Emerging neural specialization of the ventral occipitotemporal cortex to characters through phonological association learning in preschool children

Journal Article

Author(s):
Pleisch, Georgette; Karipidis, Iliana I.; Brauchli, Christian; Röthlisberger, Martina; Hofstetter, Christoph; Stämpfli, Philipp; Walitza, Susanne; Brem, Silvia

Publication date:
2019-04-01

Permanent link:
https://doi.org/10.3929/ethz-b-000325569

Rights / license:
Creative Commons Attribution 4.0 International

Originally published in:
NeuroImage 189, https://doi.org/10.1016/j.neuroimage.2019.01.046
Emerging neural specialization of the ventral occipitotemporal cortex to characters through phonological association learning in preschool children

Georgette Pleisch\textsuperscript{a,b}, Iliana I. Karipidis\textsuperscript{a,b}, Christian Brauchli\textsuperscript{a}, Martina Röthlisberger\textsuperscript{a}, Christoph Hofstetter\textsuperscript{a}, Philipp Stämpfli\textsuperscript{c}, Susanne Walitza\textsuperscript{a,b}, Silvia Brem\textsuperscript{a,b,c,*}

\textsuperscript{a} Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric Hospital, University of Zurich, Switzerland
\textsuperscript{b} Neuroscience Center Zurich, University of Zurich and ETH Zurich, Zurich, Switzerland
\textsuperscript{c} MR-Center of the Department of Psychiatry, Psychotherapy and Psychosomatics and the Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric Hospital, University of Zurich, Zurich, Switzerland

\textbf{A R T I C L E  I N F O}

\textbf{Keywords:}
Prereading children
Training
Reading acquisition
Simultaneous EEG/fMRI
ERP-N1
vOT
Expertise
Familiarity

\textbf{A B S T R A C T}

The ventral occipitotemporal (vOT) cortex serves as a core region for visual processing, and specific areas of this region show preferential activation for various visual categories such as faces and print. The emergence of such functional specialization in the human cortex represents a pivotal developmental process, which provides a basis for targeted and efficient information processing. For example, functional specialization to print in the left vOT is an important prerequisite for fluent reading. However, it remains unclear, which processes initiate the preferential cortical activations to characters arising in the vOT during child development. Using a multimodal neuroimaging approach with preschool children at familial risk for developmental dyslexia, we demonstrate how varying levels of expertise modulate the neural response to single characters, which represent the building blocks of print units. The level of expertise to characters was manipulated firstly through brief training of false-font speech–sound associations and secondly by comparing characters for which children differed in their level of familiarity and expertise accumulated through abundant exposure in their everyday environment. Neural correlates of character processing were tracked with simultaneous high-density electroencephalography and functional magnetic resonance imaging in a target detection task. We found training performance and expertise-dependent modulation of the visual event-related potential around 220 ms (N1) and the corresponding vOT activation. Additionally, trained false-font characters revealed stronger functional connectivity between the left fusiform gyrus (FFG) seed and left superior parietal/lateral occipital cortex regions with higher training performance. In sum, our results demonstrate that learning artificial-character speech–sound associations enhances activation to trained characters in the vOT and that the magnitude of this activation and the functional connectivity of the left FFG to the parieto-occipital cortex depends on learning performance. This pattern of results suggests emerging development of the reading network after brief training that parallels network specialization during reading acquisition.

1. Introduction

The development of functional specialization in specific cortical areas is critical for information processing in various domains (Houdé et al., 2010). The left ventral occipitotemporal (vOT) cortex is important for processing a variety of visual categories (Dehaene and Cohen, 2007; Hasson et al., 2003; Kourtzi and Kanwisher, 2001; Yovel and Kanwisher, 2004), and its crucial role in orthographic processing is widely accepted and has been demonstrated in numerous studies (Baker et al., 2007a; Binder et al., 2006; Bruno et al., 2008; Cohen et al., 2002; Dehaene and Cohen, 2011; Dehaene et al., 2002, 2005; Glezer et al., 2009; Martin et al., 2015). The left vOT shows increasing visual specialization to print readers along a posterior-to-anterior axis (Kronschnabel et al., 2013; Vinckier et al., 2007), and a specific area in the left mid fusiform gyrus, called the visual word form area (VWFA; Cohen et al., 2000, 2002), shows preferential activation to print (Baker et al., 2007b; Cohen et al., 2003; Vinckier et al., 2007) and whole word forms (Glezer et al., 2009; Glezer and Riesenhuber, 2013). As an electrophysiological correlate of

\textsuperscript{*} Corresponding author. Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric Hospital, Neumuensterallee 9, CH-8032, Zurich, Switzerland.

E-mail address: sbrem@kjpd.uzh.ch (S. Brem).

https://doi.org/10.1016/j.neuroimage.2019.01.046

Received 22 July 2018; Received in revised form 21 December 2018; Accepted 21 January 2019

Available online 22 January 2019

1053-8119/© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
this vOT activation, the characteristic occipitotemporal negativity in the event-related potential (ERP) after around 170 ms (N1 also referred to as N170) shows similar specialization in adults (Bentin et al., 1999; Brem et al., 2006; Maurer et al., 2005), reflected in more pronounced amplitudes with print, especially over the left hemisphere. This specialization of the vOT/N1 to print is regarded as a key component for efficient and fluent reading, which develops with reading acquisition and is often reduced in impaired readers (Ben-Shachar et al., 2011; Maurer et al., 2007; Paulesu et al., 2001; Shaywitz et al., 2002).

### 1.1. Functional specialization of the ventral occipitotemporal cortex (vOT)

Neuroimaging research has provided important insights into the functional specialization of the left vOT cortex to print during reading acquisition in children (Brem et al., 2012; James, 2010; Maurer et al., 2011; Saygin et al., 2016), adults (Dehaene et al., 2010; Pegado et al., 2014), and after symbol training in primates (Srihasam et al., 2012). Learning drives the expertise reflected in the preferential neural responses of the left vOT (Ben-Shachar et al., 2011; Boros et al., 2016; Oulilade et al., 2015; Price and Devlin, 2011; Xue et al., 2006), the corresponding N1 (Brem et al., 2009; Maurer et al., 2006, 2011), and the visually evoked fields of the magnetoencephalogram (Caffarra et al., 2017; Parvainen et al., 2006) after around 150–250 ms. The process of learning print has been simulated in adults using unknown and artificial scripts (Brem et al., 2017; Hashimoto and Sakai, 2004; Maurer et al., 2010; McCandliss et al., 1997; Song et al., 2010) and pseudoword learning paradigms (Glezer et al., 2015). Functional magnetic resonance imaging (fMRI) studies have revealed stronger activation in the bilateral vOT for trained letters and characters than magnetic resonance imaging (fMRI) studies have revealed stronger activation in the bilateral vOT for trained letters and characters than

### 1.2. Deficient functional specialization associated with poor reading

Activation in the left vOT is modulated by word reading skills in children (Ben-Shachar et al., 2011; Maurer et al., 2011) and adults (Dehaene et al., 2010; Pegado et al., 2014; Xue et al., 2006). Reading-impaired children (Araújo et al., 2012; Hasko et al., 2013; Maurer et al., 2007) and adults (Helenius et al., 1999; Mahé et al., 2012) typically show reduced preferential activation of this region to print in the vOT fMRI response (Brunswick et al., 1999; Hoefl et al., 2007a; Maisog et al., 2008; Paulesu et al., 2001; Richlan et al., 2009, 2011; Shaywitz et al., 2002; van der Mark et al., 2009) and the N1 ERP (but see Fraga Gonzalez et al., 2016; Fraga Gonzalez et al., 2014) and decreased functional connectivity to higher-order language areas (Finn et al., 2014; van der Mark et al., 2011; Zaré et al., 2017) and microstructural alterations (Linködersö et al., 2012; Williams et al., 2018). Interestingly, functional and structural alterations have been detected in pre-schoolers, such as diminished activity in response to orthographic stimuli in the occipitotemporal cortex (Specht et al., 2009) and reduced grey-matter volume in children at familial risk for dyslexia (Raschle et al., 2011b). Previous results thus suggest that the preferential response to print is likely to be initiated rapidly through suitable training and that the extent of specialization to print depends on reading performance, expertise, and risk factors for reading problems in children. What remains largely unclear so far is whether the vOT specialization is a purely perceptual process that can be explained by increased visual familiarity with a specific category, whether vOT specialization is mostly shaped by higher-order processing areas, or whether both visual familiarity and increased expertise with linguistic information contribute to the vOT specialization. Specialization of the vOT as a result of increasing visual familiarity would involve increased perceptual tuning to letters, strings, and whole words with exposure to print regardless of whether associations with phonology or lexical information have been established. Alternatively, such specialization may involve pre-existing connections with anterior language areas (Hamagan et al., 2015; Saygin et al., 2016) and phonologically guided tuning (Brem et al., 2010; Sandak et al., 2004; Schlaggar and McCandliss, 2007) or prediction errors resulting from matching and integrating feedforward sensory inputs and information from other areas (Price and Devlin, 2011; Stevens et al., 2017). The latter concept
would require some expertise to connect the visual information to higher-level linguistic or magnitude information. Here, we aim to clarify to what extent letter–speech sound learning has an impact on activation of bilateral vOT and how prereading familiarity with letters and numbers drives initial functional specialization and the consequent preferential cortical activations observed in the left vOT during child development.

1.3. Preferential processing of single characters

Relatively few studies have examined the processing of the building blocks of printed words, i.e. single letters. One important step for later reading outcomes is early automatization to letter processing through accumulating familiarity with letters in the everyday environment prior to formal literacy acquisition. Indeed, letter naming is often impaired in children with developmental dyslexia and is one of the most promising predictors of preschool reading outcome (Gern-rud-Clikeman et al., 2006; Wolf et al., 1986). Moreover, selective neural processing of letters starts early in the development of visual processing (Herdman, 2011; Miller and Wood, 1995; Rey et al., 2009) and seems to recruit a distinct, more anterior, lateral area of the vOT (Flowers et al., 2004; Gros et al., 2001; James et al., 2005; Joseph et al., 2006) than letter strings, for which preferential activation was reported in a “letter form area” (LFA), posterior to the classic VWFA (James et al., 2005; Tagamets et al., 2000; Thesen et al., 2012). However, modulation of the classic VWFA has also been reported when categorizing single letters (Pernet et al., 2005; Polk et al., 2002). Similar to words, the visual N1 response to familiar letters showed stronger amplitudes than to unfamiliar ones (Wong et al., 2005), but disfluent letter typefaces also produced stronger N1 than fluent ones, suggesting greater attentional demands for processing difficult-to-read scripts and pseudoletters (Herdman and Takai, 2013; Keage et al., 2014). However, changes in single-letter processing with reading acquisition and development have been less studied. Cantlon et al. (2011) reported coarse alphanumerical tuning when comparing letters and numbers with objects (faces, shoes), but no fine tuning between numbers and letters in four-year-old prereading children. Furthermore, sensorimotor letter training induced a stronger activation in bilateral fusiform gyri of prereading children with letters than with pseudo-letters and shapes (James, 2010). A recent study by Centanni et al. (2018) suggests that the greater specificity in the fusiform cortex activation to letters than to false-fonts is associated with reading ability and a reduction of the left fusiform face area in kindergarteners. Similar to studies on word reading (Maison et al., 2008; Martin et al., 2015; Richlan et al., 2009) one study showed that the activation to single letters in the bilateral fusiform gyr of reading-impaired school children was lower than in typical readers (Blau et al., 2010).

