from the latency and accuracy models, and Figure A2 shows the posterior distributions of the parameters. Together with a substantial transposed-letter effect in the latency data, these analyses confirmed the interaction between Script and Condition in accuracy data.

Same trials

Trained script vs. Untrained script. The general procedure was the same as above (i.e., the maximal random-effect structure model with 5,000 iterations; Rhat = 1.00 in all cases). As shown in Table A3 (estimates and 95% credible intervals) and Figure A3 (posterior distributions of the parameters), these analyses corroborated the interaction effect for between Phase and Script in the latency and accuracy data.

Experiment 2: location-specific processing

Trained vs. Untrained script

As in Experiment 1, we examined the data with Bayesian linear mixed-effects models using the brms package (Bürkner, 2016) in R. The three fixed factors were the same as in the pre-registered analyses. The initial letter was set as the reference level for the factor Position. We fitted the maximal random-effect structure model (i.e., accuracy ~ script * position * phase + (1 + script * position * phase | subject) + (1 + position * phase | item)) using 4 chains, each with 5,000 iterations after a warm-up of 1,000 iterations. The priors were the same as in Experiment 1. The model converged successfully (Rhat = 1.00 for all parameters).

As can be seen in Table A4, accuracy in the initial-letter position was substantially lower than in the middle, fixated position. The accuracy advantage of the middle position

Table A3. Parameter estimates in latency and accuracy supplemental analyses of Experiment 1 ('same' trials: trained vs. untrained script)

Latency data	Estimate	SE	95% Credible Interval
Intercept	−1.99	0.08	[−2.14, −1.84]
Script	0.00	0.03	[−0.05, 0.05]
Phase	0.13	0.06	[0.02, 0.24]
Script × Phase	−0.13	0.06	[−0.24, −0.01]
Accuracy data			
Intercept	2.70	0.21	[2.31, 3.12]
Script	−0.23	0.14	[−0.51, 0.05]
Phase	−0.43	0.19	[−0.81, −0.05]
Script × Phase	0.64	0.26	[0.11, 1.15]

Note. Those effects with 95% credible intervals beyond 0 are in bold.

