Brain sensitivity to print emerges when children learn letter–speech sound correspondences

Silvia Brem[a,b,1], Silvia Bach[a,b], Karin Kucian[c], Tomi K. Guttorm[a,d], Ernst Martin[c,e], Heikki Lyytinen[a,d], Daniel Brandeis[b,e,f,2], and Ulla Richardson[a,2]

[a]Agora Center and [d]Department of Psychology and Child Research Centre, University of Jyväskylä, 40014 Jyväskylä, Finland; [b]Department of Child and Adolescent Psychiatry, University of Zürich, 8032 Zürich, Switzerland; [c]MR Center, University Children's Hospital, 8032 Zürich, Switzerland; [e]Center for Integrative Human Physiology, University of Zürich, 8057 Zürich, Switzerland; and [f]Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Mannheim J5 68072, Germany

Edited by Michael Posner, University of Oregon, Eugene, OR, and approved March 5, 2010 (received for review April 21, 2009)

The acquisition of reading skills is a major landmark process in a human's cognitive development. On the neural level, a new functional network develops during this time, as children typically learn to associate the well-known sounds of their spoken language with unfamiliar characters in alphabetic languages and finally access the meaning of written words, allowing for later reading. A critical component of the mature reading network located in the left occipito-temporal cortex, termed the "visual word-form system" (VWFS), exhibits print-sensitive activation in readers. When and how the sensitivity of the VWFS to print comes about remains an open question. In this study, we demonstrate the initiation of occipito-temporal cortex sensitivity to print using functional MRI (fMRI) (n = 16) and event-related potentials (ERP) (n = 32) in a controlled, longitudinal training study. Print sensitivity of fast (<250 ms) processes in posterior occipito-temporal brain regions accompanied basic associative learning of letter–speech sound correspondences in young (mean age 6.4 ± 0.08 y) nonreading kindergarten children, as shown by concordant ERP and fMRI results. The occipito-temporal print sensitivity thus is established during the earliest phase of reading acquisition in childhood, suggesting that a crucial part of the later reading network first adopts a role in mapping print and sound.

learning | training | visual word-form system | functional MRI | event-related potential

Learning to read starts with the establishment of grapheme–phoneme correspondences between letters and speech sounds in alphabetic languages such as German or English. At the same time, a new functional network emerges at the neural level. Because reading is a recent cultural invention, it is unlikely that brain structures have evolved exclusively for reading. It is more plausible that some visual-processing units (1) adopt additional functions and, through practice, are increasingly sensitized to print processing in the course of childhood. The brain's initial sensitization to print in children is the focus of the present article.

Left occipito-temporal cortex regions, referred to as the "visual word-form system" (VWFS) (2), are often engaged in print processing. The activity of the VWFS, with its posterior-to-anterior sensitivity gradient to word-like stimuli (3–5), is crucial for fluent reading (2, 6, 7), particularly in beginning readers (8, 9). The term "reading skill zone" for the VWFS (10, 11) accurately emphasizes its continuous tuning with childhood development (12–14) and reading practice and its diminished activation in poor readers (15, 16).

The activation of the VWFS measured with functional MRI (fMRI) has been associated with an event-related potential (ERP) between 150 and 250 ms characterized by a focal negativity over the occipito-temporal scalp and termed "N1." The N1 is sensitive to print, with larger amplitudes over the left hemisphere for alphabetic than for nonalphabetic stimuli (17, 18). The greater sensitivity of the N1 to print than to nonsense symbols develops and peaks in the first 2 years after school enrollment, when children learn to read (9, 13, 19). During this early phase, phonological word processing is crucial, according to the dual-route model of reading (20), because direct lexical mapping of orthographic information calls for an orthographic lexicon which has yet to be built. The contribution of the VWFS in this initial phase may enable indirect lexical access through grapheme–phoneme mapping or phonological decoding as suggested by the phonological mapping hypothesis (21). This reliance on phonological processes also would explain why VWFS activity and the corresponding N1 print sensitivity (3, 13, 18, 22) is most pronounced in beginning readers relying on phonological decoding (5, 9) but is reduced in individuals with decoding problems (23) such as children with dyslexia, a severe developmental reading disorder (8, 15, 16).

Despite the importance of the VWFS for fluent reading, it still is not known at which point sensitivity to print emerges during reading acquisition. ERP evidence (13) suggests that mere visual familiarity to print is not enough for VWFS sensitization. However, it is unclear whether learning initial letter–speech sound correspondence in one's first (native) language sensitizes the brain regions that later form the VWFS to print or whether such tuning follows in-depth visual orthographic knowledge and word recognition.