When assessing the natural course of reading development by tracking the processing of written words by children in longitudinal or cross-sectional studies (Brem et al., 2009; Dehaene et al., 2010; Shaywitz et al., 2007; Turkeltaub et al., 2003), learning-related changes are likely to reflect mixed effects of age, general matura-
tion, and emerging phonological, lexical, and semantic associations that develop partly in concert. Consequently, such studies cannot fully explain the extent to which visual familiarity or emerging feedforward and feedback circuits to and from phonological and semantic association areas drive functional brain specialization to print. We consider visual familiarity as the capacity to perceptually discriminate between specific and visually similar character categories (letters, numbers, false-fonts), facilitated by prolonged exposure to stimuli. Children usually gain a high level of familiarity during the course of development due to exposure and experience in their everyday environment. In addition, we consider various types of expertise that go beyond visual categorization, as a result of formal instruction, practice and exposure and involve manipulation and association with other modalities. Expertise allows the coupling of the visual units (letters, numbers and trained false-fonts) with phonological, lexical, semantic and/or magnitude information and thus describes the effect of reciprocal interactions between the visual system and multiple top-down processes (Harel, 2016). Here, our first aim was to clarify the impact of learning new phonological associations on visual processing of characters. Through false-font speech–sound association learning, we simulated the first step of formal reading acquisition in prereaders usually taught after school enrol-
ment. After the training, we compared the brain activation to the trained false-font characters to passively viewed, visually matched false-font characters with high-density simultaneous EEG–fMRI record-
ings. We expect emerging expertise only in the trained characters in contrast with the passively viewed characters (untrained false-font), but comparable levels of moderate visual familiarity for both trained and control false-font. Studying prereaders, who do not yet possess a functional reading network, overcomes undesirable interference with an existing reading system, which is an obstacle encountered in studies on literate adults (Hashimoto and Sakai, 2004; Maurer et al., 2010). Using false-font characters instead of real letters allows the effects of phonology on character processing to be clarified more directly due to the absence of semantic and lexical associations.

The second aim of our study was to provide more detailed insights into how character types such as letters and numbers may be encoded in the brains of prereading children, who have accumulated a varying amount of visual familiarity and expertise. Given the age of these prereading children, we expect high visual familiarity with real letters and numbers, a high level of expertise with numbers, but only rudimen-
tary expertise with letters’ phonological associations and no expertise and moderate visual familiarity with passively viewed false-font characters. These diverging levels of expertise based on associations with phonological and semantic/magnitude information are assumed to shape the functional activation of the vOT and N1 ERP in preschool children.

2. Materials and methods

2.1. Participants

A group of 31 native German-speaking, prereading kindergarten children (aged 6.7 ± 3.16 males) completed four parts of an audiovisual target detection task in a simultaneous EEG/fMRI session. A group of 18 children at varying risk for developmental dyslexia (8f, mean: 6.7 ± 0.36y; Table 1) met data quality criteria for both EEG and fMRI analyses and was included in the main analysis. Additional analyses were performed with extended EEG (N = 23) and fMRI groups (N = 24, InTable 1.1, who met data quality criteria in either the EEG or fMRI modality.

Children’s risk for developmental dyslexia varied and was estimated from their parents’ reading history assessed with the Adult Reading History Questionnaire (ARHQ; Lefly and Pennington, 2000). Parental values above 0.3 in the ARHQ indicate a familial risk for the child (Lefly and Pennington, 2000). Each child’s individual familial risk score for developmental dyslexia was defined as the higher parental ARHQ value (0.53 ± 0.2) and was included in all statistical analyses as a covariate to control for familial risk.

Subjects’ nonverbal intelligence scores (IQ > 85, 108 ± 13.3) were within or above the normal range, as estimated with the block design test of the Wechsler Intelligence Scale for Children (HAWIK-IV; Petermann and Petermann, 2007). All children had normal or corrected to normal visual acuity and no diagnosis of attention-deficit/hyperactivity disorder or other neurological or cognitive impairments. The parents gave written informed consent and the children gave oral assent. The local ethics committee of the Canton of Zurich and neighboring Cantons in Switzerland approved the study. All participants received vouchers and presents as compensation.
indicated by values–awareness with a behavioral test battery. In addition, we also assessed children’s phonological knowledge of letters in experiment 3.1 words, see Table 1.)

The subjects’ behavioral characterization and learning achievements were assessed in two separate sessions preceding the simultaneous EEG/fMRI measurement (Table 1). Reading status was tested with a short list of twenty simple, one- or two-syllable (2–5 letters) words frequently found in common first-grade textbooks and written in upper-case letters. In Switzerland, formal reading instruction starts with school enrolment at age 6–7, after the end of kindergarten. As usual at this age, most children were able to name and/or partly write the (upper-case) letters of their given names, and very few children were able to decode a few short words written in upper-case letters (mean 2.9 ± 3.1 words, see Table 1) despite not having received formal reading instruction. In contrast to normally three times.

![Table 1](image) Subject characteristics and performance on behavioral measures.

- **Variables**
  - Sex (male/female): 10/8
  - Handedness (right/left): 16/2
  - False-font training set (1/2): 10/8

- **Mean ± SD Range**
  - Age (years): 6.7 ± 3.6 6.1–7.2
  - IQ estimation (nonverbal): 108 ± 13.3 85–125
  - Familial risk (ARHQ) for dyslexia: 0.53 ± 0.2 0.26–0.8
  - No. of children at increased familial risk (ARHQ >0.3): 14
  - No. of children with affected siblings/developmental language delay: 2/2

- **Reading (one- or two-syllable words in upper-case letters, max. −20)**:
  - 2.9 ± 3.1 0.10

- **Letter speech-sound knowledge (upper and lower case, max. −52)**:
  - 15.4 ± 3.6 2–35

- **Letter speech-sound knowledge of letters in experiment (b,d,t,m,z, max. −6)**:
  - 1.2 ± 1.0 0–3

- **Number knowledge (one- to three-digit numbers, max. −21)**:
  - 15.3 ± 5.0 7–21

- **Number knowledge of numbers in experiment (1,2,3,4,5,6 max. −6)**:
  - 6 ± 0

- **Rapid naming RAN (objects/second)**:
  - 27 ± 17 1–50

- **No. of children exceeding RAN risk index**: 6

- **Phonological awareness**: 47 ± 19 2–90

- **No. of children exceeding phonological awareness risk index**: 3

- **Passive vocabulary**: 49.8 ± 26.2 1–90

- **Vocabulary – word meaning**: 45.2 ± 28.9 10–93

- **Non-word repetition (Mottier Test)**:
  - 34.2 ± 22.1 1–85

- **Grapheme-Phoneme Correspondence Training (GPC)**
  - **Mean ± SD Range**
    - GPC training duration (in min): 19 ± 3.9 12–26
    - GPC training accuracy (weighted, in %): 81 ± 9 62–91

### 2.2. Behavioral assessment

The subjects’ behavioral characterization and learning achievements were assessed in two separate sessions preceding the simultaneous EEG/fMRI measurement (Table 1). Reading status was tested with a short list of twenty simple, one- or two-syllable (2–5 letters) words frequently found in common first-grade textbooks and written in upper-case letters. In Switzerland, formal reading instruction starts with school enrolment at age 6–7, after the end of kindergarten. As usual at this age, most children were able to name and/or partly write the (upper-case) letters of their given names, and very few children were able to decode a few short words written in upper-case letters (mean 2.9 ± 3.1 words, see Table 1) despite not having received formal reading instruction. In contrast to normally three times.

Training duration and achievement was calculated for the complete training session. Training duration varied between children: faster learners needed less time to successfully learn the associations in the artificial GPC training. To account for the varying number of distractors, accuracy was calculated using a weighting factor defined as the number of presented items proportional to the maximum possible number of presented items. Correlations of training parameters (duration, weighted accuracy) with behavioral assessments have already been described in Karipidis et al. (2017) but were recomputed for the specific subgroup of children analyzed in this article (see Inline Supplementary Table S2). Because the groups and results largely overlap, the results of these correlations are not further discussed, and the reader is referred to the corresponding articles (Karipidis et al., 2017, 2018).

### 2.3. Artificial grapheme-phoneme correspondence (GPC) training

Between 1 and 5 days (mean: 2.3d) prior to the simultaneous EEG/fMRI session, all participants trained the correspondences between six false-font characters and familiar sound words with an adapted version of the GraphoGame phonics training program (Karipidis et al., 2017; Lyytinen et al., 2007, 2009). To simulate reading acquisition, children were assigned to train the association between natural sound words and one of two false-font (FF) character sets using an adaptive randomization approach (Fig. 1a, training set 1 n = 10, training set 2, n = 8). During each trial of the training, the FF characters of the other set (untrained control FF) were presented on the upper part of the screen while the trained FF characters were presented in the middle of the screen. This design allowed us to visually familiarize children with both sets even though only one set was actively trained. Both sets of FF characters were based on six lower-case letters of the Latin alphabet. We rearranged parts of each letter (b, d, m, t, u, z) also appearing in the letter (LET) condition in Swiss School font with the Font Creator (Version V4.5) to form new “pseudoletters”. These false-font characters were comparable in size and width to the LET characters. In addition, we included two mirrored FF characters in each FF set to match the difficulty of mirror-image letters (b-d), which also occur in the Latin alphabet. In this way, we aimed to keep the visual content of FF and LET constant. During the training, the visual stimuli were presented on a laptop positioned in front of the child. The sounds, spoken by a female voice, were presented over headphones. The children had to choose the correct FF grapheme corresponding to the heard phoneme. The numbers of visual distractors (1–3) changed according to the accuracy rate of the previous trial. The training consisted of a single session with 131 trials divided into ten training levels. Struggling children completed supporting levels to train specific correspondences with a high error rate (84 ± 39.7 trials). The training lasted until each child was able to match the six speech sounds to their corresponding characters or until each supporting level was repeated maximally three times.

Training duration and achievement was calculated for the complete training session. Training duration varied between children: faster learners needed less time to successfully learn the associations in the artificial GPC training. To account for the varying number of distractors, accuracy was calculated using a weighting factor defined as the number of presented items proportional to the maximum possible number of presented items. Correlations of training parameters (duration, weighted accuracy) with behavioral assessments have already been described in Karipidis et al. (2017) but were recomputed for the specific subgroup of children analyzed in this article (see Inline Supplementary Table S2). Because the groups and results largely overlap, the results of these correlations are not further discussed, and the reader is referred to the corresponding articles (Karipidis et al., 2017, 2018).