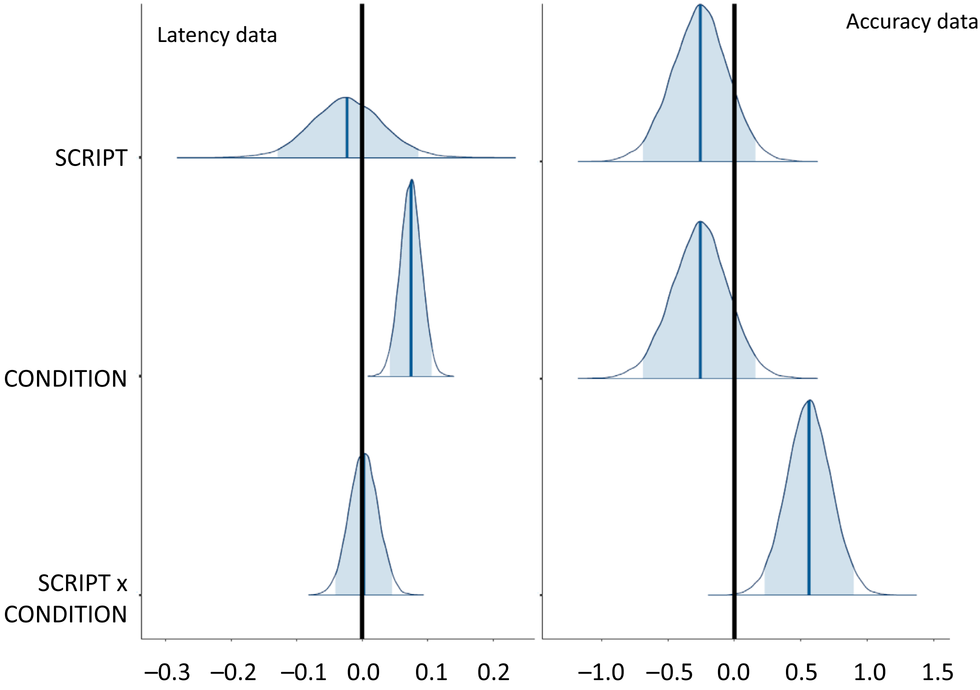

Figure A3. Posterior effects estimates from the Bayesian linear mixed models in 'same' trials of Experiment 1. The thick black line corresponds to an effect of zero, the dark grey line corresponds to the estimates, and the shaded area corresponds to the 95% credible interval. [Colour figure can be viewed at wileyonlinelibrary.com]

Table A4. Parameter estimates in the accuracy supplemental analyses of Experiment 2 (trained vs. untrained script)

		Estimate	SE	95% Credible Interval
Intercept		0.16	0.09	[−0.03, 0.34]
Script		−0.05	0.13	[−0.29, 0.20]
Position	2nd	−0.22	0.14	[−0.50, 0.06]
	3rd	**1.89**	**0.26**	**[1.38, 2.41]**
	4th	−0.20	0.14	[−0.48, 0.07]
	5th	−0.14	0.14	[−0.41, 0.13]
Phase		0.04	0.14	[−0.22, 0.31]
Script x	2nd position	0.07	0.18	[−0.29, 0.42]
	3rd position	−0.36	0.22	[−0.80, 0.08]
	4th position	0.13	0.18	[−0.22, 0.49]
	5th position	−0.03	0.18	[−0.39, 0.33]
Script × Phase		−0.06	0.18	[−0.41, 0.29]
Phase ×	2nd position	0.07	0.20	[−0.33, 0.46]
	3rd position	**−0.80**	**0.29**	**[−1.38, −0.22]**
	4th position	−0.03	0.20	[−0.42, 0.37]
	5th position	0.09	0.20	[−0.29, 0.49]
Script × Phase ×	2nd position	0.08	0.25	[−0.41, 0.58]
	3rd position	0.23	0.31	[−0.38, 0.83]
	4th position	0.21	0.25	[−0.28, 0.71]
	5th position	−0.02	0.26	[−0.52, 0.48]

Note. Those effects with 95% credible intervals beyond 0 are in bold.

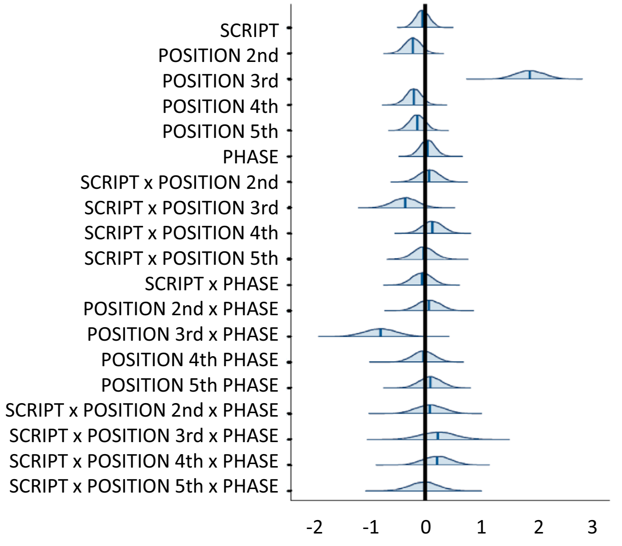

Figure A4. Posterior effects estimates from the Bayesian linear mixed models in Experiment 2. The thick black line corresponds to an effect of zero, the dark grey line corresponds to the estimate, and the shaded area corresponds to the 95% credible interval. [Colour figure can be viewed at wileyonlinelibrary.com]

increased in the post-trained phase, as deduced from the interaction with Phase. In addition, the parameter estimates showed some advantage of the initial position over the other letter positions (see Figure A4, for the posterior effects). Therefore, these analyses corroborate the findings obtained in the ANOVAs indicated in the pre-registered analyses.