This question cannot be answered by tracking training-related changes in adults learning a second script (24, 25) (because the VWFS already has been shaped by the experience with the native script) or by comparing children before and after they learn to read (9) (because literate children may rely in part on direct whole-word recognition).

We examined the emergence of VWFS print sensitivity in nonreading, (Swiss) German-speaking kindergarten children (Table 1) in a controlled, longitudinal cross-over training study with ERP (n = 32) and fMRI (n = 16) assessment at three time points (T1, T2, and T3). The two imaging techniques are suitable for examining young children and provide complementary insights into spatial and temporal aspects of brain function. Accurate localization of print-sensitive brain regions with fMRI is complemented by the precise timing from ERPs, which distinguish, for example, fast automatic from slower secondary print sensitivity. All children practiced with both a computerized grapheme–phoneme correspondence game (the Graphogame, GG) (26) and a nonlinguistic number-knowledge control game (control, NC), each for an 8-week period (total ~3.6 h). One group of children practiced first with the GG; the second group started with the NC (Table 1, *SI Text*, and Fig. S1). After the second imaging assessment (T2), the groups switched games to coun-

Author contributions: S. Brem, T.K.G., E.M., H.L., D.B., and U.R. designed research; S. Brem, S. Bach, and K.K. performed research; S. Brem and S. Bach analyzed data; and S. Brem, D.B., and U.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: sbrem@kjpd.uzh.ch.

[2]D.B. and U.R. contributed equally to this work.

This article contains supporting information online at www.pnas.org/cgi/content/full/0904402107/DCSupplemental.

Table 1. Characteristics of training groups

	ERP group				fMRI group			
	Mean ± SD				Mean ± SD			
Characteristic	All	GG-first	NC-first	*P value*	All	GG-first	NC-first	*P* value
N	32	15	17		16	8	8	
Sex (f)	17	8	9	0.98*	13	5	8	0.055*
Risk	14	8	6	0.31*	5	2	3	0.59*
Age (y)	6.4 ± 0.3	6.5 ± 0.3	6.3 ± 0.3	0.18	6.4 ± 0.3	6.5 ± 0.4	6.4 ± 0.3	0.55
IQ	114 ± 16	110 ± 12	117 ± 18	0.25	117 ± 13	113 ± 10	120 ± 16	0.35
Education (y)	16.7 ± 3.1	17 ± 2.6	16.4 ± 3.6	0.55	16.1 ± 2.9	16.4 ± 2.3	15.8 ± 3.6	0.71
UC	10 ± 7.3	11.2 ± 5.8	8.9 ± 8.4	0.39	12.1 ± 7.6	12.8 ± 6.7	11.4 ± 8.8	0.73
LC	5.8 ± 6.1	5.2 ± 5.2	6.3 ± 7.0	0.62	6.8 ± 6.1	5.6 ± 4.0	7.9 ± 7.7	0.48
Reading	0.4 ± 1.2	0.2 ± 0.4	0.7 ± 1.5	0.26	0.5 ± 1.1	0.3 ± 0.5	0.8 ± 1.5	0.39
Phonology	35.3 ± 3.6	35.8 ± 4.2	34.9 ± 3.0	0.48	36.1 ± 2.9	36.9 ± 3.0	35.4 ± 2.8	0.31
Vocabulary	16.3 ± 3.3	16.3 ± 3.7	16.3 ± 3.1	0.98	16.4 ± 3.1	16.6 ± 3.6	16.1 ± 2.9	0.76
ARHQ	0.29 ± 0.10	0.28 ± 0.12	0.30 ± 0.08	0.42	0.30 ± 0.11	0.27 ± 0.14	0.34 ± 0.06	0.25

ARHQ, mean score of adult reading history questionnaire; Education, parents' years of education; F, female; LC, knowledge of lowercase letters; Phonology, phonological awareness (sum score of four subtests of the Bielefelder screening test); Reading, number of correctly read words; Risk, children with familial risk for dyslexia; UC, knowledge of uppercase letters.
* χ^2 tests.

terbalance the order of the games so that individual training effects could be compared. Sensitivity to print processing was assessed with ERP and fMRI at all three test times using a simple modality judgment task (Fig. 1) including visual, auditory, and audiovisual word (W)/speech and false font (FF)/nonintelligible speech stimuli.