### 2.4. Audiovisual target detection task

Participants performed an implicit audiovisual target detection task (Fig. 1), which was divided into four parts of 375s each to maintain the attention of the young children (Karipidis et al., 2017, 2018). A pediatric protocol was used and the task was embedded in a story. Each part of the task included a different character type: to examine effects related

G. Pleisch et al. Neurolmage 189 (2019) 813–831

Table 1

| Variables | N   |
|-----------|-----|
| Sex (male/female) | 10/8 |
| Handedness (right/left) | 16/2 |
| False-font training set (1/2) | 10/8 |
| Age (years) | 6.7 ± 3.6 6.1–7.2 |
| IQ estimation (nonverbal) | 108 ± 13.3 85–125 |
| Familial risk (ARHQ) for dyslexia | 0.53 ± 0.2 0.26–0.8 |
| No. of children at increased familial risk (ARHQ >0.3) | 14 |
| No. of children with affected siblings/developmental language delay | 2/2 |
| Reading (one- or two-syllable words in upper-case letters, max. −20) | 2.9 ± 3.1 0–10 |
| Letter speech-sound knowledge (upper and lower case, max. −52) | 15.4 ± 3.6 2–35 |
| Letter speech-sound knowledge of letters in experiment (b,d,t,m,z, max. −6) | 1.2 ± 1.0 0–3 |
| Number knowledge (one- to three-digit numbers, max. −21) | 15.3 ± 5.0 7–21 |
| Number knowledge of numbers in experiment (1,2,3,4,5,6 max. −6) | 6 ± 0 |
| Rapid naming RAN (objects/second) | 27 ± 17 1–50 |
| No. of children exceeding RAN risk index | 6 |
| Phonological awareness | 47 ± 19 2–90 |
| No. of children exceeding phonological awareness risk index | 3 |
| Passive vocabulary | 49.8 ± 26.2 1–90 |
| Vocabulary – word meaning | 45.2 ± 28.9 10–93 |
| Non-word repetition (Mottier Test) | 34.2 ± 22.1 1–85 |
| Grapheme-Phoneme Correspondence Training (GPC) | **Mean ± SD Range** |
| GPC training duration (in min) | 19 ± 3.9 12–26 |
| GPC training accuracy (weighted, in %) | 81 ± 9 62–91 |

- **Block design test of the Wechsler Intelligence Scale for Children (HAWIK-IV).**
- **Highest parental adult reading history questionnaire value (ARHQ), risk indicated by values ≥3.**
- **Number of correctly named items.**
- **Matched percentage scores.**
to GPC training two different sets of FF characters were used of which one was trained prior to the EEG-fMRI session (TFF) whereas the other was only passively presented and served as control FF characters (CFF). The effect of expertise with different culturally meaningful character types was furthermore studied by presenting letters (LET), digits (DIG) as compared with the control false-font (CFF; Fig. 1a). All stimuli were presented in unimodal visual and auditory, and audiovisual congruent and incongruent conditions using Presentation® software (Version 16.4, www.neurobs.com). Every part consisted of 16 blocks (4 block-s/condition) whereby unimodal and bimodal blocks (15 items/block) alternated pseudorandomly separated by fixation periods of 6 or 12s. Six targets corresponding to the modalities of the conditions, i.e. either unimodal visual or unimodal auditory (cat, parrot, tortoise or shovel or their sounds) or bimodal presentation of a picture with sound (see Fig. 1a) requiring a button press were presented in addition to a total of 54 stimuli per condition to maintain children’s attention. The stimuli within each block were presented pseudorandomly for 613 ms with an interstimulus interval of 331/695 ms (Fig. 1b). Here, we focused on the unimodal visual condition (for analyses of auditory conditions see Karipidis et al. (2017)). Visual information was presented using video goggles (VisuaStimDigital, Resonance Technology, Northridge, CA), auditory information over in-ear headphones (MR confon GmbH, Magdeburg). Characters were presented in black in the middle of a grey background (mean visual angles horizontally/vertically TFF: 2.9/4.8; CFF: 2.7/4.8; LET: 2.8/4.8; DIG: 3/6.7°). In-scanner target detection accuracy (ACC) was high (89 ± 12.7%) and reasonable reaction times (RT) were recorded in all four parts and for all character types (In-line Supplementary Table S3). Performance did not significantly differ between the four parts (ACC: F (3,15) = 1.5, p = 0.227; RT: F (3, 15) = 0.9, p = 0.466). Responses of two participants were not logged due to technical problems and therefore not included in the response analysis.

2.5. EEG and fMRI acquisition

Using an MR-compatible 128-channel EEG system (Net Amps 400, EGI HydroCelGeodesic Sensor Net) simultaneous EEG-fMRI recordings were performed on a Philips Achieva 3 T scanner (Philips Medical Systems, Best, The Netherlands). Continuous EEG at a sampling rate of 1 kHz (DC-filter) was recorded with 128 scalp and two electrocardiogram (ECG) electrodes. To reduce gradient residuals during simultaneous EEG-fMRI recordings, the scanner clock and the EEG system were synchronised (Mandelkow et al., 2006). Electrode impedances were kept below 2.6 kΩ. The recording reference was located at Cz, the ground electrode (COM) posterior to Cz. Potential electrode vibration artifacts were minimized by covering the electrodes with a bandage retainer net and by turning off the helium pump of the MRI scanner during image acquisition.

A 32-elements receiver head coil was used to acquire 189 vol for each part of the task using a T2*-weighted whole-brain gradient echo-planar image sequence (EPI) with the following parameters: SofTone factor: 3, slices/volume: 31, repetition time (TR): 1.98s, echo time (TE): 30 ms, slice thickness: 3.5 mm, slice gap: 0.5 mm, flip angle: 80°, field of view (FOV): 24 × 24 cm², in plane resolution: 3 × 3mm², sensitivity-encoding reduction factor: 2.2. Specific emphasis was given on reducing scanner noise and improving auditory stimulation by using sound-absorbing over-ear headphones, a sound-absorbing mat in the MR-bore and a Sof-Tone sequence. A custom-made head pad for the EEG net was used, to reduce head movement and to ensure comfort. Additionally, a field map scan to perform B0 correction was recorded. T1-weighted images were recorded with a 3D MP-RAGE sequence (slices: 176, TR/TE: 6.8/3.2s, voxel size: 1x1x1mm³, flip angle: 9°, FOV: 27 × 25.4 cm²).

2.6. ERP analyses

Analyses were conducted using VisionAnalyzer 2.1 (BrainProducts GmbH, Munich, Germany). Channels with an overall poor data quality were topographically interpolated (range: 0–5 channels, mean: 1.57 channels SD: ±0.06). Due to continuous artifacts on the cheek electrodes, we excluded four electrodes from further processing and analyses (E43, E48, E119, E120). In addition, each data set was visually inspected and periods with major artifacts were manually excluded. After MR artefact removal using the average template subtraction method (Allen et al., 2000) and ballistocardiogram correction using sliding average template subtraction, the data was filtered (0.1–30 Hz and 50 Hz Notch) and down sampled (500 Hz). Independent component analysis (ICA; Jung et al., 2000) was applied to exclude blinks, eye movements, and residual ballistocardiogram artifacts. After artefact corrections, the data was re-ferenced to the average reference (Lehmann and Skrandies, 1980). Trials with remaining artifacts exceeding ±200 µV or identified by visual inspection were excluded. The data was segmented from −102 ms to 498 ms after visual presentation and averaged character type-wise. Core group (n = 18) grand averages included a mean of 41 epochs per character type (means: TFF = 45, CFF = 45, LET = 48, DIG = 40; range: 19–54 epochs). Using the global field power (GFP) maxima of the mean ERPs over all four character types, the interval of the N1 was defined as ±30 ms (194–254 ms) around the GFP peak. The mean amplitude values within these intervals were further analyzed. In the inline supplementary analyses 1, we also report statistics regarding effects related to training and character type processing in the preceding P1 ERP interval (102–162 ms).

2.6.1. Electrodes of interest analyses

To examine print specific activations over the posterior scalp sites for the N1 interval, the mean amplitudes over a left (LOT), middle (MO), and right (ROT) electrode cluster were computed (Fig. 2a). These left, middle and right electrode clusters comprised of the following electrodes (LOT: E65, E68, E69, E70, E73; MO: E81, E82, E75, E74; ROT: E83, E88, E89,
E90, E94). Statistical analyses for GFP and LOT/MO/ROT amplitudes were performed using linear mixed models (LMM), including repeated measurements within each subject (SAS 9.4, SAS Institute, Cary NC). To investigate training and character type effects random intercept models with fixed factors electrode cluster (LOT, MO, ROT), training (TFF, CFF) or character type (LET, DIG, CFF) and ARHQ as covariate of no interest, including the specific random intercept for each subject were computed (for details, see chapter 2.7.2). Post-hoc t-tests with Tukey-Kramer correction for multiple comparisons were performed. Correlational analysis (Fig. 2d) was performed to determine relations between trained false-font amplitudes and training performance (duration and weighted accuracy), using SPSS (Version 22.0). For normally distributed data Pearson correlation and for non-normally distributed data Spearman correlation was used.

2.7. FMRI analyses

Data was preprocessed and analyzed using SPM12 on MATLAB R2015b. After field map correction, images were spatially realigned and unwarped, slice time corrected, coregistered, segmented, and normalized using the deformations derived from the segmentation and a pediatric brain template created with the Template-O-Matic toolbox for the age range 5.9–8.5 years (Wilke et al., 2008). After resampling (3x3x3mm³), the data was smoothed with an isotropic 6 mm full width at half maximum Gaussian kernel. Volumes with more than 1.5 mm scan-to-scan movement were repaired by linear interpolation using the ArtRepair toolbox (Mazaika et al., 2011) and children with more than 10% repaired scans were excluded from analyses. None of the included data sets contained more than 6.35% repaired scans. Due to technical problems and excessive movement at the end of the task, the analyses of two children included only the first three visual blocks (instead of four blocks) for one of the parts. Including six predictors (auditory, visual, congruent, incongruent, target, and response) and six movement parameters for each participant and each part of the experiment (TFF, CFF, LET, DIG), a random-effect generalized linear model (GLM) was calculated. We report results of 2nd-level random effect analyses one-sample t-tests to characterize the general activation for each character type against baseline and 2nd-level t-tests based on first level contrasts to determine differences between visual character types of the experiment (Fig. 3a). Furthermore, we also computed whole-brain voxel-wise correlation of the BOLD response to trained false-fonts with training duration (Fig. 3b). ARHQ was included as covariate of no interest in all second level analyses.

2.7.1. Region of interest analyses

We performed ROI analyses based on the results of a meta-analysis (Martin et al., 2015) in the left ventral occipitotemporal cortex (lvOT) and its right hemispheric homologue (rvOT). A sphere with 8 mm radius (MNI coordinates x, y, z: ±52, −60, −14, Fig. 3c) was created using MarsBaR (Brett et al., 2002) with its centers in the inferior temporal gyri. Beta values of the bilateral ROIs were extracted and an LMM with fixed factors hemisphere, and training or character type and ARHQ as covariate of no interest was computed, including a specific random intercept for each subject. To compute LMMs and correlations, the same procedure as for the EEG analyses was used. To examine a potential number sensitive area in more detail, an additional analysis of a literature based spherical ROI in the number form area (NFA; Abboud et al., 2015) of the right hemisphere is presented in the inline supplementary analysis 2 (MNI coordinates x, y, z: 55, −43, −20; r = 8 mm).

Fig. 2. Training and character type effects in the N1 ERP (n = 18). (a) ERP waveforms for the four character types of the GFP and the three defined electrode clusters (LOT, MO, and ROT, see circled electrodes). The P1 (102–162 ms) and the N1 intervals (194–254 ms) are framed in grey. (b) Effect of training. Training induced N1 amplitude differences between TFF and CFF over middle and left occipital regions. c) Effect of character type. DIG showed most pronounced negativity over (right) posterior regions as also reflected in the t-maps. The color bars in b) & c) depict the mean amplitude values for each character type, pooled over the three clusters (LOT, MO, ROT) within the N1 interval. Above the bars, the corresponding potential field maps and below the bars the t-maps illustrating the statistical differences between the character types are shown. (d) Significant negative correlation of N1 GFP to TFF with training duration (r = −0.459, p = 0.043). Results of the extended EEG group are summarized in inline Supplementary Fig. S1.*
hemisphere*training: F(1,51) = 0.88, p = 0.3513; hemisphere*character type: F(2,81) = 0.08, p = 0.9193) we subsequently only included the main effects in the models (training/character type, cluster/hemisphere) in addition to the covariate ARHQ and report the results of these models in the main text. Only for the significant main effects, we also provide the effect size ($\eta^2$ Selya et al., 2012). For all LMM analyses studentized conditional residuals were computed to identify and exclude potential outliers. To correct for variance inhomogeneity, an outlier cutoff of three standard deviations from the mean was used for all analyses (Roth et al., 2007). The number of outliers for each LMM analysis is given in the online Supplementary Table S4. In addition, QQ-plots were inspected to ensure the assumption of normality and homoscedasticity of predicted versus residual conditional plots.

