AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 2, 293–306
https://doi.org/10.1037/xge0001495

The Development of Orthography and Phonology Coupling in the Ventral Occipito-Temporal Cortex and Its Relation to Reading

Agnieszka M. Dębska[1], Jin Wang[2, 3], Gabriela K. Dzięgiel-Fivet[1], Katarzyna M. Chyl[1, 4], Marta P. Wójcik[1], Katarzyna M. Jednoróg[1], and James R. Booth[2]

[1] Laboratory of Language Neurobiology, Nencki Institute of Experimental Biology, Polish Academy of Sciences
[2] Department of Psychology and Human Development, Vanderbilt University
[3] Graduate School of Education, Harvard University
[4] International Projects Unit, The Educational Research Institute

The left ventral occipito-temporal (lvOT) cortex is considered to house the brain's representation of orthography (i.e., the spelling patterns of words). Because letter-sound coupling is crucial in reading, we investigated the engagement of the lvOT cortex in processing phonology (i.e., the sound patterns of words) as a function of reading acquisition. We tested 47 Polish children both at the beginning of formal literacy instruction and 2 years later. During functional magnetic resonance imaging, children performed auditory phonological tasks from small to large grain size levels (i.e., single phoneme, rhyme). We showed that orthographically relevant lvOT areas activated during small-grain size phonological tasks were skill-dependent, perhaps due to the relatively transparent mappings between orthography and phonology in Polish. We also studied activation pattern similarity between processing visual and auditory word stimuli in the lvOT. We found that a higher similarity level was observed in the anterior lvOT compared to the posterior lvOT after 2 years of schooling. This is consistent with models proposing a posterior-to-anterior shift in word processing during reading acquisition. We argue that the development of orthography-phonology coupling at the brain level reflects writing system-specific effects and a more universal pathway of the left vOT development in reading acquisition.

Public Significance Statement
Our study showed that reading acquisition in childhood is associated with the ability to form multimodal orthographic–phonological associations in processing auditory words. This result advances our knowledge of the neurodevelopmental pathways of reading development.

Keywords: phonology, orthography, functional magnetic resonance imaging, the left fusiform, reading

Supplemental materials: https://doi.org/10.1037/xge0001495.supp

When children learn how to read in alphabetic orthographies, they need to establish associations between orthographic and phonological representations. To facilitate this mapping, children need to recognize shared grain sizes (i.e., phonemes, rhyme, whole word) between the visual and auditory systems (Ziegler & Goswami, 2005). This constitutes a basis for effective decoding at smaller

This article was published Online First November 2, 2023.
Agnieszka M. Dębska https://orcid.org/0000-0003-0848-1446
This work was supported by the National Science Center (Poland, Grant 2019/35/D/HS6/01677) and the Kosciuszko Foundation (to Agnieszka M. Dębska). Part of the data and results were presented at the Neuronus 2022: The International Brain Research Organisation and The International Research Universities Network Neuroscience Forum and 4th iWORDD—International Workshop on Reading and Developmental Dyslexia 2023.

Agnieszka M. Dębska served as lead for conceptualization, data curation, formal analysis, methodology, visualization, writing–original draft, and writing–review and editing. Jin Wang contributed equally to data curation, formal analysis, software, and writing–original draft and served in a supporting role for conceptualization. Gabriela K. Dzięgiel-Fivet served in a supporting role for conceptualization, data curation, formal analysis, methodology, project administration, software, visualization, and writing–original draft. Katarzyna M. Chyl served in a supporting role for conceptualization, data curation, and writing–original draft. Marta P. Wójcik served in a supporting role for data curation, investigation, and project administration. Katarzyna M. Jednoróg served as lead for funding acquisition, project administration, resources, software, and supervision and served in a supporting role for conceptualization, data curation, formal analysis, investigation, validation, and writing–original draft. James R. Booth served as lead for supervision, contributed equally to conceptualization, and served in a supporting role for data curation, formal analysis, methodology, resources, visualization, and writing–original draft. Agnieszka M. Dębska, Gabriela K. Dzięgiel-Fivet, and Katarzyna M. Chyl contributed equally to investigation. Jin Wang, Katarzyna M. Chyl, and Katarzyna M. Jednoróg contributed equally to methodology. Jin Wang, Gabriela K. Dzięgiel-Fivet, Katarzyna M. Chyl, Marta P. Wójcik, Katarzyna M. Jednoróg, and James R. Booth contributed equally to writing–review and editing. Agnieszka M. Dębska and Katarzyna M. Chyl contributed equally to project administration.

Correspondence concerning this article should be addressed to Agnieszka M. Dębska, Laboratory of Language Neurobiology, Nencki Institute of Experimental Biology, Polish Academy of Sciences, Pasteur 3, 02-093 Warsaw, Poland. Email: a.debska@nencki.edu.pl

grain sizes such as phonemes, and at larger grain sizes such as rhyme and whole words. It has been shown that children who struggle with acquiring strong bidirectional links between orthography and phonology later suffer from reading difficulties (Booth et al., 2000; Caravolas et al., 2012).

According to the psycholinguistic grain size theory (Ziegler & Goswami, 2005), the development and use of various grain sizes across orthography and phonology differ with respect to the transparency of the orthography. In an opaque orthography, like English, the consistency between the orthographic and phonological systems is high at larger grain sizes, such as rhyme or whole word levels, and low at smaller grain sizes, such as phonemes. Thus, reading in opaque orthographies tends to depend on orthography-to-phonology mappings at larger grain sizes (e.g., Lallier et al., 2016, 2013; Lima & Castro, 2010). In contrast, in a transparent orthography like Polish, the consistency between single grapheme and phoneme associations is high. Thus, reading tends to rely on small-grain letter-sound mapping (e.g., Ellis & Hooper, 2001). Although the granularity of mappings might differ across orthographies, learning to read generally progresses from smaller to larger units, which allows readers to establish lexical representations that can make reading faster and less effortful. Whereas, in the auditory domain, phonological awareness develops from larger to smaller grain sizes (Ziegler & Goswami, 2005), in the visual domain, reading acquisition progresses from smaller to larger grain sizes—therefore in the opposite direction (Ziegler & Ferrand, 1998). According to the developmental theory of reading by Ehri (2020), there is a change from partial-alphabetic (knowing spelling–sound unit), to full alphabetic (decoding new words), and finally to consolidation during literacy acquisition. In the final consolidation phase, children mainly operate on larger units to decode multisyllabic words.

Previous behavioral research has shown that orthography influences phonological processing. For example, both adults and children responded faster in a rhyming judgment when the spelling of two words was consistent like "pie–tie" than when it was inconsistent like "tye–tie"—a phenomenon called the "orthographic intrusion effect" (Seidenberg & Tanenhaus, 1979; Zecker, 1991). Landerl et al. (1997) showed orthographic intrusion on phonological awareness tasks in children around 12 years old, demonstrating that, in English, this effect seems to be present relatively early in literacy development. Also, Cone et al. (2008) showed involuntary orthographic influence on auditory rhyming tasks in 9- to 15-year-old children. Behavioral studies show orthographic intrusion in more transparent writing systems such as Polish as well (Awramiuk, 2006). Overall, these behavioral studies consistently suggest that orthographic representations are accessed automatically during phonological processing. Mounting evidence suggests that the orthographic intrusion effect depends on reading level (Morais et al., 1979; Ventura et al., 2007; Ziegler & Muneaux, 2007). For instance, in a lexical decision task in French that measured orthographic and phonological facilitation effects, Ziegler and Muneaux (2007) demonstrated that 7-year-old beginning readers and older children with dyslexia did not show an orthographic intrusion effect, whereas typically reading 11-year-old children did. In summary, there is evidence that the automatic influence of orthography on phonological processing is related to children's reading levels.

At the neural level, studies investigating the connections between orthography and phonology have been mostly focused on a multimodal region in the left posterior superior temporal gyrus (Cantlon et al., 2011; Centanni et al., 2017; Richlan, 2019). However, a growing body of evidence (McNorgan & Booth, 2015; Wang et al., 2018, 2021) has shown that the left ventral occipito-temporal cortex (lvOT) cortex plays a role in the coupling between phonology and orthography. The lvOT is argued to store orthographic representations in skilled readers (e.g., Glezer et al., 2016; Purcell et al., 2011, 2017 but see Planton et al., 2019; Price & Devlin, 2011). With reading acquisition, the lvOT becomes more sensitive to printed words than to other visual stimuli, such as faces, digits, or words from an unknown orthography (e.g., Baker et al, 2007; Cohen & Dehaene, 2004; Dehaene & Cohen, 2011; Dehaene-Lambertz et al., 2018). The acquisition of letter-sound correspondences and print both influence the nature of phonological representations. This dynamic relationship is facilitated by direct white matter connections between the lvOT and key speech processing areas (Saygin et al., 2016). In literate individuals, these connections play a crucial role in the categorical perception of phonemes, phonological processing, and are modulated by audiovisual speech-text stimuli (Conant et al., 2020; Hannagan et al., 2015; McNorgan & Booth, 2015; Romanovska & Bonte, 2021; Romanovska et al., 2021). Furthermore, cross-linguistic studies have demonstrated the presence of a print-speech overlap in the lvOT across various languages in both adults and children, highlighting its relevance to reading-related processes (e.g., Chyl et al., 2021; Rueckl et al., 2015). Dehaene and Cohen (2011) argued that the lvOT is automatically activated in demanding linguistic tasks in a top-down manner due to the activation of the "orthographic code" corresponding to spoken words. This top-down lvOT activation in response to speech may facilitate phonological processing as well as the integration of phonological and orthographic information crucial for reading. We argue that activation of the lvOT during spoken language processing might reflect automatic orthographic coactivation that facilitates phonological processing (Wang et al., 2018, 2021).