Results

Behavior. ***Training intensity.*** The total training time (minutes) and the training period (days) spent with each game did not differ between GG and NC training (Table S1) as shown by ANOVAs with the between-subject factor *group* [children starting with GG (GG-first) or with NC (NC-first) training] and the within-subject factor *training game* (GG, NC). The children did, however, practice longer with the first training game than with the second training game [$F(1,29) = 8.48$, $P = 0.007$, $\eta_p^2 = 0.23$]. This preference can be explained by some loss of motivation and interest over time because of the (intended) similarity of the games.

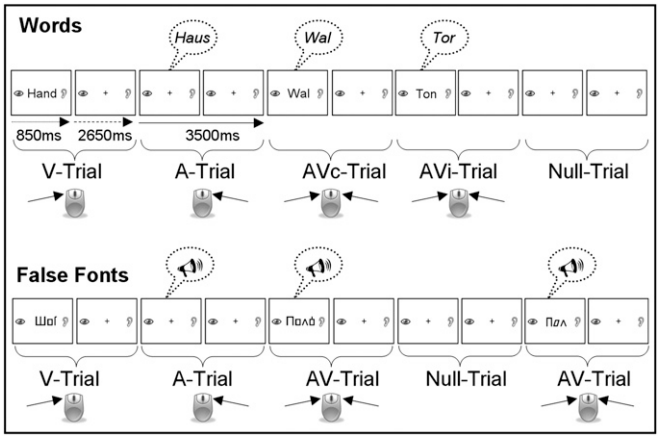

Fig. 1. Implicit audiovisual word and false font/rotated speech-processing task, divided into two separate parts: unimodal and bimodal word (*Upper*) and false font/rotated speech presentation (*Lower*). Children had to indicate whether a stimulus was presented visually (V), auditorily (A), or audiovisually (AVc/AVi), by pressing response buttons.

Letter Knowledge and Reading. Letter knowledge did not differ between groups at the beginning of the study (Table 1). A multivariate repeated-measures ANOVA (MANOVA) showed that children named or sounded out significantly more uppercase than lowercase letters [$F(1,30) = 145.3$, $P < 0.001$, $\eta_p^2 = 0.83$] and that their naming performance increased with *test time* [$F(2,29) = 51.71$, $P < 0.001$, $\eta_p^2 = 0.78$]. As expected, the improvements in letter knowledge were greater in association with GG practice than with NC practice [$F(2,29) = 13.25$, $P < 0.001$, $\eta_p^2 = 0.48$] (Fig. S2). Across all *test times*, the GG-first group knew more letters than the NC-first group [$F(1,30) = 5.1$, $P = 0.032$, $\eta_p^2 = 0.15$] because, having received GG training first, they already had improved their letter knowledge by T2.

Despite the gains in letter knowledge, reading skills improved only slightly with *time* [$F(2,29) = 4.16$, $P = 0.026$, $\eta_p^2 = 0.22$]. For all children, reading skills remained very rudimentary even after GG training (only three children were able to "decode" more than 10 words), as expected from the design and the main aim of the GG—namely, to teach grapheme–phoneme correspondences.

Task Performance ERP/fMRI. The MANOVA on changes in responding to W and FF stimuli over *time* (T1, T2, T3) (sensitivity index d'; Table S2) in the two groups revealed only a time main effect [$F(2,29) = 4.96$, $P = 0.014$, $\eta_p^2 = 0.26$] in ERP recordings. Neither *condition* nor *group* effects were apparent in reaction-time data for the ERP, or fMRI assessments.

fMRI. For second-level, whole-brain analyses, fMRI groups were pooled as follows: Pre-GG corresponded to the last test before GG training (T1 for the GG-first group and T2 for the NC-first group), and Post-GG summarized data immediately after GG training (T2 for the GG-first group and T3 for the NC-first group). Pre-NC and Post-NC groups were pooled similarly. A bilateral visual network was activated (Fig. 2*A*) according to whether the children processed W or FF stimuli (Table 2 and Table S3). In the whole-brain analyses, the contrast between W and FF stimuli did not differ in the Pre-GG and Pre-NC groups ($P < 0.005$, $k \geq 15$). Fig. S3 shows differences between the fMRI subgroups at T1.