### 2.7.3. Connectivity analyses

Seed-to-voxel functional connectivity analysis was performed using weighted GLM as implemented in the CONN toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012). The normalized anatomical image of each participant was segmented into white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF) masks. Preprocessed functional data was band-pass filtered from 0.009 to 0.08 Hz and influences of motion, WM and CSF were regressed out using the CompCor strategy (Behzadi et al., 2007). To examine functional connectivity associated with GPC training and character type differences we defined a seed region (seed FFG, see Fig. 3d) within the anatomical left fusiform gyrus (FFG: Talairach Daemon (TD) database (Lancaster et al., 2000); WFU Pickatlas, version 2.4 (Maldjian et al., 2003)) that showed functional activation to either TFF or CFF (logical operation: FFG AND (TFF OR CFF)). The seed was defined using MarsBaR (Brett et al., 2002) and for functional activation we applied cluster-level FWEcorr p < 0.05 on a cluster-defining threshold (CDT) of p < 0.001. This left FFG seed was flipped to the right to also examine functional activation of the right hemispheric homologue.

For seed-voxel analyses, the residual time course for each seed was extracted and used to generate first-level correlation maps by computing Pearson’s correlation coefficients to the time course of all other voxels. To perform second-level GLM analyses, the first-level correlation coefficients were converted to normally distributed z-scores using the Fisher transformation. The individual familial risk score was entered as a between-subject covariate of no interest. In addition, we created a grey matter mask using the tissue probability mask of grey matter of the pediatric brain template. All voxels with a probability >0.5 were defined as grey matter. Within the grey matter mask, a cluster-based FWE-corr threshold of p < 0.05 was applied on a voxel-wise uncorrected threshold of p < 0.001. In analogy to whole-brain correlational fMRI analysis for TFF, we examined the correlation of FFG seed-to-voxel connectivity with training duration (Fig. 3d). Furthermore, one-sample and paired t-tests were computed on regression coefficients to yield functional connectivity maps for each character type (TFF, CFF, LET, DIG; see online Supplementary Fig. 4 and online supplementary Table 7) against baseline and their differences (training effect, character type differences).

### 3. Results

#### 3.1. Modulation of the visual N1 ERP by false-font training and character type expertise

**3.1.1. Effect of false-font training on N1 (Fig. 2b, inline Supplementary Fig. S1)**

The LMM of the N1 (194–254 ms) GFP mean amplitude with fixed factor character type (TFF, CFF) revealed no significant effect of training [F(1,17) = 0.01, p = 0.9298] suggesting no global amplitude differences. However, a significant negative correlation of training duration and the N1 GFP to TFF (r = −0.459, p = 0.043; Fig. 2d) was found, showing that shorter training duration was associated with an enhanced N1 negativity. When examining local differences in the posterior N1 negativity, the

---

**Fig. 3.** Modulation of the BOLD signal through training and character type (n = 18). (a) Left and middle columns: Whole brain activation by character type vs. baseline. Right column: Character type contrasts. LET showed stronger activation than CFF and DIG in the right superior temporal gyrus and the left middle frontal gyrus respectively (cluster-based p (FWEcorr) < 0.05). (b) Voxel-wise whole brain correlation of BOLD response to TFF with duration of the artificial GPC training. Faster learning correlated with significantly higher activation in the left vOT (cluster-based p (FWEcorr) < 0.05; for results of the extended fMRI group, see Supplementary Fig. S3). The negative correlation of the vOT activation plotted against training duration is shown on the left hand side for illustration purposes. (c) ROI analyses in bilateral inferior temporal gyri of the vOT. Left: Mean beta values for TFF and CFF combined over the left and right vOT ROI showing a significant effect of training (TFF > CFF) as revealed in post-hoc tests (* pcorr < 0.05). Right: Mean beta values for LET, CFF and DIG combined over the left and right vOT ROI showing a significant effect of character type (LET > CFF). (d) Functional connectivity. Correlation of the left FFG seed region (blue) connectivity to superior parietal regions (red) for TFF with decreasing training duration (cluster-based p (FWEcorr) < 0.05). Error bars denote standard errors. LET = letter, TFF = trained false-font, CFF = control false-font, DIG = digit.

**2.7.2. Linear mixed model (LMM) analyses in ERP and ROI data**

As detailed in the corresponding method sections, we used linear mixed model (LMM) analyses and post-hoc t-tests (Tukey Kramer corrected) to analyze ERP N1 and fMRI ROI data regarding effects of false-font training (TFF vs CFF) or character type (LET, DIG, CFF). The random effect consists of a subject dependent random intercept. In LMM fixed and random effects explain differences between subjects and the variability within subjects respectively. First, we used LMM models including main effects and interaction terms for all models. Because none of the models (N1, vOT ROI) yielded any significant interaction effects (for N1 amplitude: cluster*training F (2,85) = 0.82, p = 0.4449; cluster*character type F (4,133) = 0.09, p = 0.9849; for ROI vOT beta values:...
LMM for the three electrode clusters with fixed factors training (TFF, CFF) and cluster (LOT, MO, ROT) revealed a significant training effect \( F(1,87) = 9.30, p = 0.0030, \hat{\beta}^2 = 0.1107 \), but no main effect of cluster \( F(2,87) = 0.83, p = 0.4375 \). The post hoc t-test analysis revealed a significantly more negative N1 for TFF than for CFF (TFF < CFF: \( t(87) = -3.05 \)).

3.1.2. Effect of character type on N1 (Fig. 2c, inline Supplementary Fig. S1)

The LMM of the N1 GFP mean amplitude with fixed factor character type (LET, DIG, CFF) revealed no significant main effect \( F(2,34) = 1.86, p = 0.1719 \) suggesting no global amplitude differences. The analysis regarding character type based on an LMM for the three electrode clusters with the fixed factor character type (LET, DIG, CFF) and cluster (LOT, MO, ROT) revealed a significant character type effect \( F(2,137) = 40.64, p < 0.0001, \hat{\beta}^2 = 0.2034 \), but no effect of electrode cluster \( F(2,137) = 1.56, p = 0.3572 \). The character type effect was driven by the significantly stronger N1 for DIG than for the other types (post hoc t-tests: DIG < LET: \( t(137) = 4.51, p < 0.0001 \), DIG < CFF: \( t(137) = 8.93, p < 0.0001 \)).

For additional analyses of training and character type in the preceding P1 interval, please see inline supplementary analysis 1. The results of the extended group \( n = 23 \) supporting the results of the core group regarding the training and the character type effects are presented in the inline Supplementary Fig. S1.

To summarize, N1 ERP results in the core and extended groups show a pronounced training-related difference between TFF and CFF and a character type-dependent modulation of the N1 amplitude when comparing DIG, LET and CFF over the occipitotemporal scalp.

3.2. Modulation of the BOLD signal related to training and character type expertise

Next, we report results of second-level whole brain voxel-wise random effect analyses to characterize activation differences evoked by visual processing of the four character types using a cluster-based family-wise error corrected (FWE\(_{\text{corr}}\)) threshold of \( p < 0.05 \) (on a cluster-defining threshold (CDT) of \( p < 0.001 \)). As expected, all four character types showed pronounced occipitotemporal activation. This activation was bilateral for TFF, CFF and LET but only reached significance in the right hemisphere for DIG (Fig. 3a & Inline Supplementary Table S5).

3.2.1. Effect of training

No cluster survived cluster extent correction for the contrast of TFF vs CFF on whole brain level, but because of a special focus on this contrast the uncorrected results \( p < 0.001 \) are reported and illustrated in the inline Supplementary Fig. S2 and Table S6).

The training duration of the artificial GPC training correlated significantly with the TFF BOLD response in the left vOT but no other region, demonstrating that a higher activation was associated with faster learning (Fig. 3b, Table 2). Importantly, this whole brain finding survived an even stronger correction \( p (\text{FWE}_{\text{corr}}) < 0.01 \), using a CDT of \( p < 0.001 \), was also confirmed by the supportive analyses of the enlarged fMRI-group (Inline Supplementary Fig. S3) and thus can clearly be considered as a robust result, unaffected by potential inflated false-positive rates (Eklund et al., 2016).

3.2.2. Effect of character type

Direct character type contrasts revealed only minor differences in whole brain analyses, mostly driven by single letter processing: LET showed more activation than CFF in the right superior temporal gyrus and the left middle frontal gyrus extending to superior frontal areas when compared with DIG (Fig. 3a, Table 2).

To summarize the whole brain results, single characters activate predominantly bilateral occipitotemporal areas, training duration correlated with activation to trained false-font characters in the left vOT and direct character type contrasts yielded only minor differences in temporal and frontal areas.

3.3. ROI analysis in the left and right vOT

The emergence of preferential activation to print in the vOT was further investigated in two a-priori bilateral, literature defined ROIs. Similar to the analysis of the visual ERP N1 we performed two separate LMMs to examine effects related to false-font training and character type expertise.

3.3.1. Effect of false-font training on vOT BOLD

First, the LMM analysis with fixed factors hemisphere (lVOt, rVOt) and training (TFF, CFF) revealed a significant main effect for training \( F(1,52) = 4.44, p = 0.0399, \hat{\beta}^2 = 0.0833 \). The training effect was driven by the significantly stronger activation for TFF compared to CFF (post-hoc t-tests: \( t(52) = 2.11 \)); see also inline Supplementary Fig. S3b for the extended fMRI group).

3.3.2. Effect of character type on vOT BOLD

The second LMM analysis with fixed factors hemisphere (lVOt, rVOt) and character type (LET, DIG, CFF) revealed significant effects of hemisphere \( F(1,83) = 4.08, p = 0.0467, \hat{\beta}^2 = 0.0488 \) and character type \( F(2,83) = 3.22, p = 0.0452, \hat{\beta}^2 = 0.0737 \). Post hoc t-tests revealed a significantly stronger activation in the right hemisphere than in the left hemisphere (rVOt > lVOt: \( t(83) = 2.02 \)) and a significantly stronger activation for LET compared to CFF (LET > CFF: \( t(83) = 2.43, p = 0.0447 \)).

Additional analyses of the character type effect in a potential number form area ROI, did not yield any differences between character types (inline supplementary analysis 2).

To summarize, in concordance with the ERP N1 results also the ROI results in the vOT exhibited a training effect in the form of enhanced beta values for TFF than CFF. An effect of character type in the vOT was specifically found for LET as compared with CFF. The ROI analysis furthermore pointed to enhanced activation in the right vOT. The effects of training and character type were confirmed by additional analyses of the extended fMRI group. Importantly, the analyses of the extended group also indicated a significant difference between DIG and CFF, suggesting initial alphanumeric tuning (see inline Supplementary Fig. S3b).