Roman script vs. learned script. The procedure was the same as above, except that the two fixed factors were Script (Roman vs. Trained) and Position – the reference level for the factor Position was also the first letter. The maximal model converged successfully

Table A5. Parameter estimates in the accuracy supplemental analyses of Experiment 2 (Roman script vs. learned script)

		Estimate	SE	95% Credible Interval
Intercept		0.15	0.10	[−0.04, 0.35]
Script		0.01	0.14	[−0.25, 0.29]
Position	2nd	−0.22	0.15	[−0.51, 0.06]
	3rd	**1.90**	**0.28**	**[1.38, 2.47]**
	4th	−0.20	0.14	[−0.49, 0.08]
	5th	−0.14	0.15	[−0.43, 0.15]
Script x	2nd position	0.15	0.21	[−0.25, 0.55]
	3rd position	−0.14	0.33	[−0.77, 0.53]
	4th position	0.29	0.20	[−0.11, 0.69]
	5th position	0.18	0.21	[−0.24, 0.59]

Note. Those effects with 95% credible intervals beyond 0 are in bold.

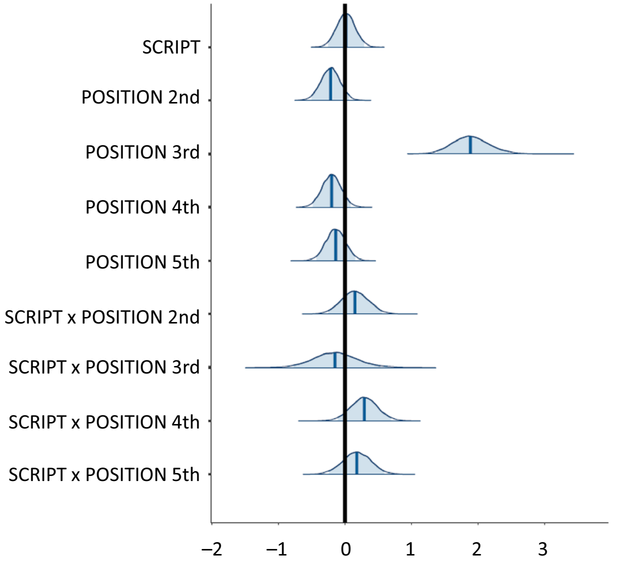

Figure A5. Posterior effects estimates from the Bayesian linear mixed models in Experiment 2 (Roman vs. Trained script). The thick black line corresponds to an effect of zero, the dark grey line corresponds to the estimate, and the shaded area corresponds to the 95% credible interval. [Colour figure can be viewed at wileyonlinelibrary.com]

(Rhat = 1.00 for all parameters). As can be seen from the 95% credible intervals (see Table A5; see also Figure A5, for the posterior distributions), we found a substantial advantage of the third letter position. In addition, there was a numerical advantage of the first-letter position relative to the second letter position – this effect did not interact with script.

Appendix B:

Figures B1 and B2 provide a visual representation of how training improved participants' performance along the learning days (from day 2 to day 6 – note that, on day 1, participants only had to listen to and write down the phoneme–grapheme correspondences).