Training effects determined by ANOVA ($P < 0.001$, $k \geq 15$) with factors *time* (Pre, Post) and *training* (GG, NC) for the W–FF contrast were in line with our hypothesis: A more pronounced

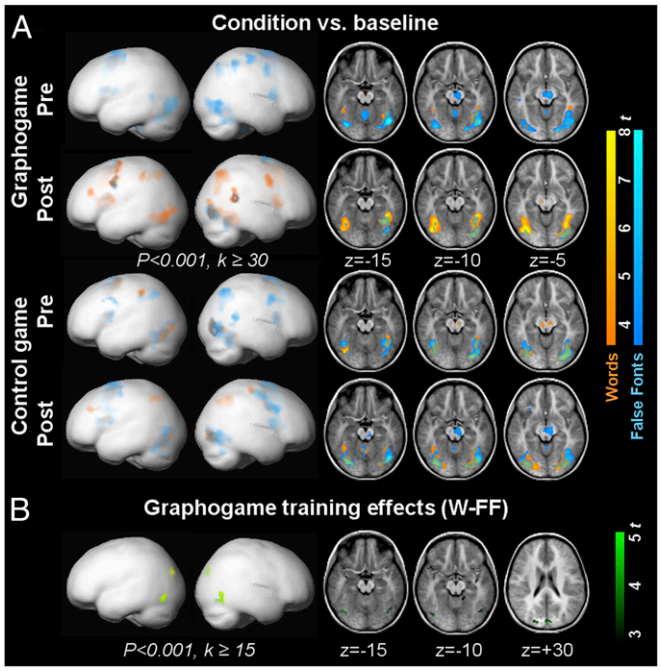

Fig. 2. Emerging print sensitivity seen as differential fMRI activation to words and false fonts and projected onto a pediatric brain template (*Left*) and on three sections ($k \geq 0$) of the mean structural image of the group (*Right*). (*A*) fMRI activation to words (orange) and false fonts (blue) for the whole fMRI sample ($n = 16$) before (Pre) and after (Post) GG training (*Upper two rows*) or NC training (*Lower two rows*). For both conditions, Pre- and Post-training activity was predominantly bilateral occipito-temporal. Note that no difference was detected for words vs. false font stimuli between Pre-GG and Pre-NC, but slight group differences in the right inferior occipital gyrus were found between the fMRI subgroups at T1 (Fig. S3). (*B*) Interaction (two-factorial ANOVA) of *training* (GG, NC) and *time* (Pre, Post) revealed areas with more pronounced activity for the word–false font contrast after GG but not after NC training (green). On the lateral views (*Left*), Post vs. Pre GG training of W-FF is superimposed in yellow.

increase in print sensitivity was detected in two clusters in the left and right occipito-temporal lobes (fusiform gyrus and inferior temporal gyrus) as well as in the cuneus over the course of GG (Pre-Post) than during the course of NC training. The changes in print sensitivity (Post vs. Pre) for each game separately confirmed this result by pointing to GG-related increases in activation in the left posterior fusiform gyrus, right inferior temporal gyrus, and cuneus (Fig. 2*B* and Table 2), whereas no effect was seen over the course of NC.

Table 2. Training effects on brain activity (activity greater for the contrast W-FF after than before training)

		MNI coordinate				
Training	Brain area	*x*	*y*	*z*	*Z*	*k*
GG	L FFG*	−45	−78	−12	4.68	26
	L Cuneus	−21	−90	30	3.93	29
	R ITG, FFG	57	−69	−3	3.88	29
	R Cuneus	3	−90	33	3.73	52
NC	No significant voxel at $P < 0.001$, $k \geq 15$					
GG > NC	L FFG*	−45	−75	−15	4.73	19
	Cuneus	6	−93	30	4.14	34
	R ITG, FFG	57	−72	−3	4.08	33
NC > GG	No significant voxel at $P < 0.001$, $k \geq 15$					

Listed are MNI coordinates (*x*, *y*, *z*) of cluster maxima for $P < 0.001$, $k \geq 15$ (uncorrected). FF, false fonts; FFG, fusiform gyrus, ITG, inferior temporal gyrus; L, left; R, right, W, words.
*Maxima of clusters survive family-wise error correction at $P < 0.05$.

Region-of-Interest Analysis. To elucidate further the effects of training on print-sensitive processing, we performed a region-of-interest (ROI) analysis of the percent signal change across five spherical ROIs (R1–R5) along the anterior–posterior axes of the occipito-temporal cortex (Fig. 3). Notably, at T1 group differences in the W–FF contrast were already present in three posterior ROIs [R3: left (l), $P = 0.069$, right (r), $P = 0.072$; R4: l, $P = 0.016$, r, $P = 0.005$, R5: l, $P = 0.030$, r, $P = 0.078$; Fig. S4]. The main MANOVA included the factors *group* (GG-first, NC-first), *test time* (T1, T2, T3), *hemisphere*, *condition*, and *ROI* (R1-5). Because of the aforementioned between-group *condition* differences at T1, we verified the results with posthoc MANCOVAs.