---

Table 2

| Correlations | p(FWEcorr) | k | T | x | y | z | Brain region |
|--------------|-----------|---|---|---|---|---|--------------|
| Training duration & TFF BOLD connectivity left FFG seed | 0.00901 | 63 | -6.76 | -35 | -81 | -9 | L IOG/ LING |

Listed are cluster size k, peak voxel T and the MNI coordinates (x, y, z) using a cluster-based family-wise error corrected (FWE\(_{\text{corr}}\)) threshold of \( p < 0.05 \) (on a cluster-defining threshold (CDT) of \( p < 0.001 \)).
3.4. Enhanced functional connectivity to superior parietal/lateral occipital regions for trained false-fonts

Second level random effect results (FWERcor, of $p < 0.05$, using a CDT of $p < 0.001$) of bilateral FFG seed-based functional connectivity analyses are summarized in the online Supplementary Fig. S4 and inline Table 7 for each character type and the corresponding contrasts.

There was no significant difference in the functional connectivity from the left or right FFG seed regions between TFF and CFF. The functional connectivity of the left FFG seed region for TFF showed a significant negative correlation with training duration to a cluster in the left superior parietal gyrus/lateral occipital cortex (LOC) (Fig. 3d, Table 2).

4. Discussion

Studying the emerging functional activation preference to print in visual areas as a result of letter-speech sound training in the developing brain is highly relevant because it can inform about future reading outcomes of children. Such insights are especially important for children at an increased risk for developing reading problems such as the group examined in this study. Preschool neural measures related to language processing and learning performance may critically contribute to the early identification of children with poor reading outcomes (Bach et al., 2013; Brem et al., 2013; Hoeft et al., 2007b; Karipidis et al., 2018; Maurer et al., 2009; Raschle et al., 2012) and to provide timely and efficient support for struggling learners.

Focused on the case of the well-known functional specialization of the left vOT/N1 to print in literates, this study aimed to clarify which processes drive the often reported preferential activation to print by comparing the effect of varying expertise and visual familiarity on the processing of characters in prereading children at varying risk for developmental dyslexia. Our approach included, first, an artificial GPC training that led to learning of phonological associations to print. After training, we examined the sensitivity of vOT/N1 activations to the trained false-fonts compared to control false-fonts that children had passively viewed but not actively associated with speech sounds. We further relate the activation and functional connectivity to trained false-font characters to learning performance in the preceding training session. Secondly, we compared activation patterns of artificial false-font characters to culturally meaningful characters, i.e., letters and digits, for which preschool children exhibit different levels of expertise. This comparison yielded further insights into how different levels of expertise related to phonological, semantic and magnitude associations may shape the vOT and N1 specialization in the brain of preschool children.

4.1. Neural specialization to print after phonological association training and modulation by training performance

Learning to associate artificial characters to speech sounds was reflected in preferential activation for trained compared with passively viewed characters in the visual N1 ERP and in the vOT cortex BOLD response. This result demonstrates the development of preferential activation after short artificial GPC training and critically extends previous knowledge about the development of functional specialization to print in children (Brem et al., 2010; Gentanni et al., 2018; Dehaene-Lambertz et al., 2018; Fraga Gonzalez et al., 2016; Fraga Gonzalez et al., 2014; James, 2010; Maurer et al., 2006). We show that a short (<30min) artificial GPC training already induces rapid adaptation in the functional response of the ventral occipitotemporal cortex in preschool children. Changes in the vOT activity (Hashimoto and Sakai, 2004; Song et al., 2010; Xue and Poldrack, 2007) and stronger N1 activation have previously been reported when adults learn print-like stimuli (Brem et al., 2017; Maurer et al., 2010; McCandliss et al., 1997) and with increasing literacy (Dehaene et al., 2010; Pegado et al., 2014). Studies in poor reading children within the first years of school, related N1 sensitivity to words to the outcome of a relatively short grapheme-phoneme training (Fraga Gonzalez et al., 2016) and to the children’s response to reading intervention (Molfese et al., 2013). In addition, preferential activation to real words has also been shown in prereaders by training real grapheme-phoneme correspondences over several weeks, even though the children’s reading skills were still rudimental after training (Brem et al., 2010).

Such training effects suggest that learning grapheme - phoneme correspondences is a key factor of specialization processes and preferential activation to print in the vOT. In addition, here we quantified the ease of learning by the duration of the adaptive artificial GPC training. Accounting for the ease of learning in the GPC training revealed that the faster children learned new associations, the stronger was the corresponding preferential response of the vOT and N1 to single trained false-fonts and the functional connectivity of the left FFG seed region to the left SPL/LOC. This result supports the notion of expertise-dependent activation in the vOT (Maurer et al., 2005, 2011; Maurer and McCandliss, 2007; Price and Devlin, 2011) emerging in the prereading brain through graphophonological association training. The preferential vOT activation to trained characters may be indicating a process of building up abstract neural letter form representations, similar to the notion of the VWFA as a “prelexical hub” for computation and storing abstract orthographic word forms (Dehaene et al., 2002, 2005).

Increasing visual familiarity through training may partly explain the results of this study. The negative correlations of the N1 and BOLD signals with training duration, however, clearly argue against a pure visual familiarity effect because longer training durations, and thus more exposure to the false-fonts, were associated with decreased and not increased brain responses to trained false-font characters in the children. Instead, the negative correlation of the N1 and vOT activation with training duration, reflecting performance in the artificial GPC training, supports the notion of a visual expertise effect related to learning of phonological associations to print. Thus, the preferential activation to trained characters in the vOT may result from establishing novel circuits to higher order cognitive areas in the prereading brain, which may in turn, provide predictions about the content of the visual input (Price and Devlin, 2011). In such a model, the left vOT could be best described as an “integrative hub”, matching and evaluating the information received from different brain regions (Carreiras et al., 2014; Price and Devlin, 2011). The negative correlation of increased functional connectivity between the left FFG seed and the left superior parietal/lateral occipital cortex with training duration is in line with previous findings of character-speech sound association learning in adults indicating that this region exerts a critical role in cortical plasticity related to learning novel letters (Hashimoto and Sakai, 2004). Our results in young children show that this is also the case during the prereading stage. Together with the ERP and the functional activation findings our connectivity results indicate a phonologically guided tuning of the preferential vOT/N1 activation.

4.2. vOT/N1 activation to character types with varying expertise

In addition to the experimental manipulation of expertise through phonological association training, we also compared character types (LET, DIG vs. CFF) naturally varying in their level of expertise regarding phonological, semantic and/or magnitude associations that is built up and refined over the course of child development. Our analyses revealed that different character types show distinct encoding in bilateral vOT and in the corresponding visual N1 at preschool age already.

First, the N1 ERP mean amplitude showed the most pronounced activation over all character types for digits. N1 amplitude modulations have been strongly associated with the level of expertise in specific visual categories (Rosston et al., 2002; Tanaka and Curran, 2001), including print (Brem et al., 2005; Maurer et al., 2005). The pronounced ERPs to digits may thus simply reflect the advanced neural specialization to this visual character category, with high expertise at preschool age. However, purely visual specialization developed through increased familiarity and
exposure would hardly explain why letters showed weaker N1 amplitudes than digits, as we would expect ample visual familiarity with both letters and digits at this age. Thus, the increased N1 responses of prereaders to digits seem to be driven by expertise related to phonological, semantic, and magnitude associations.

This coincides nicely with the description that gaining visual experience goes along with the recruitment and involvement of multiple high-level areas and the notion of visual expertise as an “enhanced, ecological form of visual object recognition that emerges from the reciprocal interactions between multiple top-down factors, such as semantic knowledge, attention, and task relevance, and the visual system.” (Harel (2016), p 88). Unlike the ERP results, an equal BOLD response was detected for digits and letters in the vOT cortex. Our results are thus in line with the findings of Cantlon et al. (2011), showing no fine tuning between numbers and letters in prereaders in our implicit task. Neither was such fine tuning observed for digits and letters in a region specialized for number-form processing in adults or school children (Abboud et al., 2015; Shum et al., 2013). This suggests the absence of preferential activation to digits over letters or false-fonts in the putative number-form area at this developmental stage. Our results coincide with those of Cantlon et al. (2011) in showing higher activation to alphanumeric characters (letters, digits) than to false-font characters. This alphanumeric tuning only approached a trend in the main analysis but was significant in the enlarged fMRI sample (see online supplement Fig. 3b). Moreover, our pattern of alphanumeric tuning in bilateral vOT extends the findings of the former study in showing an even finer level of tuning, because the processing difference between objects and alphanumeric characters (Cantlon et al., 2011) may be expected to be greater than between the alphanumeric and matched false-font characters used in our study.

Given the absence of any processing difference between letters and digits in the vOT, it is difficult to disentangle the contribution to vOT specialization of visual familiarity from that of expertise with phonological and magnitude information. Visual familiarity clearly contributes to the specialization of vOT activation, but mere visual familiarity does not explain why letters and trained false-fonts resulted in similar vOT BOLD and N1 activation (additional direct comparisons of TFF and LET yielded no significant differences neither in the N1 ERP nor vOT BOLD responses: both pcorr > 0.6). The children were exposed to the trained false-fonts for only about 20 min; it is unlikely that such brief training led to visual familiarity comparable to the accumulated visual exposure to letters over the first six years of life.

Alternatively, with regard to the suggestion of the VOT as an integrative hub, temporally co-occurring predictions from multiple cognitive systems such as phonology, semantics and/or magnitude may induce higher prediction errors as compared to character types where none (untrained false-fonts) or only weak (letters) predictions are expected given the low level of phonological or lexical expertise for such characters (Brem et al., 2010; Maurer et al., 2005; Park et al., 2014). Such a model would also account for the differential N1 amplitude to trained and untrained false-fonts reported in the previous section, because relatively stable phonological associations may initiate feedback projections from phonological areas.

4.3. Contribution of attentional mechanisms and lateralization of functional specialization

A series of previous studies suggest that the activation of the vOT or N1 is modulated by attentional mechanisms through its link to the dorsal attention network (Christophel et al., 2018; Cohen et al., 2008; Luck et al., 2000; Vogel et al., 2012; Vogel and Luck, 2000; Yoncheva et al., 2015). Vogel et al. (2012) used resting state functional connectivity analyses to show that the posterior part of the visual word form system is strongly connected to the dorsal attention system including the frontal and inferior parietal cortex. This connectivity suggests an important role in directing attention to the critical information within words and sentences. Importantly, this connectivity to the dorsal attention system was diminished in children and dependent on reading ability, indicating that the contribution of the attentional network becomes more important with fluent reading (Vogel et al., 2012).

Other studies examined the effects of attentional processes on the visual N1. The ERP study by Yoncheva et al. (2015) showed that attentional focus on either single grapheme–phoneme correspondences or whole words during training has an effect on how words consisting of such trained characters are explicitly processed in the N1 after training: A more left-lateralized N1 negativity was found after training graphophonological associations but not after whole-word association training. Our findings of activation preference to specific visual character types (TFF, DIG, LET) reflected in both the vOT and N1 ERP activation may thus also be attributed to the contribution of attentional mechanisms during the graphophonological association training or during the implicit task. The greater functional connectivity from the vOT to the left SPL/LOC for trained false-font characters with shorter training duration may correspondingly indicate the increased involvement of general attentional mechanisms of a parietal network in processing the newly learned false-font characters. Previous studies emphasized the increased involvement of the posterior parietal cortex in preorthographic processing of multi-element character strings (Lobier et al., 2012) and showed higher connectivity from VWFA to SPL in normal reading as compared to children with dyslexia (van der Mark et al., 2011). Importantly, effects on the N1/vOT in previous studies were especially pronounced when attention was specifically manipulated, such as in spatial attention allocation tasks (Luck et al., 2000), in working memory tasks with explicit recall of attended or unattended stimuli (Christophel et al., 2018), when comparing choice reaction to simple reaction time tasks (Vogel and Luck, 2000), and in object expertise tasks with attended and unattended items (Harel, 2016). No such explicit manipulation of attentional focus on our core stimuli (TFF, CFF, LET, DIG) was done in our implicit target detection task. Moreover, attention was directed to unrelated and qualitatively different target items (drawings and environmental sounds), and it is thus rather unlikely that activation differences in character type processing are solely attributable to attention allocation in our task. However, experimental paradigms that specifically manipulate control over attentional top-down influences during task performance would be useful in improving our understanding and clarifying the potential impact of attention on vOT specialization.