Different grain sizes of orthographic representations appear to involve different locations in the lvOT. Vinckier et al. (2007) showed a posterior-to-anterior gradient in selectivity of the lvOT: false fonts, letters, bigrams, trigrams, and real words were processed with increased selectivity in a posterior-to-anterior direction along the lvOT. This is in line with Dehaene et al.'s (2005) proposal where the posterior vOT ($y = -68$) processes single letters, the middle vOT ($y = -56$) processes bigrams and trigrams, and finally, the anterior vOT ($y = -48$) processes lexical units. Support for a gradient also comes from an functional magnetic resonance imaging (fMRI) study on handwriting (Beeson et al., 2003) where writing names of objects (whole words) showed a peak in anterior lvOT (the Montreal Neurological Institute [MNI], −42, −54, −12) whereas a posterior part showed a peak in writing alphabet letters (MNI, −48, −68, −18).

Similar to behavioral studies (Morais et al., 1979; Ventura et al., 2007; Ziegler & Muneaux, 2007) which suggest that the orthographic intrusion effect appears as a function of reading skill, neural studies have shown that the engagement of the lvOT during speech-related tasks is also associated with reading expertise (e.g., Conant et al., 2020; Dębska et al., 2016, 2019; Desroches et al., 2010; Raschle et al., 2012; Wang et al., 2018, 2021). A pattern emerges from these studies. The posterior lvOT appears to be engaged in younger children (e.g., 5- to 6-year-olds in Raschle et al., 2012, peak, MNI −16, −86, −10; 6- to 7-year-olds in Dębska et al., 2016, peak, MNI −20, −58, −6), whereas the anterior lvOT seems engaged in older children (e.g., 9- to 15-year-olds in Desroches et al., 2010, peak, MNI −48 −51 −15; 8- to 17-year-olds in Conant et al., 2020, peak, MNI, −47, −54, −13; 8- to 13-year-olds in Dębska et al., 2019, peak, MNI −42 −50 −10). However, the majority of the previous studies used only one task, making it difficult to interpret the grain size effects at different ages.

An exception to studies using a single grain size in reading is Wang et al. (2018, 2021) who used an in-scanner auditory task including a word onset condition (i.e., alliteration), where word pairs only shared the first sound at the phoneme level, and a rhyme condition, where word pairs shared rhymes. Children were asked to judge whether the word pairs they heard shared a sound or not. Results showed that reading ability in children aged 5–6 years old was positively linked to the activation of the lvOT for the onset versus rhyme contrast. This effect was found only in the posterior part of lvOT, a region believed to be involved in processing smaller units (Wang et al., 2018). In their follow-up study, Wang et al. (2021) showed that in 7- to 8-year-old children, reading fluency measured by rapid automatized naming was positively correlated with activation in the anterior part of the lvOT for the rhyme versus onset contrast corresponding to the larger grain size.

Overall, the neural results of Wang et al. (2018, 2021) suggest a shift in processing from the posterior to anterior lvOT. Younger readers relied on the orthographic coactivations for small grain sizes in the posterior lvOT, whereas older and more skilled readers relied on coactivations for larger grain sizes in the anterior lvOT. This argument is consistent with developmental theories of transition from grapheme-phoneme to whole-word level mappings (Ehri, 2020; Frith, 1986). However, the studies by Wang et al. (2018, 2021) only examined English-speaking children cross-sectionally, and thus were unable to provide strong evidence for developmental changes across languages. In addition, the measure of reading ability in Wang et al. (2021) was rapid naming of letters which may limit the results to small grain size mappings. Using a whole-word reading fluency task should be a better measure to tap into the mapping between phonology and orthography at both lexical and single-letter levels.

We examined developmental changes related to phonological and orthographic coupling at the neural level in Polish children. Similar to Wang et al. (2018, 2021), we measured activity in the lvOT during phonological auditory tasks that operated at different grain sizes, including alliteration, rhyming, and word matching. However, our study was unique in several dimensions. First, we conducted a longitudinal fMRI study to examine developmental changes. Children were tested twice, at Time-point 1 (TIME1) as beginning readers and again after 2 years of schooling at Time-point 2 (TIME2). Second, in contrast to previous studies (Wang et al., 2018, 2021), which defined their lvOT region of interest (ROI) purely based on literature, we applied a visual word localizer to independently identify orthographically relevant voxels in the lvOT in each individual (Pattamadilok et al., 2019). Third, univariate analysis, which averages brain activation across voxels, could fail to detect weak but informative signals distributed in multivariate patterns (Tong & Pratte, 2012). Thus, in addition to the commonly used univariate analyses (e.g., Conant et al., 2020; Dębska et al., 2016, 2019; Desroches et al., 2010; Raschle et al., 2012; Wang et al., 2018, 2021), a multivoxel pattern analysis (MVPA) was used to test the coupling between phonological and orthographic representations.

We investigated how brain activations during phonological tasks in the orthographically relevant part of the lvOT are related to reading skills. Previous studies on beginning and more advanced English readers (Wang et al., 2018, 2021) showed skill-dependent activity of the posterior lvOT during small grain size processing in beginning readers and skill-dependent activity of the anterior lvOT during larger grain size processing in more advanced readers. Polish has a transparent orthography, and therefore the importance of smaller grain size for orthography-to-phonology binding should be accentuated. Thus, we hypothesized that Polish-speaking children would rely longer on small-grain grapheme-to-phoneme mapping than English-speaking children. Specifically, in the case of beginning readers (TIME1), we expected to see—similar to English—a positive relationship between reading level and brain activity in the posterior lvOT during a small grain alliteration task. After 2 years of schooling (at TIME2), we might observe prolonged reliance on the small grain mapping. If bigger units are still processed on a small grain size level we might expect reading-related effects in the anterior lvOT.

To better understand phonological and orthographic coupling, we applied a multivariate approach to examine the similarity of neural patterns. Previous studies using MVPA (e.g., Anzellotti & Caramazza, 2017; Devereux et al., 2013; Kriegeskorte et al., 2008; Murphy et al., 2017) have suggested that higher similarity of activation patterns means that a particular brain region is able to generalize information about stimuli from different modalities. Thus, in the current study, we assumed that the higher the similarity of activation patterns in the lvOT during a visual reading task and a spoken language task, the stronger the coupling between orthographic and phonological representations.

Method

Participants

The first screening phase included 354 children recruited from kindergartens in Warsaw. Children who met the inclusion criteria were invited to participate in the fMRI study ($N = 98$) at the Nencki Institute of Experimental Biology (Warsaw, Poland). Those criteria included: (a) being born at term (>37 weeks), right-handedness, no history of neurological illness or brain damage, no ADHD symptoms, normal hearing, and normal or corrected-to-normal vision, Polish as a first language as reported by their parents in the online questionnaire and (b) reading zero words in a sight-word reading task administered individually in kindergarten. Parents completed an online questionnaire with demographic questions like sex (options: male/female) or first language (Polish/other).

Sixty five children completed a full fMRI session at both TIME1 and TIME2. Among those children, 12 were excluded from analyses due to excessive motion during fMRI scanning at TIME1 or TIME2 (see fMRI Data Analysis). Six were excluded due to not meeting accuracy criteria in the phonological in-scanner task (see the Phonological Auditory Task section). In the end, analyses and results from 47 children[1] are presented in the current study, M_{age} TIME1 = 6.8 (6.05–7.36), M_{age} TIME2 = 8.8 (8.21–9.37), 20 girls, 27 boys.

Prior to fMRI scanning, nonverbal IQ was assessed individually with the Cattell Culture Fair Intelligence Test (Koć-Januchta et al., 2013). All children in the final group scored above 85 in standard scores ($M = 108.67 \pm 12.89$). On the day of the fMRI session at

[1] Among the 47 children, five children did not have behavioral data for the fMRI tasks due to technical problems with response pads during fMRI scanning. Results remained the same when excluding them.

TIME1 and TIME2, participants were administered the Decoding Test with the subtests including letter knowledge, sight word reading, and pseudoword reading (Szczerbiński & Pelc-Pękała, 2013). The averaged time difference between TIME1 and TIME2 fMRI sessions was 680 days (min. 697 days–max. 790 days).

The study was approved by the Research Ethics Committee at the SWPS University of Social Sciences and Humanities in Warsaw and was carried out following the provisions of the World Medical Association Declaration of Helsinki. Parents of children who participated signed an informed consent form and children agreed verbally.

Phonological Auditory Task

The phonological auditory task used a block design and had three runs. Run orders were counterbalanced across participants in TIME1 and TIME2 fMRI sessions. During each trial, two words in a pair were presented sequentially through headphones. Each run included one experimental condition/task. In the first task, children needed to judge whether a stimulus pair rhymed or not (e.g., młot:płot—RHYME task) by pressing corresponding buttons: left for "yes," right for "no." In the second task, children were asked to judge if a stimulus pair started with the same sound or not (e.g., kawka: krata—ALLITERATION task). In the third task, children were required to judge if the two words in a stimulus pair were the same or not (e.g., kino:kino—WORD task). A full list of stimuli is available in the online supplemental materials 1.

Each run included four blocks of five word pairs and five pseudoword pairs resulting in 40 pairs per run (20 word pairs, 20 pseudoword pairs). Each stimulus pair (words/pseudowords) was presented for 1 s, followed by a 2-s response interval. The whole task took approximately 20 min. There were no written words. Instead, a headphone symbol was presented on the screen when children needed to give an answer during the response interval. Although pseudowords were included in our tasks, because the goal of the current study was to examine the developmental changes in Polish word processing and compare the findings with previous studies on word processing in English-speaking children (Wang et al., 2018, 2021), only words were analyzed in the current study. Frequent short Polish words were selected as the stimuli for the current study from the Polish CHILDES database (Haman et al., 2011). All words were regular words, including either one or two syllables. There was no difference across tasks in the number of syllables and phonemes or the frequency of the words.

To make sure that children were engaged and understood the task well, we included only participants with acceptable accuracy and without a response bias. An acceptable accuracy was defined as a minimum of 60% accuracy in the WORD task with no more than 50% of missing responses. The lack of a response bias was defined as the accuracy difference between the "yes" and "no" trials in the WORD task being lower than 50%.