No main effect or interaction with *hemisphere* was found, but more anterior ROIs showed less activation than posterior ROIs [$F(4,11) = 30.64$, $P < 0.001$, $\eta_p^2 = 0.92$]. This regional difference in activation was modulated by *condition* [$F(4,11) = 3.87$, $P = 0.034$, $\eta_p^2 = 0.58$]. The *ROI* × *condition* × *group* interaction [$F(4,11) = 4.5$, $P = 0.021$, $\eta_p^2 = 0.62$] further indicated *group* differences in regional W–FF differentiation. Finally, the *ROI* × *test time* × *condition* × *group* interaction [$F(8,7) = 3.82$, $P = 0.047$, $\eta_p^2 = 0.81$] indicated that this difference in activation between W stimuli and FF stimuli also depended on time.

Based on these interactions, we assessed the training effects in separate MANOVAs for each ROI. The expected training effect emerged only in R4 [$F(2,13) = 5.67$, $P = 0.017$, $\eta_p^2 = 0.47$], and separate analyses of the left and right R4 revealed that the training effect was confined to the left hemisphere [$F(2,13) = 5.86$, $P = 0.015$, $\eta_p^2 = 0.47$, right: $P > 0.1$]. Posthoc MANCOVAs including the baseline covariates confirmed training effects within the occipito-temporal cortex [*condition* × *test time* × *group*: $F(2,10) = 6.12$, $P = 0.018$, $\eta_p^2 = 0.55$] (Fig. S5) as well as the regional training effect in the left R4 [$F(2,12) =$, $P = 0.028$, $\eta_p^2 = 0.45$; right: $P > 0.5$]. The training effect mainly reflects the growing response to W stimuli during GG and not during NC training, consistent with post hoc *t* tests showing that changes in print sensitivity occurred predominantly when children played with GG (Fig. 3 and Fig. S4).

ERPs. ***Amplitude analyses.*** Our method for analyzing the N1 for training-related increases of print-sensitive activity was similar to

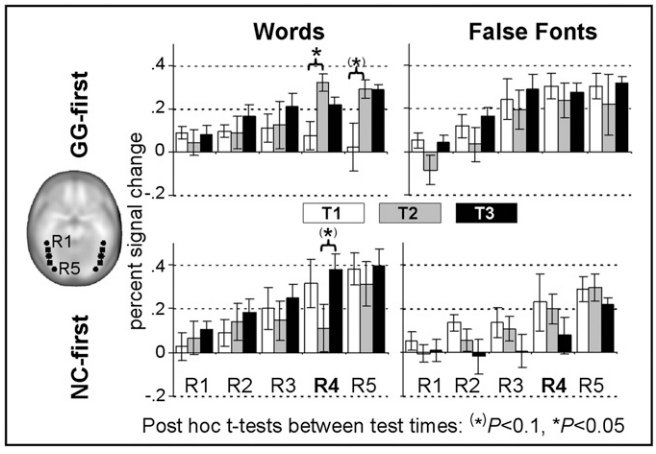

Fig. 3. ROI analyses along the occipito-temporal cortex. Mean percent signal change (error bars, ±1 SEM) for the GG-first (*Upper*) and NC-first (*Lower*) *groups* in five consecutive spherical ROIs plotted at each *test time* for the left hemisphere (Fig. S4). The location of the five ROIs is illustrated on the axial slice at the left. Significant training effects (*) were detected in R4.