4.4. Neural specialization to characters beyond the vOT

In our fMRI analysis, direct character type contrasts at whole-brain level revealed only minor differences, mostly driven by single letter processing. First, the activation pattern in the right superior temporal gyrus (STG) was mainly driven by higher activation to letters than to passively viewed false-font characters. The STG has shown anatomical alterations in relation to reading skills, such as reductions in grey matter volume in dyslexic children (Richlan et al., 2013) and in at-risk children at prereading age (Black et al., 2012; Raschle et al., 2011a) and increased grey matter volume in subjects who learned to read in adulthood (Carreiras et al., 2009). Learning to read and increased reading performance is usually accompanied by increases in lateralization of occipitotemporal reading networks to the left hemisphere (Maurer et al., 2007; Shaywitz et al., 2007; Turkeltaub et al., 2003). In our young children, BOLD analyses of the vOT and whole brain analyses indicate some stronger engagement of the right than the left hemisphere. Whether or not this stronger right-hemispheric activation and the enhanced activation for letters in the right STG reflect some compensatory processing (Shaywitz et al., 2003) or atypical lateralization of language processing in children at heightened risk for dyslexia (Brem et al., 2013; Hoefle et al., 2011) needs to be clarified in future studies. Nevertheless, the STG region has been associated with the processing of speech and phonology, and the left STG in particular has been implicated in graphophonological decoding (Jobard et al., 2003). Given the lower hemispheric
specialization in young children (Ossowski and Behrmann, 2015) increased activation in the right hemisphere could also indicate early attempts to match letter and speech-sound information (Blau et al., 2010; van Attevelde et al., 2004). Second, letters showed stronger activation in the left middle frontal gyrus (MFG) of the dorsolateral prefrontal cortex than did digits, which could indicate increased attentional resources and/or enhanced control functions implicated in processing characters with less expertise (Houdé et al., 2010).

4.5. Limitations and further questions

Several limitations of the present study should be discussed. The first is the rather small sample examined here. The application of simultaneous EEG and fMRI in children as young as kindergarteners is unique and allows the analysis and comparison of emerging specializations with high temporal and spatial resolution. This approach made visible convergent patterns of vOT BOLD and N1 ERP, such as enhanced activation to false-fonts after training and modulation by training duration. However, it also revealed divergent ERP/BOLD patterns such as the dominant N1 negativity over the posterior scalp to digits as compared with the more general alphanumeric tuning in the vOT BOLD signal. However, such simultaneous EEG-fMRI is also challenging: Motion artifacts distort both fMRI and EEG data quality, and the limited test time available with such young children affects the duration, number of trials, signal-to-noise ratio and robustness of the task. As a consequence, and given also the stringent motion artefact exclusion criteria that we applied, only 18 children with sufficient EEG and fMRI data quality remained for the core analyses, and this sample included two left-handed children. Despite a higher incidence of right dominant and bilateral language networks (Szafarski et al., 2002) in left-handed subjects than right-handed ones, we still included the two left-handed children. Language lateralization is not expected to be fully mature in our preschool children and left-hemispheric language lateralization dominates in left-handed subjects (Somers et al., 2015; Szafarski et al., 2002). We addressed these problems by providing additional analyses of extended ERP and fMRI groups including 23 or 24 children in the inline supplement and by verifying these analyses without the left-handed children. These analyses largely support the analyses of the smaller core sample.

Secondly, our preschool children all had a varying risk for developmental dyslexia, most of them either from the parents’ or a sibling’s reading history. A better understanding of graphophonological association learning and emerging neural specialization in such children is of special relevance for improving the early identification of children with poor reading outcomes. However, our focus on at-risk children raises the question of how representative our results are for the general population, given that at-risk children have shown alterations in brain structure and function in previous studies (Karipidis et al., 2017; Maurer et al., 2003; Raschle et al., 2011a). It should be noted that our children had a varying familial risk level, as can be seen in the widely differing ARHQ values. To address the varying risk level, we corrected for the individual familial risk score by including the maximum ARHQ risk factor in all statistical analyses as a covariate of no interest, and we also address the potential differences of such a risk group in the discussion section. However, it is essential that future studies examine emerging specialization in a low-risk sample.

Finally, we acknowledge that it is difficult to quantify the degree of visual familiarity or expertise children gained with the control false-font characters during the training, because they were only passively presented and not actively trained. It may have been the case that some children did not attend to these stimuli and did not gain visual familiarity with these control characters. Therefore, we cannot fully disentangle the impact of phonological association learning from the effect of increasing visual familiarity by comparing trained and passively viewed false-font characters in this study. Even though it is unlikely, given our explanations in the previous sections, that the difference between trained and control false-fonts relied entirely on visual processes, this aspect should be tested in a future study by having children train two sets of characters, one with phonological associations and the other in a purely visual approach. Such a design, when also controlling for attentional factors, could further help to quantify the contributions of different aspects and processes to neural specialization of the vOT.

5. Conclusion

Here, we show altered visual character processing in the prereading brain after short artificial grapheme-phoneme correspondence training (<30 min) as a model mimicking one of the first steps in reading acquisition. Emerging functional specialization was reflected in more pronounced activation for trained false-font characters than for untrained ones, irrespective of imaging modality, and a strong relationship between the ease of learning and activation in the left vOT. Moreover, we show distinct expertise-dependent activation differences for the vOT and the visual N1, including emerging alphanumeric tuning in the visual word from system at that early age. As grapheme-phoneme learning is considered the core principle of acquiring alphabetic languages (Blomert, 2011), learning performance in such model training and the associated level of neural specialization in prereaders might indicate later success in reading acquisition (Karipidis et al., 2018). These novel insights into learning-dependent early specialization of the vOT to print may thus contribute to the identification of prereaders at risk for reading difficulties and hence, allow the development of early targeted interventions. Finally, the current insights also pinpoint the parallel development of functional circuits and cortical specialization in the developing brain during learning.

Competing financial interests

The authors declare no competing financial interests.

Further information and requests for resources should be directed and will be fulfilled by the Lead Contact, Silvia Brem (sbrem@kjpd.uzh.ch). Some restrictions apply for the sharing of raw and processed data for ethical reasons, because this would compromise participant confidentiality and privacy. The artificial grapheme-phoneme correspondence training was developed at the Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric Hospital, University of Zurich using the GraphoGame platform provided by the University Jyväskylä and is subjected to copyright.

Acknowledgements

This study was financed by the Swiss National Science Foundation (grant: 32003B_141201), the Hartmann-Müller Foundation (grant: 1912) and the Olga Mayenfisch Foundation. We thank A. Baur, D. Dornbierer, A. Roth, and M. Schneebeli for their assistance during recruitment, investigation, and data analyses and D. Brandeis for his valuable inputs on the text. We are grateful to U. Richardson and I. Kapanen (University of Jyväskylä, Finland) for their support in implementing the artificial GPC training. We thank G. Fraga Gonzalez for critical reading and commenting of this manuscript. We thank all the participating children with their families.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.01.046.
Appendix B. Inline supplement proofs

Inline Supplementary Table S1
Characteristics and performance on behavioral measures for 23 (enlarged EEG-group) and 24 (enlarged fMRI-group) subjects included in supplemental analyses

| Enlarged groups for supportive analyses | EEG-group | fMRI-group |
|----------------------------------------|-----------|-----------|
| N                                      | 23        | 24        |
| Sex (male/female)                      | 13/10     | 13/11     |
| Handedness (right/left)                | 20/3      | 21/3      |
| False-font training set (1/2)          | 12/11     | 12/12     |
| Mean ± SD                             | 6.8 ± 3.6 | 6.1 ± 3.7 |
| Range                                  | 80–140    | 6.8 ± 3.3 |
| False-font training set (1/2)          | 20/3      | 21/3      |
| Mean ± SD                             | 107 ± 10  | 85–140    |
| Range                                  | 0.26–0.8  | 0.52 ± 0.14 |
| Number of correctly named items        | 3         | 2         |
| Highest parental adult reading history questionnaire value (ARHQ), risk indicated by values > .3 | 0.53 ± 0.17 | 0.26–0.8 |
| IQ estimation (nonverbal)              | 105 ± 14  | 107 ± 10  |
| Number of children exceeding phonological awareness risk index | 4         | 3         |
| No. of children at increased familial risk (ARHQ > 0.3) | 19        | 20        |
| Reading (one- or two-syllable words in upper case letters, max. – 20) | 4.3 ± 4.4 | 0.09–19 |
| No. of children with siblings/developmental language delay | 2/2       | 2/2       |
| Passive vocabulary                     | 53.8 ± 26.6 | 56.1 ± 27.3 |
| Vocabulary - word meaning              | 10.3 ± 5.2 | 85.0–100 |
| Letter-sound knowledge of upper case letters | 16.7 ± 10.4 | 17.5 ± 8.2 |
| No. of children exceeding RAN risk index | 6 ± 0     | 6 ± 0     |
| No. of children exceeding phonological awareness risk index | 7         | 6         |
| Letter-sound knowledge of lower case letters | 49 ± 22.4 | 50.9 ± 22.1 |
| N. of children exceeding phonological awareness index | 3         | 3         |
| Letter-sound knowledge of upper case letters | 31.2 ± 23.3 | 35.7 ± 23.3 |
| Letter-sound knowledge of lower case letters | 29.5 ± 20.0 | 30.1 ± 19.1 |
| GPC Training accuracy (weighted in %)  | 81 ± 9.4  | 81 ± 10.6 |
| GPC Training duration (min)            | 1147      | 1118      |
| IQ estimation                          | –0.055    | 0.087     |
| Familial risk for dyslexia             | –0.453(∗) | –0.313    |
| RAN objects                            | 0.370     | –0.469∗   |
| Letter-sound knowledge of upper case letters | 0.494∗    | –0.160    |
| Letter-sound knowledge of lower case letters | 0.294    | 0.043     |
| Non-word repetition (Mottier Test)     | 0.301     | –0.492∗   |
| Passive vocabulary                     | –0.038    | 0.180     |
| Vocabulary (word meaning)              | 0.265     | –0.200    |
| Training achievement %                 | –0.567∗   | –         |
| Training rate in minutes               | 0.185     | –0.495∗   |

Correlation matrix with r values: ***P < 0.001 **P < 0.01, *P < 0.05, (∗) P < 0.1.
1 highest parental ARHQ score.
2 percentile scores.
3 raw value.
4 trial-wise item-weighted accuracy.
5 mean amplitude of the global field power of the TFF condition.