Orthographic Visual Localizer Task

During the orthographic visual task, children were instructed to lay still, carefully watch the images appearing on the screen, and press the left button on the response pad whenever they saw a red dot appear on the screen. The task was composed of four runs (5.13 min each, 20.5 min in total). Stimuli were different in each run. Each of the four runs consisted of 30 blocks including five categories of stimuli, with each category presented with three repetition blocks and three mixed blocks. In each block, six items were presented, with each item having a visual stimulus presentation time of 700 ms followed by a 500 ms interstimulus interval. After each block, a blank screen was presented for 1–3 s. The five different categories of stimuli included (a) words, defined as highly frequent Polish nouns consisting of three to five letters and written in courier font, (b) false fonts, defined as the same Polish words written in letter-like Brussels Artificial Character Sets (BACS) false font (Vidal et al., 2017). BACS is a public false font that was created based on character strokes from existing writing systems and controlled relative to major features of the Latin alphabet (number of strokes, junctions). (c) Faces, defined as neutral children's faces (Meuwissen et al., 2017), (d) houses (Filliter et al., 2016), and (e) objects, which are known from everyday life (Brodeur et al., 2014). The Shine toolbox was used to equalize the luminance and contrast of the visual stimuli (Willenbockel et al., 2010).

In the repetition blocks, the same image was presented six times in a row. In the mixed blocks, six different items from the same category were presented. To make sure participants were engaged in the orthographic visual task, in each run, 10 out of the 30 blocks were designed with a randomly presented red dot serving as a target that required children to press a button. The red dot appeared simultaneously with the stimulus in one of the six trials in a block and could appear in any position in the block except for the first one. To avoid bias in attention, the 10 target trials (in 10 blocks, respectively) were designed to only appear once in each stimulus category (i.e., words, false fonts, faces, houses, and objects) and in each block type (i.e., repetition and mixed). The experiment was programed in Presentation Software (Neurobehavioral Systems, Albany, California).

fMRI Data Acquisition

All children were familiarized with the task and MRI environment in a mock scanner. Different stimuli were used in the mock and real scanning sessions. During the fMRI task, children were instructed to lay still with left and right response pads. The stimuli were provided with MRI-compatible noise-canceling headphones at approximately 70 dB, and visual stimuli were presented on a computer screen.

fMRI data were acquired on a 3T Siemens Trio scanner (Siemens, Berlin, Germany) using a whole-brain echo-planar imaging sequence with a 12-channel head coil in case of phonological task (32 slices, slice-thickness 3.5 mm, repetition time (TR) = 2,000 ms, echo time (TE) = 30 ms, flip angle = 90°, FOV = 192 mm^2, matrix size: 64 × 64, voxel size 3 × 3×4 mm) and 32 channel coil in case of visual task and flip angle = 80 with other parameters equal. Anatomical data were acquired using a 32-channel coil and a T1-weighted sequence (176 slices, slice-thickness 1 mm, TR = 2,530 ms, TE = 3.32 ms, flip angle = 7°, matrix size: 256 × 256, voxel size 1 × 1 × 1 mm).

fMRI Data Analysis

Preprocessing

The neuroimaging data preprocessing and analyses were performed using Statistical Parametric Mapping (SPM12, Welcome Trust Center for Neuroimaging, London, United Kingdom) run on MATLAB R2016b (The Math-Works Inc. Natick, Massachussetts, United States).

In the phonological auditory task, images from TIME1 and TIME2 were realigned to the mean. Next, a pairwise longitudinal registration was performed on T1-weighted images from two time points and a midpoint average image was created. The outcome of the pairwise longitudinal registration was coregistered to the mean functional image. Coregistered images were segmented using pediatric tissue probability maps, while the Template-O-Matic toolbox was used with the matched pairs option. The functional images were normalized using compositions of flow fields and a group-specific template. Finally, the normalized functional images were smoothed with an 8 mm isotropic Gaussian kernel. The data was modeled for each run using the canonical hemodynamic response function convolved with the tasks. ART toolbox was used to reject motion-affected volumes surpassing the movement threshold of 3 mm and a rotation threshold of 0.05 radians (Dębska et al., 2019, 2021; Raschle et al., 2012). Subjects were included if a minimum of 80% of volumes from each run were artifact-free. Images from the orthographic visual localizer data from TIME2 and TIME1 were preprocessed with the standard pipeline described above and modeled.

ROI

In both univariate and multivariate approaches, we tested our hypotheses taking into account two main regions of interest within the left hemisphere: the posterior vOT and anterior vOT. ROIs were prepared based on Wang et al. (2018). First, the anatomical left fusiform gyrus and left inferior temporal gyrus from the automated anatomical labeling template (Rolls et al., 2020) were used to create the left vOT mask. Then, the left vOT mask was divided into two parts based on y coordinates to make the posterior ($-75 < y < 61$) and anterior ($-61 < y < -48$) left vOT next to each other.

For the purpose of validation of the visual orthographic task, the 100 most active voxels were extracted separately within the posterior and anterior lvOT masks (see the Region of Interest section) using the contrast words—false fonts task from the TIME2, mixed blocks fMRI session. We decided to use only the TIME2 session to localize orthographic voxels because TIME1 data failed to reveal any significant activation clusters in the vOT at the group level for words-false fonts, in line with previous observations in young, beginning readers (Chyl et al., 2018, 2023). To make sure that we were successful in extracting voxels relevant to orthographic processing, we performed a correlation analysis with word reading at TIME1 and TIME2 and the 100 most active voxels in the posterior and anterior vOT in the contrast of interest (words-false fonts). Posterior and anterior vOT beta values from a visual task positively correlated with the reading level at TIME1 (posterior lvOT, $r = .3$, $p < .05$, anterior lvOT, $r = .3$, $p < .01$) and at TIME2 (posterior lvOT, $r = .2$, $p < .05$, anterior lvOT, $r = .2$, $p < .05$). All p-values for correlations were adjusted using the Holm–Bonferroni correction for multiple comparisons.

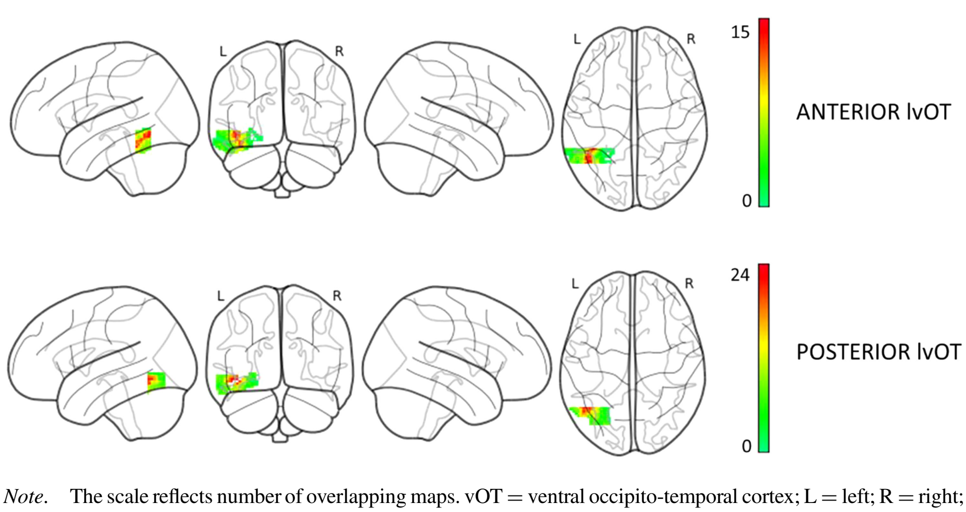
Univariate Analysis

We focus on experimental trials contrasted with explicitly modeled REST periods for TIME1 and TIME2. Therefore, first-level statistical analysis was calculated for TIME1 and TIME2; ALLITERATION-rest, RHYME-rest, WORD-rest. At the group level, one-sample t-tests for TIME1 and TIME2 for each task were performed as well as paired t-tests between TIME1 and TIME2 to track changes in activation with time and schooling. The results are family-wise error (FWE)-corrected ($p < .05$).

For univariate as well as multivariate analysis, we used the 100 most active voxels (highest t-values from spmT maps) from the visual localizer words-false fonts contrast (TIME2) in the anterior and posterior vOT (see the Orthographic Visual Localizer Task section). Figure 1 presents the number and localization of overlapping individual masks within the posterior and anterior vOT. Then, beta values associated with each task minus rest were extracted using MATLAB script. R software (R Core Team, 2012) and lme4 were used to perform a linear mixed effects analysis of the relationship

Figure 1

Localization of the 100 Most Active Voxels Overlapping Between Participants (Dark Gray) Within the Posterior and Anterior vOT (Light Gray) Based on t-Maps From the Orthographic Visual Localizer Task (Word > False Font)



Note. The scale reflects number of overlapping maps. vOT = ventral occipito-temporal cortex; L = left; R = right; lvOT = left ventral occipito-temporal cortex. See the online article for the color version of this figure.

Table 1

Final Group Results (N = 47): Mean (Standard Deviation); Range and Statistical Comparison (Paired t-Tests) for Tasks of Interest

Time points comparison	Time1	Time2	Test
Age (years)	6.8 (0.3); 6.1–7.4	8.8 (0.3); 8.2–9.3	$t = 241.8$, $p < .001$
Letter knowledge[a]	42.0 (1.5);11–64	63.4 (1.1); 60–64	$t = 9.6$, $p < .001$
Words per minute[b]	7 (8); 0–35	52 (20); 17–96	$t = 17.2$, $p < .001$
Pseudoword per minute[b]	7 (7); 0–32	34 (9); 15–55	$t = 21.6$, $p < .001$
Rhyme ACC raw (max 20)	16 (3); 9–20	18 (2); 9–20	$t = 2.5$, $p < .01$
Alliteration ACC raw (max 20)	11 (3); 3–18	17 (2.5); 9–20	$t = 10.1$, $p < .001$
Word ACC raw (max 20)	18 (1); 14–20	19 (1); 6–20	$t = 4.8$, $p < .001$

Note. ACC = accuracy.
[a] Letter naming, raw scores, max. 64 (lower and upper case). [b] Number of words/pseudowords read correctly in 60 s.

between the average beta value for each subject (ACTIVATION), ROIs (posterior, anterior vOT), time points (1, 2), TASK (rhyme, alliteration, word), and SKILL (reading level). As random effects, we added intercepts for subjects.