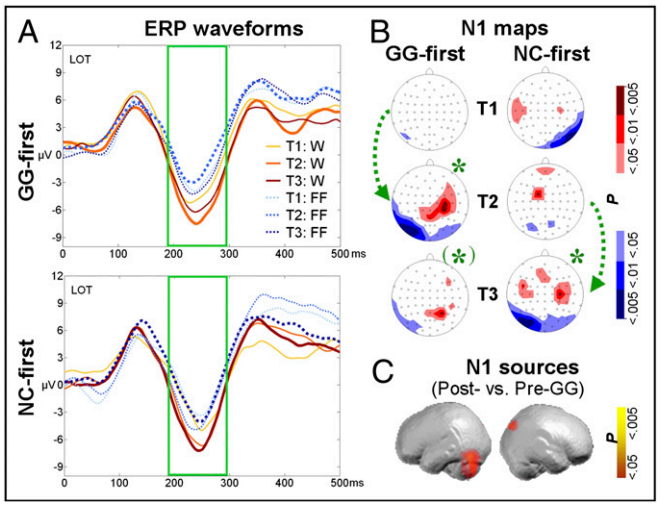

Fig. 4. Print-sensitive ERP activity in the visual N1 after GG training. (*A*) Training effects on the N1 (189–295 ms) at LOT sites for the GG-first (*Upper*) and NC-first (*Lower*) groups, respectively. Waveforms after GG training are plotted with thicker lines. Print sensitivity emerged as a pronounced difference in amplitude between W and FF in the N1 interval after GG training. (*B*) Statistical N1 ERP t-maps illustrating the W–FF contrast for the GG-first (*Left*) and NC-first (*Right*) groups. According to TANOVA, W and FF map topographies differed after GG training at T2 for the GG-first group and at T3 for both groups. *, $P < .05$; (*), $P < 0.1$). Green arrows illustrate periods with GG training. (*C*) Estimated N1 Post- vs. Pre-GG sources for W–FF in the occipito-temporal cortex and right cuneus.

our analysis of fMRI ROI data. Mean ERP amplitudes in left occipito-temporal (LOT) and right occipito-temporal (ROT) electrode clusters were compared over an interval of 195–289 ms following W and FF stimuli. No group difference in print sensitivity occurred at T1 (LOT $P > 0.8$, ROT $P = 0.095$). The MANOVA for N1 amplitudes with factors *group*, *test time, condition*, and *hemisphere* pointed to significant differences in *condition* [$F(1,30) = 22.39$, $P < 0.001$, $\eta_p^2 = 0.43$] and *test time* [$F(2,29) = 3.38$, $P = 0.048$, $\eta_p^2 = 0.19$]. No main effect or interaction with *hemisphere* reached significance. Most importantly, and in accordance with the fMRI results, the difference between W and FF responses emerged only after GG training [$F(2,29) = 3.56$, $P = 0.042$, $\eta_p^2 = 0.2$] (Fig. 4). Posthoc *t* tests on the differential N1 amplitudes at LOT electrode clusters corroborated that the MANOVA reflected a significant increase in the W–FF differentiation only after GG (pooled groups: Post-GG vs. Pre-GG (W–FF), $P = 0.016$; trends for subgroups: GG-first T1–T2, $P = 0.086$; NC-first T2–T3, $P = 0.084$). The print sensitivity effect, however, declined again after discontinuation of grapheme–phoneme training, as shown in the GG-first group between T2 and T3 ($P = 0.022$), suggesting that continued practice is required to consolidate the sensitivity gained with GG training.

Further *t* tests showed no significant amplitude differences between W and FF stimuli in the LOT before GG training, but response amplitudes to W stimuli were larger at T1 (but not at T2) for the NC-first group in the right hemisphere (T1: $P = 0.003$).

Topographic analyses. The overall topographic analysis of variance (TANOVA), which detects pure topographic differences in ERP maps, concurred with the N1 amplitude analyses. Accordingly, differences in map topography between W and FF stimuli (Fig. 4*B*) were found at T2 ($P = 0.017$) only for the GG-first group and were found at T3 for both groups (NC-first: $P = 0.041$; GG-first: trend $P = 0.052$).

ERP sources. The ERPs also allow estimation of source locations. In analogy to the whole-brain fMRI analyses, we tested whether the source of print-sensitive N1 activity is located within the VWFS by estimating the training-related sources (Fig. 4*C*) and comparing pre- and posttraining data of the print-sensitive ERP N1 activity, pooled over both groups. The sources of significant ($P < 0.05$) print-sensitive N1 activity after GG training were localized to the left occipito-temporal cortex (fusiform gyrus and lingual gyrus), right cuneus, and posterior cingulate, concurring nicely with the fMRI results.

Discussion

Here, we demonstrated that the learning of letter–speech sound correspondences in young, nonreading kindergarten children results in an initial sensitization to print of specific areas within the occipito-temporal cortex. Emerging print sensitivity was assessed by comparing the processing of linguistically relevant print (W) and FF stimuli in the purely visual conditions before and after grapheme-phoneme training.

Words and false fonts activated a bilateral and predominantly ventral posterior occipito-temporal network before training, as previously found for symbol-string and object processing in both adults (3, 4, 27, 28) and children (5, 14). After the brief (3–4 h) grapheme–phoneme training in linking print to speech sounds, print-sensitive activation was enhanced mainly in the posterior VWFS, with corresponding fMRI results and N1 ERP effects at around 200 ms. This finding indicates that specific brain regions in the emerging VWFS are tuned for print and adopt print-specific functions when phonological mapping of graphemes becomes feasible. The emergence of print sensitivity in cortical areas during the acquisition of grapheme–phoneme correspondences is in line with the inverse U-shaped developmental trajectory of print sensitivity of the ERP N1, which peaks in beginning readers (9), i.e., a phase in which models of reading acquisition emphasize the importance of grapheme–phoneme decoding (29).