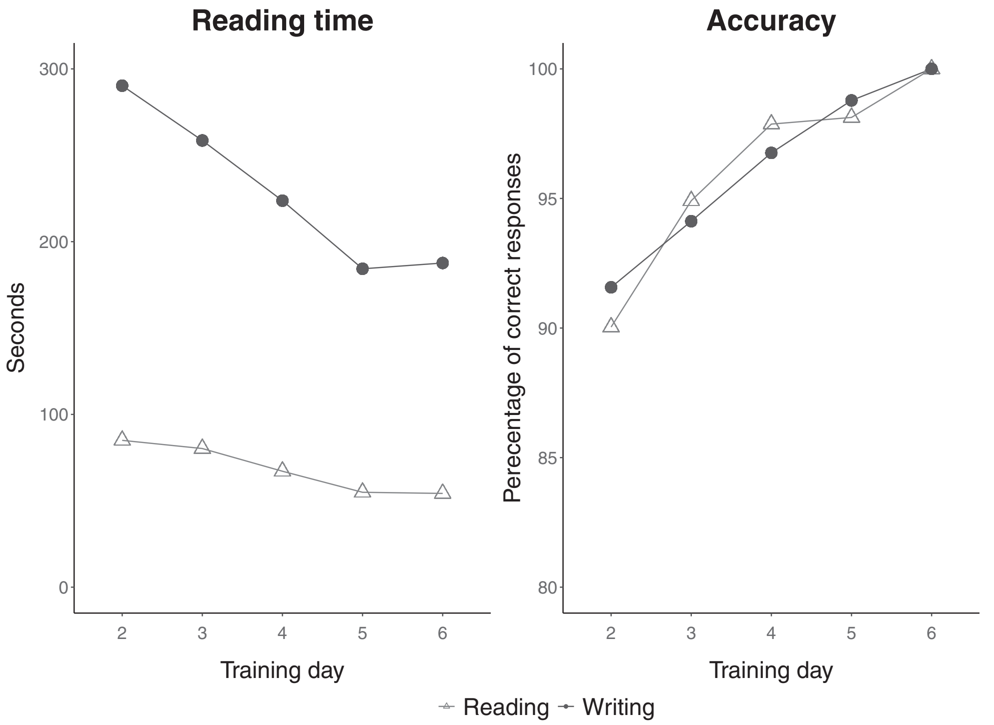

Figure B1. Participants performance on the reading aloud and writing tasks along the training days (from Day 2 to Day 6).

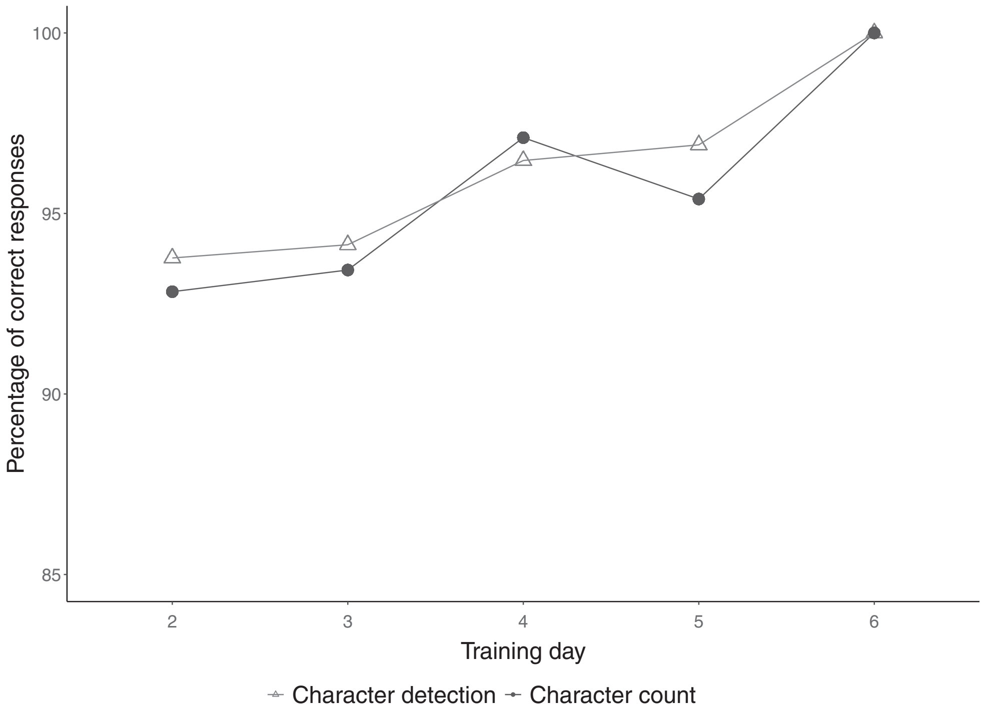

Figure B2. Participants performance on the visual familiarization tasks along the training days (from Day 2 to Day 6).