Multivariate Pattern Analysis

In the multivariate analysis, general linear models like those used for the univariate approach were made on the first level of analysis, but nonsmoothed data was used for the orthographic visual localizer and phonological auditory task (see Wang et al., 2021). We defined the orthographically relevant voxels in the same way as for the univariate approach (see above). We extracted the *t*-values from spmT maps for TIME2 for orthographic processing (from visual localizer task) for contrast words-false fonts (100 voxels) and phonological processing (from phonological auditory task) for contrasts rhyme-rest, alliteration-rest, word-rest (TIME1 and TIME2). Then we calculated the correlations between the two spmT maps within ROIs masks and, as a next step, conducted a fisher-*z* transformation on the correlational values. Higher correlation values suggest a higher representational similarity between orthographic and phonological information on the neural level.

We performed a linear mixed effects analysis of the relationship between SIMILARITY, ROIs (posterior, anterior vOT), time points (1, 2) reading level (SKILL) and TASK using R software (R Core Team, 2012) and lme4. As random effects, we added intercepts for subjects.

Transparency and Openness

Behavioral and ROI data as well as the data analysis codes used in the current study are available online on the Open Science Framework data repository https://osf.io/uydv6/?view_only=e96f9

Table 2

One-Sample t-Tests for Alliteration, Rhyme, and Word Matching Tasks at TIME1 and TIME2, FWE for Voxels, p < .05

Clusters	Peak_*x*	Peak_*y*	Peak_*z*	Cluster_mean	Number of voxels	AAL definition (>5%)
Rhyme-rest TP1						
1	−52	−18	4	8.08	60,512	25.95% Temporal_Sup_L; 13.63% no_label; 12.00% Insula_L; 10.83% Temporal_Mid_L; 9.72% Frontal_Inf_Tri_L; 5.91% Frontal_Inf_Oper_L;
2	58	−20	8	7.50	40,584	37.89% Temporal_Sup_R; 12.56% Insula_R; 11.43% no_label; 6.33% Putamen_R; 5.78% Temporal_Pole_Sup_R;
3	−4	14	48	7.96	18,040	37.12% Supp_Motor_Area_L; 21.02% Supp_Motor_Area_R; 15.96% Cingulate_Mid_R; 11.53% Cingulate_Mid_L; 7.58% Frontal_Sup_Medial_L;
4	−46	0	50	6.17	960	97.50% Precentral_L;
5	30	−94	8	5.74	600	97.33% Occipital_Mid_R;
6	−10	−14	12	5.83	464	100.00% Thalamus_L
Alliteration-rest TP1						
1	−60	−32	8	7.36	89,920	17.62% Temporal_Sup_L; 16.50% no_label; 11.67% Frontal_Inf_Tri_L; 9.45% Temporal_Mid_L; 7.42% Insula_L; 7.21% Frontal_Inf_Oper_L; 6.94% Precentral_L;
2	54	−20	2	7.43	57,440	32.24% Temporal_Sup_R; 10.93% no_label; 8.52% Frontal_Mid_2_R; 7.81% Temporal_Mid_R;
3	2	6	58	7.71	31,304	29.41% Supp_Motor_Area_L; 21.72% Supp_Motor_Area_R; 14.21% Cingulate_Mid_R; 9.00% Frontal_Sup_Medial_L; 8.59% Cingulate_Mid_L;
4	26	−88	−10	6.02	14,672	28.14% Cerebelum_6_R; 24.05% Occipital_Inf_R; 10.47% Occipital_Mid_R; 8.62% Fusiform_R; 8.29% Lingual_R;
5	−20	−98	−4	6.06	5,048	38.83% Occipital_Inf_L; 35.82% Occipital_Mid_L; 7.13% Lingual_L;
6	12	−14	16	5.87	4,144	49.23% Caudate_R; 32.63% Thalamus_R;
7	34	−52	40	5.82	2,920	48.77% Parietal_Inf_R; 33.70% Angular_R;
8	−44	−52	−26	5.77	2,568	46.11% Cerebelum_6_L; 42.06% Cerebelum_Crus1_L; 6.85% Fusiform_L;
9	−6	−20	−10	5.75	2,240	95.00% no_label; 5.00% Thalamus_L

(*table continues*)

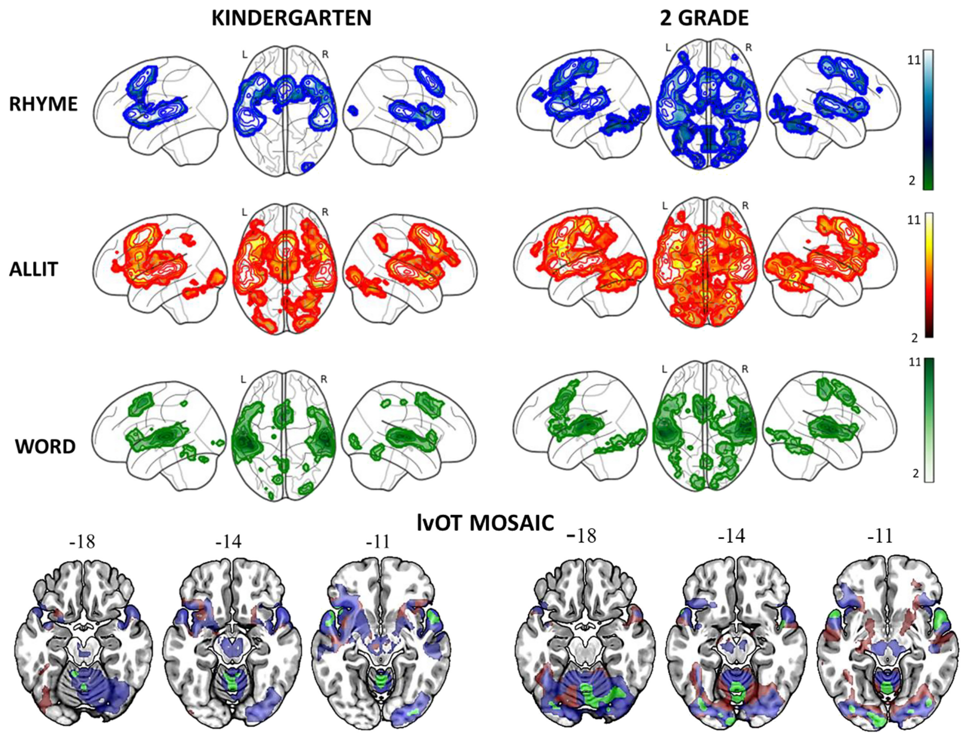
Table 2 (*continued*)