Increasing visual familiarity with letters goes hand in hand with grapheme–phoneme expertise and might contribute to the coemergent print sensitivity. In the extreme, a more pronounced sensitivity effect would be expected in adults, who have greater letter familiarity, than in children, but the N1 and the VWFS sensitivity to alphabetic over nonalphabetic stimuli typically becomes less pronounced in skilled adult readers (3). Maurer et al. (13) have shown that familiarity with letters of kindergarten children is not sufficient for left-lateralized print sensitivity but may account for some precursors reflected by a posterior N1 negativity atypically confined to the right hemisphere, such as the topography seen in the NC-first group at T1. Accordingly, the children in our study exhibited considerable print familiarity and letter-naming abilities before starting the training (at T1) but did not exhibit the significant VWFS or N1 responses that would be expected if visual familiarity alone were responsible for these markers of print sensitivity. Prior fMRI data also favor phonological mapping over visual familiarity: Visual word-form training with corresponding increases in visual familiarity and expertise resulted in diminished fusiform activation in adults (24, 25), whereas enhanced fusiform activation was found after phonological training (24). Taken together, visual familiarity to letters might explain the lesser activity in R4 in response to W as compared with FF stimuli before grapheme–phoneme training but cannot explain the increase in activity in response to W stimuli seen after training. On the other hand, developing visual expertise with case- and font- invariant print features may allow effective grapheme–phoneme mapping and contribute indirectly to the N1 enhancement and to emerging fusiform sensitivity (6).

We cannot rule out word-recognition processes as contributors to the initial print sensitivity after training. A major role for word recognition seems unlikely, because even the few children who started to practice word reading during the training exhibited only very rudimentary reading skills. The children's effortful letter-by-letter decoding in the reading test clearly reflected slow, indirect word processing (20). The print sensitivity we observed, on the other hand, reflected fast and implicit processing, because

the differentiation occurred rapidly (~250 ms) after stimulus presentation as shown by the time-sensitive ERPs and the correspondence of the estimated N1 source with fMRI activity.

In the present study, print-sensitive activation was located posterior to the core of the adult VWFS (2), a region associated with processing nonsense consonant strings in adults (4). A shift of print-sensitive activation to more anterior regions that become selectively tuned for word processing (6) would corroborate the change of VWFS functionality in the process of learning to read when children's print-processing strategies change from letter-by-letter decoding to fluent whole-word reading. Future studies may clarify whether such an anterior shift occurs during reading development.

Limitations of the present study include the unbalanced gender distribution of the fMRI subgroup, the variability in children's training times, the inclusion of children at familial risk for developmental dyslexia, and group differences in fMRI print sensitivity at outset (T1). Previous studies have not reported consistent gender differences in the laterality of activation patterns (30), suggesting that the emerging print sensitivity was not affected by the gender imbalance among our subjects. Supplementary analyses (*SI Text*) confirmed our core training effects when taking into account variable training intervals or the familial-risk status of the children. Additional analyses controlling for pretraining group differences confirmed the training effect, particularly in the left hemisphere. The young age of the children forced us to use short tasks to maintain children's attention and achieve fMRI (and ERP) data of acceptable quality but limited the statistical power. The striking convergence of complementary brain imaging at high temporal and high spatial resolution regarding the emerging VWFS sensitivity to print argues for the validity of our results.

In summary, print sensitivity in the VWFS emerges rapidly during acquisition of grapheme–phoneme correspondences in young children before full-word reading. The onset of print sensitivity of the VWFS after grapheme–phoneme training supports its central role in initial phonological decoding as predicted by the phonological mapping hypotheses (21). Because phonological decoding is a known core problem for individuals with dyslexia, a critical role of the VWFS in storing grapheme–phoneme correspondences also could explain its reduced activity in dyslexia. Whether strengthening the VWFS sensitivity to print through grapheme–phoneme training may alleviate potential reading problems in children with developmental dyslexia remains to be examined.

Methods

Further methodological details are given in *SI Methods*.

Subjects. Thirty-two young, healthy, non reading, right-handed children in kindergarten completed training as well as ERP and behavioral testing (Table 1). Sixteen of these children also completed the fMRI assessments (fMRI subgroup). To identify at-risk children, both parents reported in a questionnaire whether any first-degree relative of the child was dyslexic; no formal diagnosis was required. In addition, parents completed an adult reading history questionnaire (ARHQ) (31). At-risk and no-risk children were pooled for all analyses because of the small number in each group; mean parental ARHQ scores were used as a covariate for supplementary analyses (*SI Text*) to control for individual familial risk status. The years of parental education served as an estimate of the children's socioeconomic background.