Inline Supplementary Table S2
Correlational analyses (n = 18)

| Measure                             | Training achievement % | Training duration in minutes |
|-------------------------------------|------------------------|-----------------------------|
| IQ estimate                         | –0.055                 | 0.087                       |
| Familial risk for dyslexia          | –0.453(∗)              | –0.313                      |
| RAN objects                         | 0.370                  | –0.469∗                     |
| Letter-sound knowledge of upper case letters | 0.494∗              | –0.160                      |
| Letter-sound knowledge of lower case letters | 0.294           | 0.043                       |
| Non-word repetition                 | 0.301                  | –0.492∗                     |
| Passive vocabulary                  | –0.038                 | 0.180                       |
| Vocabulary (word meaning)           | 0.265                  | –0.200                      |
| Training achievement %              | –0.567∗                | –                           |
| Training rate in minutes            | 0.185                  | –0.495∗                     |

| Measure                             | Mean ± SD | Range |
|-------------------------------------|-----------|-------|
| In-scanner accuracy TFF (%)         | 92 ± 11.2 | 60–100 |
| In-scanner reaction time TFF (ms)   | 755 ± 166.9 | 579–1118 |
| In-scanner accuracy CFF (%)         | 89 ± 12.7% | 54–100 |
| In-scanner reaction time CFF (ms)   | 719 ± 132.2 | 533–905 |
| In-scanner accuracy ET (%)          | 97 ± 7.3 | 71–100 |
| In-scanner reaction time ET (ms)    | 711 ± 133.5 | 571–1005 |
| In-scanner accuracy DIG (%)         | 90 ± 13.4% | 66–100 |
| In-scanner reaction time DIG (ms)   | 761 ± 188.6 | 512–1147 |

In-scanner performance in the audiovisual target detection experiment for the core group
### Inline Supplementary Table S4

Number of outlier data (not cases) excluded for each LMM

| Analyses of the main text and inline supplement | N = 18 | N = 23 (ERP)/N = 24 (ROI) |
|-----------------------------------------------|-------|---------------------------|
| Number of excluded outlier data sets (>3SD)   | N = 18 | N = 23 |
| N1 GFP                                        | 0     | 1             |
| N1 cluster                                    | 0     | 0             |
| ROI VWFA                                      | 0     | 6             |

### Inline Supplementary Table S5

Activations for each character type vs baseline in the core group (n = 18)

| Coord     | cluster size | peak voxel | MNI coordinates | Hemisphere | Brain region |
|-----------|--------------|------------|-----------------|------------|--------------|
| 0.00001   | 168          | 6.86       | -47             | -69        | 1            | MOG          |
| 0.00000   | 471          | 5.93       | 43              | -63        | -12          | ITG          |
| 0.00000   | 449          | 7.03       | 43              | -72        | -6           | ITG          |
| 0.02993   | 44           | 4.88       | -32             | -93        | -3           | MOG          |

LET = letter, TFF = trained false font, CFF = control false font, MOG = middle occipital gyrus, MOG = middle occipital gyrus, IOG = inferior occipital gyrus, FFG = fusiform gyrus, OFC = orbitofrontal cortex, ITG = inferior temporal gyrus, MTG = middle temporal gyrus, SFG = superior frontal gyrus, SPG = superior parietal gyrus, l = left, r = right.

### Inline Supplementary Table S6

Main activation peaks of the second level t-tests for the contrast TFF vs. CFF

| Coord     | cluster size | peak voxel | MNI coordinates | Hemisphere | Brain region |
|-----------|--------------|------------|-----------------|------------|--------------|
| 0.58486   | 17           | 5.32       | 49              | -12        | -15          | r MTG        |
| 0.46249   | 20           | 4.98       | -56             | -6         | -18          | l MTG        |

Listed are MNI coordinates (x, y, z) of cluster maxima at an uncorrected p < 0.001, k = 15 corresponding to a cluster-wise error corrected threshold (p(FWEcorr) < 0.05). LET = letter, TFF = trained false font, CFF = control false font, MTG = middle temporal gyrus, l = left, r = right.
## Functional connectivity seed-voxel analyses

| Brain region       | p (FWEcorr) | k | T  | x   | y   | z   | p (FWEcorr) | Brain region       |
|--------------------|-------------|---|----|-----|-----|-----|-------------|--------------------|
| **Left FFG seed**  |             |   |    |     |     |     |             |                    |
| TFF vs baseline    | 0.001       | 180| 9.53|    |    | -46| -66         | -16                |
| CFF vs baseline    | <0.001      | 429| 7.84|    | -40| -64| -12         | ITG/FFG            |
| DIG vs baseline    | 0.009       | 116| 7.48| 50  |    | 48 | -6          | Frontal pole       |
| CFF vs baseline    | <0.001      | 193| 6.75| 50  |    | -72| 6           | LOC                |
| DIG > CFF          | <0.001      | 220| 6.10| -8  | 50 | -12| 0           | ACG paracingulate gyrus |
| CFF > DIG          | 0.008       | 114| 6.19| 40  |    | -36| 30          | MFG                |
|                   | 0.030       | 8  | 5.81| -40 | 36 | 24 | MFG         |
| **Right FFG seed** |             |   |    |     |     |     |             |                    |
| TFF vs baseline    | <0.001      | 274| 7.21| 52  |    | -66| -18         | ITG/FFG            |
| CFF vs baseline    | <0.001      | 215| 6.88| 46  |    | -60| -18         | ITG/FFG            |
| CFF vs baseline    | <0.001      | 134| 5.15| -2  | 48 | 0 | ACG paracingulate gyrus |
| DIG vs baseline    | 0.009       | 121| 6.08| -46 |    | -66| 20          | MFG                |
|                   | 0.040       | 89 | 5.45| 58  |    | -60| 12          | STG                |
| DIG vs baseline    | 0.009       | 123| 7.74| 38  |    | -70| -12         | ITG/FFG            |

Listed are MNI coordinates (x, y, z) of cluster maxima at $p(CD)<0.001$, FWE<$0.05$. TFF = trained false font, CFF = control false font, LET = letter, DIG = digit; MTG = middle temporal gyrus, ITG = inferior temporal gyrus, FFG = fusiform gyrus, STG = superior temporal gyrus, ACG = anterior cingulate gyrus, MFG = middle frontal gyrus, AG = angular gyrus, LOC = lateral occipital cortex, l = left, r = right.

### Inline Fig. S1

**N1 ERP amplitude analyses with extended group (n = 23).** Verification of N1 EEG analyses with an enlarged group of 23 subjects (Inline Table S1). In this group, 12 children trained FF set 1, and 11 children with FF set 2. a) The N1 amplitude analyses substantiate the results of the main text in showing the same character type and training effects. N1 potential field maps of the grand averages (first row, in μV) for each character type and the statistical t-maps (below), illustrating the differences between character types. Effect related to training: The LMM for the electrode clusters with fixed factor training (TFF, CFF) and cluster (LOT, MO, ROT) revealed a significant effect of training $[F(1,112)=9.64, p=0.0024, \hat{\eta}^2=0.086];$ TFF $< CFF, t(112)=-3.1$ but no effect of cluster $[F(2,112)=0.53, p=0.592]$. Effect related to character type: The LMM for the electrode clusters with fixed factor character type (LET, DIG, CFF) and cluster (LOT, MO, ROT) revealed a significant effect of character type $[F(2,174)=37.89, p<0.0001, \hat{\eta}^2=0.287]$ but no effect of cluster $[F(2,174)=2.06, p=0.1308]$. Contrasts between character types indicated significantly stronger N1 amplitudes to DIG as compared with LET $t(174)=4.35, p<0.0001$ and CFF $t(174)=8.69, p<0.0001$. b) The GFP for the N1 TFF correlated marginally with training measures (training duration: Pearson $r=-0.40, p=0.073$).
fMRI whole brain analysis of the core group (n = 18): Training effect. The whole brain analysis of the contrast TFF vs CFF showed a bilateral middle temporal gyrus (MTG) activation (MNI x = 49/-56, y = 12/-6, z = 15/-18, Inline supplementary table S6) which was more pronounced for trained as compared to control characters. Given that this activation did not survive the a priori defined cluster correction for multiple comparisons, this marginal effect shall be interpreted with care. Activation in the MTG has been related to integration of speech sounds and print in the process of reading acquisition and the bilateral activation could be explained by a diminished lateralization in language processing in young children (activation on axial and sagittal slices shown for p < 0.001 (uncorrected) cluster size left MTG k = 20, right MTG k = 17).

To verify the training effect, fMRI data of an enlarged group, including 24 children (Inline Table S1), was analysed. In this group, 12 children trained with FF set 1 and the other with FF set 2. This enlarged fMRI group consisted of data sets meeting data quality criteria for TFF and CFF. The additional analyses confirmed the results of the core group. (a) Correlation of BOLD responses on whole brain level to TFF in the left vOT ROI with children's training duration in the GPC training. Faster learning was correlated with significantly higher activation in the left vOT (small volume corrected p(FWEcorr) < 0.05). Below, the negative correlation of the significant vOT cluster activation is plotted against training duration for illustration purposes. (b) ROI analyses in the vOT. Top: The LMM with fixed factors hemisphere (l, r), and training (TFF, CFF) showed a significant main effect of training [F(1,70) = 8.43, p = 0.0049, f² = 0.1175]. When lowering our stringent motion quality inclusion criteria we were able to rerun the same LMM including for the character type effect (LET, DIG, CFF) and hemisphere (lvOT, rvOT) in the extended sample (n = 24), whereby one child did not conduct the DIG part of the experiment. This analysis showed a significant main effect of character type [F(2,107) = 6.02, p = 0.0033, f² = 0.1069] and no significant main effect of hemisphere [F(1,107) = 2.03, p = 0.1573]. Comparing character types revealed stronger activation for LET than CFF (t(107) = 2.72, p = 0.0209), and for DIG than CFF (t(107) = 3.2, p = 0.0051).
Despite differences between visual categories in the N1 ERP interval, previous studies also indicated some effects in the preceding posterior P1 (~100 ms) ERP. Category effects on the P1 have been demonstrated between words and objects (e.g. Schendan et al. (1998)) and also between words and symbol strings in adults (Maurer et al., 2005) though less consistently than in the visual N1. Based on previous findings on P1 differences we also examined the P1 interval in the current study. In analogy to the N1 analyses we used the global field power (GFP) maxima of the mean ERPs over all four character types to define the interval of the P1 as ±30 ms (102–162 ms) around the GFP peak. The same three posterior electrode clusters (LOT, MOT, ROT) were used for analyses.

The LMM analysis of the P1 global field power (GFP) mean amplitude (102–162 ms) with fixed factor training (TFF, CFF) revealed neither a significant training effect \[F(1,17) = 0.14, p = 0.7112\] nor a significant character type effect \[F(2,34) = 1.98, p = 0.1543\] and thus no global differences in the scalp potential related to graphophonological association training or character type. The LMM for the predefined posterior electrode clusters with fixed factor training and cluster (LOT, MO, ROT) revealed neither a training \[F(1,87) = 2.60, p = 0.1103\] nor a cluster effect \[F(2,87) = 1.09, p = 0.3395\]. The LMM for the predefined posterior electrode clusters with fixed factor character type and cluster (LOT, MO, ROT) showed a significant character type effect \[F(2,140) = 4.57, p = 0.0120, \eta^2 = 0.045827\] driven by a stronger P1 mean amplitude to LET compared to CFF [post-hoc t-tests: \(LET > CFF: t(140) = 3, p < 0.0089\)]. These results indicate some character type differences prior to the N1 in the posterior bilateral positivity P1 after around 130 ms. In contrast to the N1, no modulation of the P1 through phonological association learning was detected for false fonts.

To examine preferential activation to numbers in more detail, activation of a potential number sensitive area was analysed, by determining the beta values to all four character types in a literature based spherical ROI in the number form area (NFA; Abboud et al. (2015)) of the right hemisphere (MNI coordinates \(x, y, z = 55, -43, -20; r = 8\ mm)). This additional ROI analysis in the NFA did not reveal number sensitive activation differences in our kindergarten children [character type: \(F(3,50) = 0.99, p = 0.4834\)].
Maurer, U., Sprek, B., Kranz, F., Benz, R., Steinhausen, H.C., Brandeis, D., 2007. Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. Brain 130, 3200–3210.