Clusters	Peak_*x*	Peak_*y*	Peak_*z*	Cluster_mean	Number of voxels	AAL definition (>5%)
10	−44	−50	52	5.49	1,056	97.73% Parietal_Inf_L;
11	−42	48	8	5.45	552	57.97% Frontal_Mid_2_L; 42.03% Frontal_Inf_Tri_L
12	64	10	16	5.43	304	57.89% Frontal_Inf_Oper_R; 31.58% Precentral_R; 10.53% Rolandic_Oper_R
Word-rest TP1						
1	−58	−26	6	7.15	32,552	41.12% Temporal_Sup_L; 22.39% Temporal_Mid_L;
2	58	−16	4	6.94	24,544	54.43% Temporal_Sup_R; 14.93% Temporal_Mid_R;
3	−4	6	56	6.40	11,688	41.82% Supp_Motor_Area_L; 28.41% Supp_Motor_Area_R; 10.54% Cingulate_Mid_R; 7.32% Frontal_Sup_Medial_L;
4	0	−56	−10	5.66	2,160	48.15% Vermis_4_5; 34.81% Vermis_6; 11.48% Cerebelum_4_5_L;
5	36	−90	−4	5.64	1,672	53.11% Occipital_Inf_R; 42.11% Occipital_Mid_R;
6	−6	−76	−26	5.55	560	61.43% Cerebelum_Crus1_L; 30.00% Cerebelum_Crus2_L; 8.57% Vermis_7
7	48	−48	52	5.28	328	100.00% Parietal_Inf_R
8	38	−20	54	5.40	240	100.00% Precentral_R
Rhyme-rest TP2						
1	−62	−28	12	7.53	59,808	24.02% Temporal_Sup_L; 13.10% no_label; 10.62% Frontal_Inf_Tri_L; 10.41% Insula_L; 9.72% Temporal_Mid_L; 8.75% Frontal_Inf_Oper_L; 7.49% Rolandic_Oper_L;
2	−4	−2	62	7.32	40,888	22.23% Supp_Motor_Area_L; 19.02% Supp_Motor_Area_R; 16.98% Precentral_R; 9.43% Cingulate_Mid_R; 7.22% Cingulate_Mid_L; 7.16% no_label; 6.93% Postcentral_R;
3	62	−20	6	6.98	38,968	39.91% Temporal_Sup_R; 12.40% Insula_R; 9.83% Rolandic_Oper_R;
4	36	−64	−24	6.00	15,616	30.74% Cerebelum_6_R; 19.26% Cerebelum_Crus1_R; 11.17% Occipital_Mid_R; 8.97% Occipital_Inf_R; 8.56% Fusiform_R; 7.94% Lingual_R;
5	−32	−62	−24	5.91	11,552	32.48% Cerebelum_6_L; 23.89% Cerebelum_Crus1_L; 11.22% Fusiform_L; 6.79% Occipital_Inf_L; 6.72% Lingual_L; 6.16% Calcarine_L; L
6	−44	−22	60	6.56	10,560	61.97% Precentral_L; 35.15% Postcentral_L;
7	8	−76	−24	6.02	6,896	23.43% Vermis_6; 17.05% Vermis_4_5; 15.31% Vermis_7; 9.63% Cerebelum_4_5_L; 8.70% Cerebelum_Crus1_R; 7.77%
8	−12	−16	10	6.16	1,760	100.00% Thalamus_L
9	−34	48	20	5.84	1,440	95.00% Frontal_Mid_2_L; 5.00% Frontal_Sup_2_L
10	14	−26	−6	6.13	656	87.80% no_label; 9.76% Lingual_R; 2.44% Hippocampus_R
11	12	−14	10	5.58	640	100.00% Thalamus_R
12	−12	10	0	5.46	584	42.47% no_label; 30.14% Caudate_L; 20.55% Putamen_L; 6.85% Pallidum_L
Alliteration-rest TP2						
1	58	−12	0	5.57	706,712	21.41% no_label; 3.01% Precentral_L; 2.97% Temporal_Sup_R; 2.56% Frontal_Inf_Tri_L; 2.49% Postcentral_L; 2.41% Temporal_Sup_L; 2.15% Precentral_R; 2.11% Calcarine_L; 1.87% Supp_Motor_Area_R;
2	0	−30	−42	3.62	200	100.00% no_label
3	−24	46	−8	3.39	96	58.33% no_label; 41.67% Frontal_Sup_2_L
4	48	48	−6	3.39	88	54.55% Frontal_Inf_Orb_2_R; 45.45% Frontal_Mid_2_R
5	6	−72	46	3.33	56	100.00% Precuneus_R
6	−30	14	52	3.34	40	100.00% Frontal_Mid_2_L
7	−14	−62	46	3.31	16	100.00% Parietal_Sup_L
8	−28	−32	−30	3.55	8	100.00% Cerebelum_4_5_L
9	−28	−34	−32	3.38	8	100.00% Cerebelum_4_5_L
Word-rest TP2						
1	−62	−24	8	6.86	31,408	35.33% Temporal_Sup_L; 14.29% Temporal_Mid_L; 8.46% no_label; 7.41% Rolandic_Oper_L; 6.93% Frontal_Inf_Oper_L; 6.32% Frontal_Inf_Tri_L; 5.68% Insula_L; 5.02% Heschl_L;
2	56	−10	2	7.24	27,896	51.16% Temporal_Sup_R; 9.41% Rolandic_Oper_R; 8.03% Temporal_Mid_R; 7.80% Insula_R; 7.00% no_label; 6.77% Temporal_Pole_Sup_R; 6.14% Heschl_R;
3	−4	8	46	6.51	11,664	34.50% Supp_Motor_Area_L; 23.11% Supp_Motor_Area_R; 19.34% Cingulate_Mid_L; 11.52% Cingulate_Mid_R;
4	32	−24	68	6.35	6,144	63.54% Precentral_R; 23.96% Postcentral_R; 12.50% no_label
5	36	−46	−26	5.85	4,440	66.31% Cerebelum_6_R; 14.41% Fusiform_R; 10.99% Cerebelum_4_5_R; 5.41% Cerebelum_Crus1_R
6	42	−78	−6	5.86	2,464	52.27% Occipital_Inf_R; 25.97% Lingual_R; 13.96% Temporal_Inf_R; 3
7	−4	−94	−14	5.74	2,280	66.67% Calcarine_L; 17.19% Lingual_L; 6.67% Occipital_Mid_L
8	6	−66	−14	5.57	1,672	61.24% Vermis_6; 19.62% Vermis_4_5; 12.92% Cerebelum_6_R; 5.74% Cerebelum_6_L;
9	14	−14	6	5.80	1,136	81.69% Thalamus_R; 8.45% Pallidum_R; 5.63% Putamen_R;
10	−36	−42	−28	5.63	920	53.91% Cerebelum_Crus1_L; 40.00% Cerebelum_6_L

Note. L = left; R = right; AAL = automated anatomical labeling; FWE = family-wise error; Inf = inferior; Sup = superior; Mid = middle; TP1 = time point1; TP2 = time point 2.

Figure 2

One-Sample t-Tests (Task vs. Explicitly Modeled Rest) for the Whole Group of Participants at Both Time Points



Note. FWE = family-wise error; ALLIT = alliteration; lvOT = left ventral occipito-temporal cortex, $p < .05$. Blue = rhyme; red = alliteration; green = word. See the online article for the color version of this figure.

95bcf094a24aaf5709ca9ae6b5d and https://github.com/wangjinvandy/PhonoOrthoCoupling. Unthresholded fMRI *t*-maps are available on NeuroVault (https://identifiers.org/neurovault.collection:12960). We report all data exclusions, all manipulations, and all measures and statistical tools in the study in the Method section. This study's design and its analysis were not preregistered.

Results

Behavioral Results

All means, standard deviations, and comparisons between TIME1 and TIME2 for behavioral tests of interest and in-scanner performance are shown in Table 1. The series of paired *t*-tests show an increase in performance in all tasks with time and schooling.

We looked more closely at the differences between reading words and pseudowords at TIME1 and TIME2. Repeated measure ANOVA (RM ANOVA) revealed a significant effect of task (word, pseudoword), $F(1, 46) = 391$, $p < .001$, $\eta^2 = .6$, time point, $F(1, 46) = 68$, $p = .001$, $\eta^2 = .88$, and interaction effect, $F(1, 46) = 83$, $p = .001$, $\eta^2 = .64$. Post hoc tests with the Bonferroni correction showed that children were faster in reading words than pseudowords at TIME2 ($p < .001$) but not at TIME1 ($p = .24$).

Next, we looked at the accuracy level between TIME1 and TIME2 in an in-scanner phonological task. RM ANOVA with time points (1, 2) and TASK (alliterations, rhymes, word) as factors revealed the main effects of time points, $F(1, 41) = 69$, $p < .001$, $\eta^2 = .65$, and TASK, $F(2, 82) = 82$, $p < .001$, $\eta^2 = .66$, and a significant interaction effect, $F(2, 82) = 39$, $p < .001$, $\eta^2 = .48$. Post hoc tests with the Bonferroni correction revealed that the highest accuracy was achieved in the case of the WORD task, lower accuracy in the case of the RHYME task, and the lowest accuracy in the case of ALLITERATIONS (all $p < .001$). Children had higher accuracy at TIME2 than TIME1 ($p < .001$). As for the interaction effect, the WORD task showed significantly higher accuracy than RHYME and ALLITERATION only at TIME1 ($p < .001$). Tasks did not differ in accuracy at TIME2. For all means and test results, see Table 1.

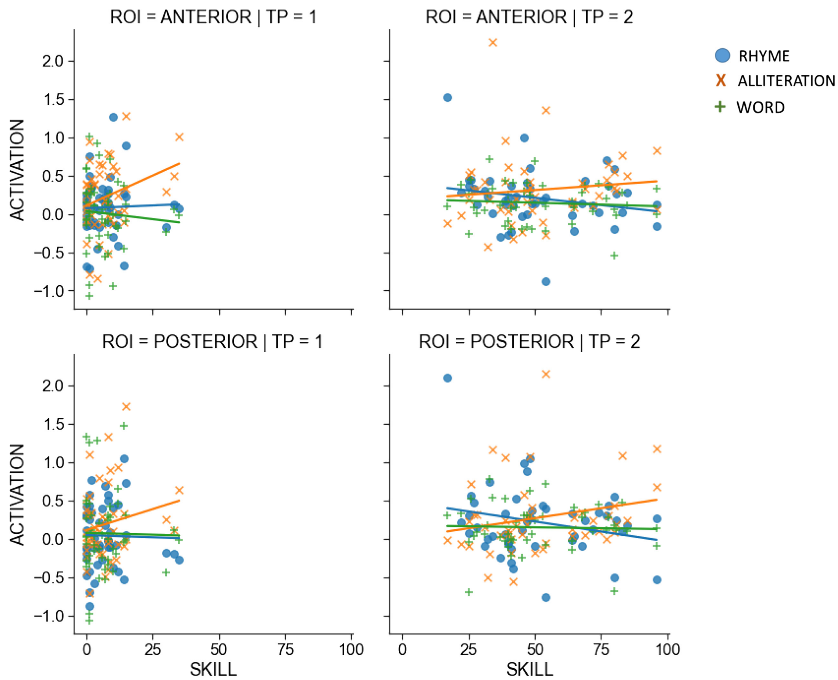
fMRI Results

Whole Brain

Results from one-sample *t*-tests for TIME1 and TIME2 for three tasks minus REST are presented in Table 2 and Figure 2. Paired *t*-tests between TIME1 and TIME2 for each task separately did not yield significant results at a chosen threshold (FWE, $p < .05$). However, on a more liberal threshold (FWE-cluster corrected, $p = .001$, $p < .05$), we found that children in a TIME2 > TIME1 comparison activated the bilateral lingual gyrus and right precentral/postcentral in the case of RHYME (lingual: −12 −66 −8, $T = 5.3$, 6,705 voxels, PreC/PostC: 20 −32 56, $T = 4.8$, 352 voxels) and the left lingual gyrus in the case of ALLITERATION (−22 −42 0, $T = 5.2$, 791 voxels).

Figure 3

Relation of Reading Skill to Activation Patterns for Each Phonological Task for the Two Time Points and for the Anterior and Posterior Parts of the lvOT



Note. Two interaction effects were found: SKILL:TASK (6.7, $p < .001$, ALLITERATION > WORD, RHYME) and TIME POINTS:TASK (5.7, $p < .01$, TIME1 ALLITERATION > WORD, RHYME). lvOT = left ventral occipito-temporal cortex; ROI = region of interest; TP = time point. See the online article for the color version of this figure.