Maurer, U., Sprek, B., Kranz, F., Benz, R., Steinhausen, H.C., Brandeis, D., 2006. Coarse neural tuning for print peaks when children learn to read. Neuroimage 33, 749–757.

Maurer, U., Bucher, K., Sprek, B., Benz, R., Kranz, F., Schulz, E., van der Mark, S., Steinhausen, H.C., Brandeis, D., 2009. Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. Biol. Psychiatry 66, 336–348.

Maurer, U., Bucher, K., Sprek, B., Brandeis, D., 2003. Altered responses to tone and phoneme mismatch in kindergartners at familial dyslexia risk. Neuroreport 14, 2245–2250.

Maurer, U., McCandliss, B., 2007. The Development of Visual Expertise for Words: The Contribution of Electrophysiology. Maurer, U., Sprek, B., Kranz, F., Benz, R., Steinhausen, H.C., Brandeis, D., 2011. The development of print tuning in children with dyslexia: evidence from longitudinal ERP data supported by fMRI. Neuroimage 57, 714–722.

Mayer, N. 2011. Test zur Erfassung der phonologischen Bewusstheit und der Benenngeschwindigkeit (TEPHOE). Ernst Reinhardt Verlag, München.

Mazaka, P.K., Glova, G.H., Reis, A.L., 2011. Rapid motions in pediatric and clinical populations. Psychiatry 65, 1315–1323.

McCandliss, B.D., Posner, M.I., Givon, T., 1997. Brain plasticity in learning visual words. Cogn. Psychol. 33, 88–110.

Miller, S.L., Wood, F.B., 1995. Electrophysiological indicators of black-white discrimination performance for letter and non-letter patterns. Int. J. Neurosci. 80, 299–316.

Molleson, P.J., Fisher, J.M., Denton, C.A., 2013. Adequate versus inadequate response to reading intervention: an event-related potentials assessment. Dev. Neuropsychol. 38, 281–296.

Olalude, O.A., Flowers, D.L., Napolioni, E.M., Eden, G.F., 2015. Dyslexic children lack word selectivity gradients in occipito-temporal and inferior frontal cortex. Neuroimage 7, 742–754.

Oswalik, J., Behrmann, M., 2008. Left hemisphere specialization for word reading potentially results, rather than causes from, a left lateralized bias for high spatial frequency visual information. Cortex 27, 29–39.

Park, J., Chiang, C., Brannon, E.M., Woldorff, M.G., 2014. Experience-dependent hemispheric specialization of letters and numbers is revealed in early visual processing. J. Cogn. Neurosci. 26, 2239–2249.

Park, J., van den Berg, B., Chiang, C., Woldorff, M.G., Brannon, E.M., 2018. Developmental trajectory of neural specialization for letter and number visual processing. Dev. Sci. 21, e1257A.

Parvainen, T., Helenius, P., Poskiparta, E., Niemi, P., Salmelin, R., 2006. Cortical outcome and multimodal brain signatures of letter-speech sound learning in dyslexic adolescents under minimized phonological demands. Neuroimage 74, 1126–1138.

Paulesu, E., Démonet, J.-F., Fazio, F., McCrosy, E., Chauvine, E., Brunswick, N., Capra, S., Conus, G., Habib, M., Frith, C.D., 2001. Dyslexic cultural diversity and biological unity. Science 291, 2165–2167.

Peğado, F., Comerolato, E., Ventura, F., Jobert, A., Nakamura, K., Bulatti, M., Ventura, P., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., 2014. Timing the impact of literacy on visual processing. Proc. Natl. Acad. Sci. United States Am. 111, E5233–E5242.

Pernet, C., Celsis, P., Démonet, J.-F., 2005. Selective response to letter categorization within the left fusiform gyrus. Neuroimage 28, 738–744.

Rispens-Trambauer, I., Brands, M., Woldorff, M.G., 2004. Hemispheric specialization for letter-speech sound learning during word perception. Neuropsychologia 42, 1232–1242.

Reckless, N., Brandeis, D., 2009. Neurophysiology in preschool improves contributions to reading. Trends Cognit. Sci. 13, 250–253.

The interactive account of ventral occipitotemporal contributions to reading. Trends Cognit. Sci. 15, 246–253.

Raschle, N.M., Chang, M., Gaab, N., 2011a. Structural brain alterations associated with dyslexia predate reading onset. Neuroimage 57, 742–749.

Raschle, N.M., Chang, M., Gaab, N., 2011b. Structural brain alterations associated with dyslexia predict reading onset. Neuroimage 57, 742–749.

Rey, A., Dufau, S., Massol, S., Grainger, J., 2009. Testing computational models of letter perception with item-level event-related potentials. Cogn. Neurosci. 26, 7–22.

Richlan, F., Kliegl, R., Wimmer, H., 2009. Functional abnormalities in the dyslexic brain: a quantitative meta-analysis of neuroimaging studies. Hum. Brain Mapp. 30, 289–306.

Richlan, F., Kliegl, R., Wimmer, H., 2011. Meta-analyzing brain dysfunctions in dyslexic children and adults. Dev. Neuropsychol. 36, 159–176.

Richlan, F., Kliegl, R., Wimmer, H., 2013. Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies. Hum. Brain Mapp. 34, 945–956.

Roth, A., Roesch-Ely, D., Bender, S., Weisbrod, M., Kaiser, S., 2007. Increased event-related potential latency and amplitude variability in schizophrenia detected through wave-based single-trial analysis. Int. J. Psychophysiol. 66, 244–254.

Price, C.J., Devlin, J.T., 2011. The interactive account of ventral occipitotemporal contributions to reading. Trends Cognit. Sci. 15, 246–253.
Saccchi, E., Lazzio, S., 2016. An event-related potential study of the relationship between N170 lateralization and phonological awareness in developing readers. Neuropsychologia 91, 415–425.

Sandak, R., Mencel, W.E., Frost, S.J., Rueckl, J.G., Katz, L., Moore, D.L., Mason, S.A., Fulbright, R.K., Constable, R.T., Pugh, K.R., 2004. The neurobiology of adaptive learning in reading: a contrast of different training conditions. Cognit. Affect. Behav. Neurosci. 4, 67–88.

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. Nat. Neurosci. 19, 1250–1255.

Schendan, H.E., Ganis, G., Kutas, M., 1998. Neuropsychological evidence for visual perceptual categorization of words and faces within 150 ms. Psychophysiology 35, 240–251.

Schlaggar, B.L., McCandliss, B.D., 2007. Development of neural systems for reading. Annu. Rev. Neurosci. 30, 475–503.

Selya, A.S., Rose, J.S., Dierker, L.C., Hedeker, D., Mermelstein, R.J., 2012. A practical guide to calculating Cohen’s f2, a measure of local effect size, from PROC MIXED. Front. Psychol. 3, 111.

Semrud-Clikeman, M., Goy, K., Griffin, J.D., Hynd, G.W., 2000. Rapid naming deficits in children and adolescents with reading disabilities and attention deficit hyperactivity disorder. Brain Lang. 74, 70–83.

Shaywitz, B.A., Shaywitz, S.E., Pugh, K.R., Mencel, W.E., Fulbright, R.K., Skudlarski, P., Constable, R.T., Marchione, K.E., Fletcher, J.M., Lyon, G.R., 2002. Disruption of posterior brain systems for reading in children with developmental dyslexia. Biol. Psychiatry 52, 101–110.

Shaywitz, B.A., Shaywitz, S.E., Pugh, K.R., Mencel, W.E., Fulbright, R.K., Skudlarski, P., Constable, R.T., Marchione, K.E., Fletcher, J.M., Lyon, G.R., Gore, J.C., 2003. Neural systems for compensation and persistence: young adult outcome of childhood reading disability. Biol. Psychiatry 54, 25–33.

Shum, J., Hermes, D., Foster, B.L., Dautjerji, M., Rangarajan, V., Winawer, J., Miller, K.J., Parviz, J., 2013. A brain area for visual numerals. J. Neurosci. : Off. J. Soc. Neurosci. 33, 6709–6715.

Somers, M., Aukes, M.F., Ophoff, R.A., Boks, M.P., Fleer, W., Kahn, R.S., Sommer, I.E., 2015. On the relationship between degree of hand-preference and degree of language lateralization. Brain Lang. 144, 10–15.

Song, Y., Huo, S., Li, X., Li, W., Liu, J., 2010. The role of top-down task context in learning to perceive objects. J. Neurosci. 30, 9869–9876.

Specht, K., Hugdahl, K., Olte, S., Nygaard, M., Bjernerud, A., Plante, E., Helland, T., 2009. Brain activation on pre-reading tasks reveals at-risk status for dyslexia in 6-year-old children. J. Cognit. Psychol. 50, 79–91.

Sthamam, K., Mandeville, J.B., Morocz, I.A., Sullivan, K.J., Livingston, M.S., 2002. Development of neural mechanisms for reading. Nat. Neurosci. 6, 767–773.

van Attovaert, N.M., Formisano, E., Goebel, R., Blomert, L., 2004. Integration of letters and speech sounds in the human brain. Neuroimage 47, 271–282.

van der Mark, S., Buijse, K., Maurer, U., Schulz, E., Brem, S., Buckelmuller, J. Kronbichler, M., Lonneker, T., Klaver, P., Martin, E., Brandeis, D., 2009. Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. Neuroimage 47, 1940–1949.

van der Mark, S., Klaver, P., Buijse, K., Maurer, U., Schulz, E., Brem, S., Martin, E., Brandeis, D., 2011. The left occipitotemporal system in reading: disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. Neuroimage 54, 2426–2436.

Vinkler, F., Dehaese, 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. Neuroimage 55, 143–156.

Vogel, A.C., Miezin, F.M., Petersen, S.E., Schlaggar, B.L., 2012. The putative visual word form area is functionally connected to the dorsal attention network. Cerebr. Cortex 22, 537–549.

Vogel, E.K., Luck, S.J., 2000. The visual N1 component as an index of a discrimination process. Psychophysiology 37, 190–203.

Whitfield-Gabrieli, S., Nieto-Castanon, A., 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connect. 2, 125–141.

Wild, N., Fleck, C., 2013. Neuromorphing des Mottier-Tests für 5-bis 17-jährige Kinder mit Deutsch als Erst-oder als Zweitsprache. Praxis Sprache 9, 1–12.

Wolf, M., Bally, H., Morris, R., 1986. Automaticity, Retrieval Processes, and Reading: A Electrophysiological Response Associated with Expertise in Letter Perception. Cognit. Affect Behav. Neurosci. 5, 306–318.

Yoncheva, Y.N., Wise, J., McCandliss, B., 2015. Hemispheric specialization for visual words is shaped by attention to sublexical units during initial learning. Brain Lang. 145, 23–29.

Zhao, J., Curran, T., 2001. A neural basis for expert object recognition. Psychol. Sci. 12, 43–47.

Xue, G., Poldrack, R.A., 2007. The neural substrates of visual perceptual learning of words: implications for the visual word form area hypothesis. J. Cognit. Neurosci. 19, 1643–1655.

Zaric, G., Correia, J.P., Binder, J.R., Possing, E.T., McKiernan, K.A., Ward, B.D., Hammeke, T.A., Eden, G.F., 2003. Development of the visual word form area hypothesis. J. Cognit. Neurosci. 19, 1643–1655.