Univariate ROI

A linear mixed effects analysis with subject as a random effect revealed significant main effects of time points (TIME2 > TIME1) and TASK (ALLITERATION > RHYME, WORD) on ACTIVATION, time points, $F(1, 501)= 4.5$, $p < .03$, correlation coefficient = 3.75, TASK, $F(2, 505) = 6.5$, $p < .001$, correlation coefficient = 5.39. Significant interaction effects were found in the case of TIME POINTS:TASK, $F(2, 505) = 5.7$, $p < .005$, correlation coefficient = 4.6, and SKILL:TASK, $F(2, 505) = 6.7$, $p < .001$, correlation coefficient = 9.0. Variance of random effects was 0.029 (564 observations, 47 groups). Post hoc pairwise comparisons showed that at TIME1 but not TIME2, the ALLITERATION task evoked more activation than RHYME and WORD independent of ROIs (ALLITERATION-WORD, $t(507) = 4.36$, $p < .001$, ALLITERATION-RHYME, $t(507) = 3.838$ $p < .001$). Furthermore, ALLITERATION was more related to the reading level than two other tasks independent of time points and ROIs. ALLITERATION-RHYME, $t(507) = 2.9$, $p < .01$, ALLITERATION-WORD, $t(507) = 4.02$, $p < .001$. All degrees of freedom were estimated using the Kenward–Roger method (Halekoh & Højsgaard, 2014). See Figures 3 and 4. To show the relation between task performance and VOT activations, we ran an additional model where accuracy in all tasks and time points was added as a variable of interest. The influence of accuracy on activation was not significant, $F = 0.0072$, $p = .9322$. Overall, adding a new variable to the model did not change the pattern of results.

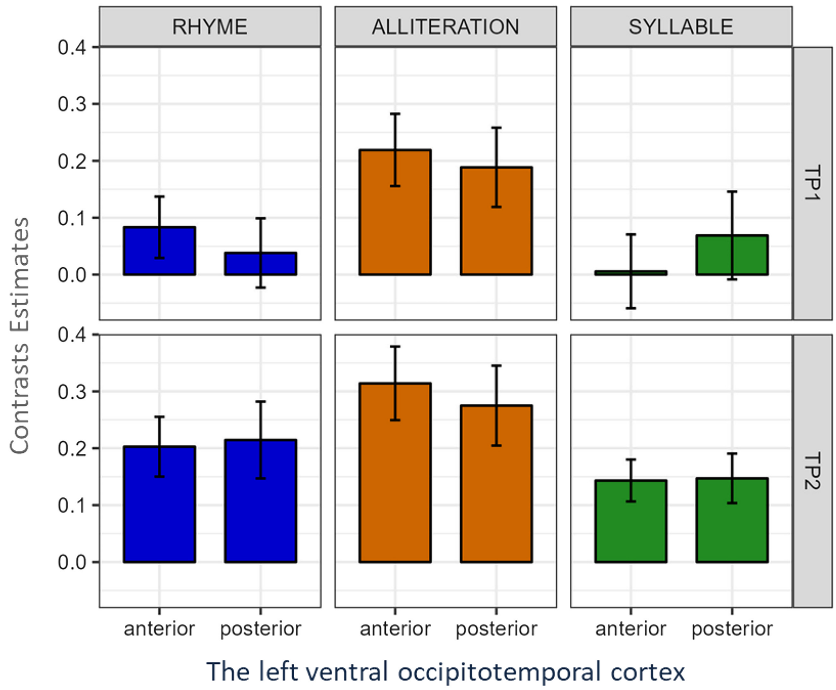
Multivariate Pattern Analysis

A linear mixed effects analysis with subject as a random effect revealed significant negative effect of reading level (SKILL) on SIMILARITY, $F(1, 284) = 5.9$, $p < .01$, correlation coefficient = −6.7, and significant interaction effects for SKILL:TIME POINTS, $F(1, 399) = 6.8$, $p < .01$, correlation coefficient = 3.6, and ROI:TIME POINTS, $F(1, 488) = 5.3$, $p < .01$, correlation coefficient = −6.2. The variance of random effects was 0.05 (552 observations, 46 groups). Post hoc pairwise comparisons showed that the main effect of SKILL on similarity level was negative overall. A significant interaction effect of SKILL with TIME POINTS showed that similarity level was less negatively related to the reading level in TIME2 than in TIME1, $t(177) = 2.3$, $p < .05$. Furthermore, at TIME2 but not TIME1 the anterior lvOT shows more similarity level between processing auditory and visual stimuli than posterior lvOT, $t(500) = 3.9$, $p < .001$. All degrees of freedom were estimated using the Kenward–Roger method. See Figure 5.

Discussion

Our longitudinal fMRI study in 6- to 9-year-old children suggested that beginning readers of a transparent orthography rely on activation of small grain grapheme-to-phoneme mapping in the lvOT cortex, with greater reliance on better readers. Our MVPA analysis found that the similarity between auditory and visual word processing in the lvOT was higher for the anterior compared

Figure 4
Mean Beta Change in Region of Interests, Tasks, and Time Points



Note. TP1 = time point 1; TP2 = time point 2. See the online article for the color version of this figure.

to posterior lvOT in TIME2, consistent with the argument of a transition from posterior to anterior lvOT involvement with reading acquisition (Wang et al., 2021). Also, contrary to our expectations, the similarity between auditory and visual stimuli was overall negatively related to reading skills but significantly less negatively after 2 years of formal education than in beginning readers (Figure 6).

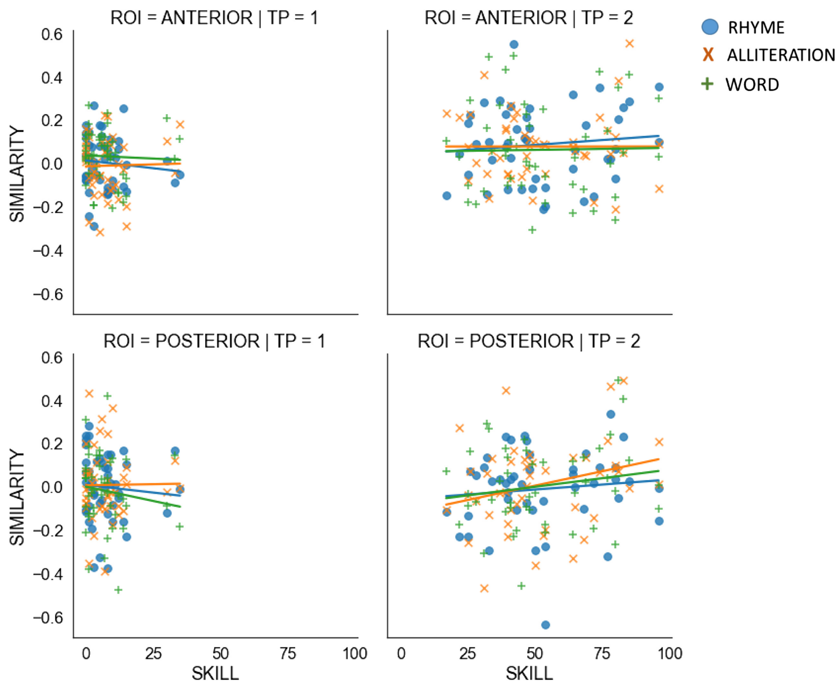
Our results showed that the activity of the lvOT during the alliteration task, but not the rhyme or word task, was positively related to reading skill independent of time of schooling. This finding suggests a general role of small-grain orthographic-phonological coupling in reading development in Polish. This is different from previous findings in English-speaking children reported by Wang et al. (2018, 2021) in which they found a developmental transition from relying on small grain sizes to large grain sizes. Specifically, Wang et al. (2018) showed a correlation between reading skill and brain activation for phonemic processing in the posterior vOT in 5- to 6-year-old children, whereas their later study (2021) using the same design with an older 7- to 8-year-old cohort showed skill-related brain activation for rhyme processing in the anterior vOT. Polish was expected to be different from English because it is a transparent orthography with regular mappings between letters and phonemes. For reading development in Polish, it is most beneficial to establish phonology-orthography coupling of small grain sizes. Therefore, even when Polish-speaking children grow older and become more efficient in reading, their reading skill continues to rely on fine-grain letter-sound coupling. When the task demands processing of larger grain sizes, as in the case of rhyme and word tasks, we do not see a relation between lvOT activity and reading level.

Although the transparency of a language may drive greater reliance on certain grain sizes, reading seems to universally follow a progression from small to large units according to the reading stage theory by Ehri (2020). Thus, we expected to see a progression of skill-related involvement from the posterior to the anterior vOT as reading develops. However, we did not observe skill-dependent activity in the anterior lvOT in older children as was found for English (Wang et al., 2021). One explanation is that, in our sample at the second time point, reading was not sufficiently automatic, so older children will need to be studied to determine whether this relation appears. Indeed, in a study on older Polish children (8–13, $M =$ 10.5; Dębska et al., 2019), we found a positive correlation between reading skill and activity of the anterior lvOT during a large grain size auditory phonological task (words matching). The posterior to anterior transition in a transparent orthography thus might occur later than in English.

Another possibility for the lack of transition may be the insensitivity of using univariate analysis. We did find that children responded faster in word than pseudoword reading at the second compared to the first time point, suggesting that children start to rely on automatized whole word processing as suggested by Ehri (2004). Using MVPA, we also found greater similarity of processing auditory and visual word stimuli in the anterior lvOT at the second, but not first, time point. This was not observed in the posterior lvOT. At the first time point, when children were mostly at the grapheme-phoneme stage of processing, there was no difference in similarity patterns for whole-word stimuli between the posterior and anterior lvOT. However, at the second time point when children processed larger units, we see higher similarity patterns in the anterior than in the posterior lvOT. This suggests that, even though we did not confirm the shift from posterior to anterior lvOT processing in our activation-based analysis, the MVPA analysis shows that the

Figure 5

Relation of Reading Skill to Multivariate Similarity Patterns for Each Phonological Task for the Two Time Points and for the Anterior and Posterior Part of the lvOT



Note. Two interaction effects were found: TIME POINTS:ROI (5.3, $p < .05$, ANTERIOR > POSTERIOR in TIME2 but not TIME1) and TIME POINTS: SKILL (6.7, $p < .001$, TIME2 > TIME1). lvOT = left ventral occipito-temporal cortex; ROI = region of interest; TP = time point. See the online article for the color version of this figure.

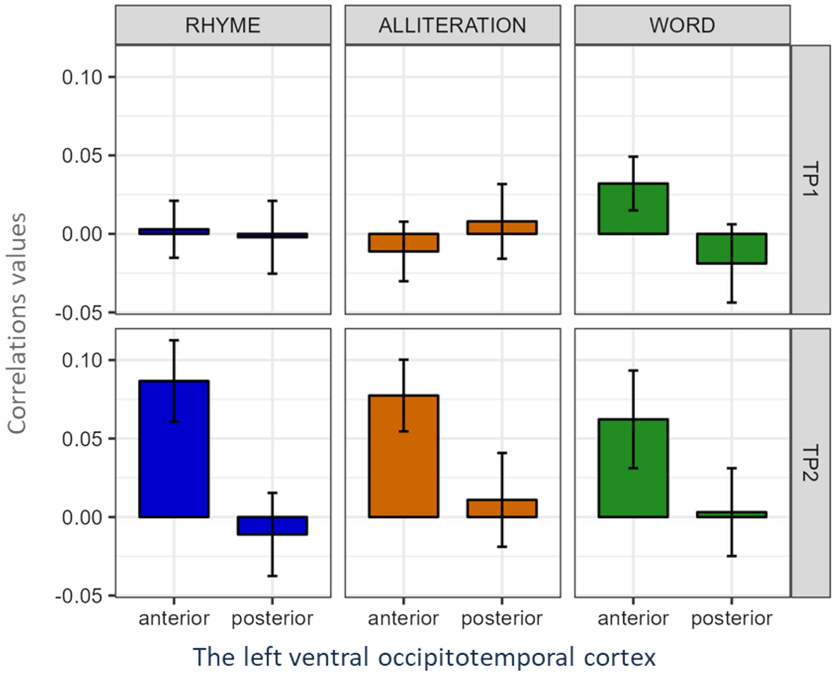
orthography-to-phonology coupling develops in the anterior lvOT after 2 years of formal education. The latest connectivity studies showed that anterior lvOT is connected with higher-language areas as opposed to the posterior part which is connected with visual features of words. This suggests a functional division between the posterior and anterior parts. The pattern observed in the current study—no clear posterior to anterior shift of activations but the growing similarity patterns for multimodal stimuli in the anterior compared to the posterior part in TIME2 may be a sign of developmental changes in the connectivity of the lvOT with higher language areas supporting the emergent functional division. More studies on the role of the posterior lvOT in early processing of single letters and sounds are therefore needed, as well as longitudinal studies that would track the connectivity pattern over time.

Finally, by examining activation patterns, we see that the similarity between visual and auditory words is negatively related to the reading skill, but more negatively related at TIME1 than TIME2. This seems inconsistent with McNorgan and Booth (2015) studied older 8- to 13-year-old children and showed that audiovisual integration is positively related to reading-related skills in the left vOT. However, other studies in adults (Pattamadilok et al., 2019; Zhao et al., 2017) claim that the lvOT represents both orthographic and phonological information unimodally. Using our method, we cannot be certain if the similarity of brain activation patterns is due to increased multimodality or increased sensitivity to the same information from both modalities. An increase in similarity with respect to reading level over time might be explained by the idea that unimodal activations need to specialize first to allow later multimodal enhancement. However, further longitudinal studies are needed to directly distinguish the effects of multimodality and unimodality on reading acquisition. Also, our study tests developmental processes on real words (nouns) which have a strong semantic aspect as opposed to grammatical words or pseudowords. Because it was shown that lvOT processes are modulated by semantic factors (Wang et al., 2018) a different developmental pathway might be expected in the case of grammatical words or pseudowords (McNorgan & Booth, 2015; Zhang & Peng, 2022).

Conclusion

By investigating the coupling of orthographic and phonological representations in the lvOT, the current study contributes to the orthographic depth hypothesis proposed by Ziegler and Goswami (2005). Our study shows that reliance on small grain sizes is crucial for reading in a transparent orthography. Mappings may rely on small grain sizes longer than on larger grain sizes in transparent orthographies. This is different from the development observed in opaque orthographies (i.e., English), where there is an earlier shift from small to large grain size processing. However, the growing orthography-to-phonology coupling in the anterior lvOT seems to be present in opaque English and relatively transparent Polish

Figure 6
Changes in Similarity Patterns in Region of Interests, Tasks, and Time Points



Note. TP1 = time point 1; TP2 = time point 2. See the online article for the color version of this figure.

orthographies. Thus, our results also show universal developmental effects.

Constraints on Generality

The final sample included in our study was recruited from kindergartens in Warsaw, Poland, which limited the generalization of findings to children from relatively large, industrial areas. Also, the results were only included in the analysis when children completed all experimental procedures, including sessions in the fMRI scanner, which is a very demanding environment for 6-year-olds. Therefore, the context of the phonological tasks and cognitive as well as psychological demands of the procedure were different from school or home-based situations. Lastly, our results are only generalized to children learning to read in a relatively transparent orthography, comparable to Polish. We have no reason to believe that the results depend on other characteristics of the participants, materials, or context.

References

Anzellotti, S., & Caramazza, A. (2017). Multimodal representations of person identity individuated with fMRI. *Cortex*, *89*, 85–97. https://doi.org/10.1016/j.cortex.2017.01.013

Awramiuk, E. (2006). *Lingwistyczne podstawy początkowej nauki czytania i pisania po polsku*. Trans Humana Wydawnictwo Uniwersyteckie.

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092. https://doi.org/10.1073/pnas.0703300104

Beeson, P., Rapcsak, S., Plante, E., Chargualaf, J., Chung, A., Johnson, S., & Trouard, T. (2003). The neural substrates of writing: A functional magnetic resonance imaging study. *Aphasiology*, *17*(6–7), 647–665. https://doi.org/10.1080/02687030344000067

Booth, J. R., Perfetti, C. A., MacWhinney, B., & Hunt, S. B. (2000). The association of rapid temporal perception with orthographic and phonological processing in children and adults with reading impairment. *Scientific Studies of Reading*, *4*(2), 101–132. https://doi.org/10.1207/S1532799XSSR0402_02

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of standardized stimuli (BOSS) phase II: 930 new normative photos. *PLoS ONE*, *9*(9), Article e106953. https://doi.org/10.1371/journal.pone.0106953

Cantlon, J. F., Pinel, P., Dehaene, S., & Pelphrey, K. A. (2011). Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebral Cortex*, *21*(1), 191–199. https://doi.org/10.1093/cercor/bhq078

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G., & Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, *23*(6), 678–686. https://doi.org/10.1177/0956797611434536

Centanni, T. M., King, L. W., Eddy, M. D., Whitfield-Gabrieli, S., & Gabrieli, J. D. E. (2017). Development of sensitivity versus specificity for print in the visual word form area. *Brain and Language*, *170*, 62–70. https://doi.org/10.1016/j.bandl.2017.03.009

Chyl, K., Gentile, F., Dębska, A., Dynak, A., Łuniewska, M., Wójcik, M., Bonte, M., & Jednoróg, K. (2023). Early reading skills and the ventral occipito-temporal cortex organization. *Cortex*, *160*, 134–151. https://doi.org/10.1016/j.cortex.2023.01.004

Chyl, K., Kossowski, B., Dębska, A., Łuniewska, M., Banaszkiewicz, A., Żelechowska, A., Frost, S. J., Mencl, W. E., Wypych, M., Marchewka, A., Pugh, K. R., & Jednoróg, K. (2018). Prereader to beginning reader:

Changes induced by reading acquisition in print and speech brain networks. *Journal of Child Psychology and Psychiatry*, *59*(1), 76–87. https://doi.org/10.1111/jcpp.12774

Chyl, K., Kossowski, B., Wang, S., Dębska, A., Łuniewska, M., Marchewka, A., Wypych, M., Bunt, M. V. D., Mencl, W., Pugh, K., & Jednoróg, K. (2021). The brain signature of emerging reading in two contrasting languages. *NeuroImage*, *225*, Article 117503. https://doi.org/10.1016/j.neuroimage.2020.117503

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *Neuroimage*, *22*(1), 466–476. https://doi.org/10.1016/j.neuroimage.2003.12.049

Conant, L. L., Liebenthal, E., Desai, A., Seidenberg, M. S., & Binder, J. R. (2020). Differential activation of the visual word form area during auditory phoneme perception in youth with dyslexia. *Neuropsychologia*, *146*, Article 107543. https://doi.org/10.1016/j.neuropsychologia.2020.107543

Cone, N. E., Burman, D. D., Bitan, T., Bolger, D. J., & Booth, J. R. (2008). Developmental changes in brain regions involved in phonological and orthographic processing during spoken language processing. *Neuroimage*, *41*(2), 623–635. https://doi.org/10.1016/j.neuroimage.2008.02.055

Dębska, A., Banfi, C., Chyl, K., Dzięgiel-Fivet, G., Kacprzak, A., Łuniewska, M., Plewko, J., Grabowska, A., Landerl, K., & Jednoróg, K. (2021). Neural patterns of word processing differ in children with dyslexia and isolated spelling deficit. *Brain Structure and Function*, *226*(5), 1467–1478. https://doi.org/10.1007/s00429-021-02255-2

Dębska, A., Chyl, K., Dzięgiel, G., Kacprzak, A., Łuniewska, M., Plewko, J., Marchewka, A., Grabowska, A., & Jednoróg, K. (2019). Reading and spelling skills are differentially related to phonological processing: Behavioral and fMRI study. *Developmental Cognitive Neuroscience*, *39*, Article 100683. https://doi.org/10.1016/j.dcn.2019.100683

Dębska, A., Łuniewska, M., Chyl, K., Banaszkiewicz, A., Żelechowska, A., Wypych, M., Marchewka, A., Pugh, K. R., & Jednoróg, K. (2016). Neural basis of phonological awareness in beginning readers with familial risk of dyslexia—Results from shallow orthography. *NeuroImage*, *132*, 406–416. https://doi.org/10.1016/j.neuroimage.2016.02.063

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, *9*(7), 335–341. https://doi.org/10.1016/j.tics.2005.05.004

Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLoS Biology*, *16*(3), Article e2004103. https://doi.org/10.1371/journal.pbio.2004103

Desroches, A. S., Cone, N. E., Bolger, D. J., Bitan, T., Burman, D. D., & Booth, J. R. (2010). Children with reading difficulties show differences in brain regions associated with orthographic processing during spoken language processing. *Brain Research*, *1356*, 73–84. https://doi.org/10.1016/j.brainres.2010.07.097

Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *The Journal of Neuroscience*, *33*(48), 18906–18916. https://doi.org/10.1523/JNEUROSCI.3809-13.2013

Ehri, L. C. (2004). Teaching phonemic awareness and phonics: An explanation of the national reading panel meta-analyses. In P. Mccardle, V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 153–186). Paul H. Brookes Publishing.

Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly*, *55*(1), 45–73. https://doi.org/10.1002/rrq.256

Ellis, N. C., & Hooper, A. M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics*, *22*(4), 571–599. https://doi.org/10.1017/S0142716401004052

Filliter, J. H., Glover, J. M., McMullen, P. A., Salmon, J. P., & Johnson, S. A. (2016). The DalHouses: 100 new photographs of houses with ratings of typicality, familiarity, and degree of similarity to faces. *Behavior Research Methods*, *48*(1), 178–183. https://doi.org/10.3758/s13428-015-0561-8

Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, *36*(1), 67–81. https://doi.org/10.1007/BF02648022

Glezer, L. S., Eden, G., Jiang, X., Luetje, M., Napoliello, E., Kim, J., & Riesenhuber, M. (2016). Uncovering phonological and orthographic selectivity across the reading network using fMRI-RA. *Neuroimage*, *138*, 248–256. https://doi.org/10.1016/j.neuroimage.2016.05.072

Halekoh, U., & Højsgaard, S. (2014). A Kenward–Roger approximation and parametric bootstrap methods for tests in linear mixed models—The R Package pbkrtest. *Journal of Statistical Software*, *59*(9), 1–32. https://doi.org/10.18637/jss.v059.i09

Haman, E., Etenkowski, B., Łuniewska, M., Szwabe, J., Dąbrowska, E., Szreder, M., & Łaziński, M. (2011). *Polish CDS corpus*. CHILDES/TalkBank.

Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., & Dehaene, S. (2015). Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences*, *19*(7), 374–382. https://doi.org/10.1016/j.tics.2015.05.006

Koć-Januchta, M., Cattell, R. B., Weiss, R. H., & Osterland, J. (2013). *CFT 1-R neutralny kulturowo test inteligencji Cattella-wersja 1: zrewidowana przez Rudolfa H. Weißa i Jürgena Osterlanda: podręcznik* [CFT 1-R culturally neutral intelligence test Cattell version 1: revised by Rudolf H. Weiß and Jürgen Osterland: manual]. Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego [Psychological Testing Laboratory wof the Polish Psychological Society].

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *4*, Article 2008. https://doi.org/10.3389/neuro.06.004.2008

Lallier, M., Acha, J., & Carreiras, M. (2016). Cross-linguistic interactions influence reading development in bilinguals: A comparison between early balanced French-Basque and Spanish-Basque bilingual children. *Developmental Science*, *19*(1), 76–89. https://doi.org/10.1111/desc.12290

Lallier, M., Carreiras, M., Tainturier, M. J., Savill, N., & Thierry, G. (2013). Orthographic transparency modulates the grain size of orthographic processing: Behavioral and ERP evidence from bilingualism. *Brain Research*, *1505*, 47–60. https://doi.org/10.1016/j.brainres.2013.02.018

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition*, *63*(3), 315–334. https://doi.org/10.1016/S0010-0277(97)00005-X

Lima, C. F., & Castro, S. L. (2010). Reading strategies in orthographies of intermediate depth are flexible: Modulation of length effects in Portuguese. *European Journal of Cognitive Psychology*, *22*(2), 190–215. https://doi.org/10.1080/09541440902750145

McNorgan, C., & Booth, J. R. (2015). Skill dependent audiovisual integration in the fusiform induces repetition suppression. *Brain and Language*, *141*, 110–123. https://doi.org/10.1016/j.bandl.2014.12.002

Meuwissen, A. S., Anderson, J. E., & Zelazo, P. D. (2017). The creation and validation of the developmental emotional faces stimulus set. *Behavior Research Methods*, *49*(3), 960–966. https://doi.org/10.3758/s13428-016-0756-7

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*(4), 323–331. https://doi.org/10.1016/0010-0277(79)90020-9

Murphy, C., Rueschemeyer, S. A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *NeuroImage*, *147*, 19–31. https://doi.org/10.1016/j.neuroimage.2016.11.067

Pattamadilok, C., Planton, S., & Bonnard, M. (2019). Spoken language coding neurons in the visual word form area: Evidence from a TMS adaptation paradigm. *NeuroImage*, *186*, 278–285. https://doi.org/10.1016/j.neuroimage.2018.11.014

Planton, S., Chanoine, V., Sein, J., Anton, J. L., Nazarian, B., Pallier, C., & Pattamadilok, C. (2019). Top-down activation of the visuo-orthographic system during spoken sentence processing. *NeuroImage*, *202*, Article 116135. https://doi.org/10.1016/j.neuroimage.2019.116135

Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. https://doi.org/10.1016/j.tics.2011.04.001

Purcell, J. J., Jiang, X., & Eden, G. F. (2017). Shared orthographic neuronal representations for spelling and reading. *Neuroimage*, *147*, 554–567. https://doi.org/10.1016/j.neuroimage.2016.12.054

Purcell, J. J., Turkeltaub, P. E., Eden, G. F., & Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Frontiers in Psychology*, *2*, Article 239. https://doi.org/10.3389/fpsyg.2011.00239

Raschle, N. M., Zuk, J., & Gaab, N. (2012). Functional characteristics of developmental dyslexia in left-hemispheric posterior brain regions predate reading onset. *Proceedings of the National Academy of Sciences*, *109*(6), 2156–2161. https://doi.org/10.1073/pnas.1107721109

R Core Team. (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Richlan, F. (2019). The functional neuroanatomy of letter-speech sound integration and its relation to brain abnormalities in developmental dyslexia. *Frontiers in Human Neuroscience*, *13*, Article 21. https://doi.org/10.3389/fnhum.2019.00021

Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *Neuroimage*, *206*, Article 116189. https://doi.org/10.1016/j.neuroimage.2019.116189

Romanovska, L., & Bonte, M. (2021). How learning to read changes the listening brain. *Frontiers in Psychology*, *12*, Article 726882. https://doi.org/10.3389/fpsyg.2021.726882

Romanovska, L., Janssen, R., & Bonte, M. (2021). Cortical responses to letters and ambiguous speech vary with reading skills in dyslexic and typically reading children. *NeuroImage: Clinical*, *30*, Article 102588. https://doi.org/10.1016/j.nicl.2021.102588

Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W.-J., Bick, A., Frost, S. J., Hancock, R., Wu, D. H., Mencl, W. E., Duñabeitia, J. A., Lee, J. R., Oliver, M., Zevin, J. D., Hoeft, F., Carreiras, M., Tzeng, O. J., Pugh, K. R., & Frost, R. (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proceedings of the National Academy of Sciences*, *112*(50), 15510–15515. https://doi.org/10.1073/pnas.1509321112

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D., & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*(9), 1250–1255. https://doi.org/10.1038/nn.4354

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546–554. https://doi.org/10.1037/0278-7393.5.6.546

Szczerbiński, M., & Pelc-Pękała, O. (2013). *Zestaw metod do diagnozy trudności w czytaniu–test dekodowania* [The decoding test-A set of tools for diagnosing reading difficulties]. PTPiP, Gdansk.

Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, *63*, 483–509. https://doi.org/10.1146/annurev-psych-120710-100412

Ventura, P., Morais, J., & Kolinsky, R. (2007). The development of the orthographic consistency effect in speech recognition: From sublexical to lexical involvement. *Cognition*, *105*(3), 547–576. https://doi.org/10.1016/j.cognition.2006.12.005

Vidal, C., Content, A., & Chetail, F. (2017). BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior Research Methods*, *49*(6), 2093–2112. https://doi.org/10.3758/s13428-016-0844-8

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Wang, J., Joanisse, M. F., & Booth, J. R. (2018). Reading skill related to left ventral occipitotemporal cortex during a phonological awareness task in 5–6-year old children. *Developmental Cognitive Neuroscience*, *30*, 116–122. https://doi.org/10.1016/j.dcn.2018.01.011

Wang, J., Joanisse, M. F., & Booth, J. R. (2021). Letter fluency in 7-8-year-old children is related to the anterior, but not posterior, ventral occipitotemporal cortex during an auditory phonological task. *Developmental Cognitive Neuroscience*, *47*, Article 100898. https://doi.org/10.1016/j.dcn.2020.100898

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Zecker, S. G. (1991). The orthographic code: Developmental trends in reading-disabled and normally-achieving children. *Annals of Dyslexia*, *41*(1), 178–192. https://doi.org/10.1007/BF02648085

Zhang, Z., & Peng, P. (2022). Reading real words versus pseudowords: A meta-analysis of research in developmental dyslexia. *Developmental Psychology*, *58*(6), 1035–1050. https://doi.org/10.1037/dev0001340

Zhao, L., Chen, C., Shao, L., Wang, Y., Xiao, X., Chen, C., Yang, J., Zevin, J., & Xue, G. (2017). Orthographic and phonological representations in the fusiform cortex. *Cerebral Cortex*, *27*(11), 5197–5210. https://doi.org//10.1093/cercor/bhw300

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*(4), 683–689. https://doi.org/10.3758/BF03208845

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., & Muneaux, M. (2007). Orthographic facilitation and phonological inhibition in spoken word recognition: A developmental study. *Psychonomic Bulletin & Review*, *14*(1), 75–80. https://doi.org/10.3758/bf03194031

Received November 16, 2022
Revision received August 15, 2023
Accepted August 26, 2023 ■