Behavioral Screening Battery. The behavioral screening test battery (Table 1) administered before training started covered IQ [with Raven's colored matrices (32)], phonological skills (Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten, BISC) (33), vocabulary (receptive vocabulary), and word comprehension (Marburger Sprachverständnistest für Kinder, MSVT) (34). Letter knowledge (producing letter names or letter sounds) and reading skills were tested repeatedly using the word-reading subtest of the Salzburger Lesetest, SLT) (35) to track training-related improvements. Only children who could "read" (usually several attempts and much time were necessary for decoding) fewer than 6/30 high-frequency nouns at the first behavioral test were included in the study; these children are referred to as "nonreaders."

Study Design. Each of the three test times (T1, T2, and T3) included a behavioral assessment plus separate ERP and fMRI sessions in a counterbalanced order (nine children started with fMRI; seven children started with ERP). Between assessments, the children trained at home (Table S1) with either the noncommercial GG or the NC (average training period: GG, 58.1 ± 12.5 days; NC, 56.9 ± 14.1 days). GG is a computerized, child-friendly, grapheme–phoneme association training program developed for nonreading children (*SI Text*) (26). In the NC, number knowledge and basic addition and subtraction were trained. Both games were matched in terms of motivational aspects, game type, and visual appearance (Fig. S1). Children were assigned randomly to matched groups starting with either GG (the GG-first group) or NC (the NC-first group) (Table 1).

ERP and fMRI Stimulation and Task. We used a modality judgment task (Fig. 1) consisting of two experimental parts to examine print processing in nonreading children. In one part, visual (V), auditory (A), audiovisually congruent (AVc), and audiovisually incongruent (AVi) word (W) processing was assessed. The other part examined responses to false fonts (FF) and non intelligible speech. Children judged the modality by pressing the corresponding button(s) using their left, right, or both index fingers.

Stimuli in black on white appeared for 850 ms every 2650 ms; between stimuli static pictures of an eye on one side of the screen and an ear on the other reminded children where to press. The ERP task included the pseudorandom presentation of 42 stimuli per condition and 42 null events. To reduce movement-related artifacts, the fMRI task was shorter and consisted of 28 stimuli per condition and 40 null events plus an initial and final rest block of 10.5 s.

fMRI and ERP Task Performance. The behavioral data (Table S2) served only to monitor children's attention during task performance; no child was excluded because of poor performance. We also confirmed that the results were not biased by lack of attention and poor performance as shown by a supplementary analysis (*SI Text*) that included only the best-performing children (sensitivity index $d' \geq 1$).

fMRI Recording and Processing. FMRI was performed on a 3-T scanner (GE Medical Systems) with an EPI sequence covering the whole brain (25 axial slices, TR = 1.5 s, TE = 31 ms, matrix = 64 × 64, slice thickness/gap = 4.6/0.4 mm, flip angle = 50°, FOV = 240 × 240 mm). Visual stimuli were presented through video goggles, and auditory stimuli were presented via headphones. Responses were collected with a response box.

A high-resolution structural data set also was acquired, using a 3D T1-weighted gradient echo sequence (172 slices, TR = 6.31 ms, TE = 2.93 ms, flip angle = 12°, voxel size 0.94 × 0.94 × 1).

Data were analyzed using SPM5 software (Wellcome Department of Cognitive Neurology, http://www.fil.ion.ucl.ac.uk/spm) in the following order: functional data first were motion corrected and then were coregistered to the child's individual anatomical data. Normalization of the individual anatomical image to a pediatric (5–9.5 y) template provided by the Cincinnati Children's Hospital Medical Center (https://irc.cchmc.org) was followed by normalization (fourth-degree spline interpolation) of the functional images with the estimated parameters, resampling to isotropic 3-mm^3 voxels, and spatial smoothing with a Gaussian kernel of 9-mm FWHM. No individual run had more than 2.5 mm/degree translation/rotation displacement in the x, y, or z plane. The subject-specific first-level model included both experimental parts (W stimuli and FF stimuli). The event-related activation evoked by all conditions (V, A, AVc, and AVi) was modeled using the standard SPM hemodynamic response function and filtered with a 128-s high-pass filter. Realignment parameters were included in the model to account for motion.

fMRI Statistics. fMRI whole-brain analyses were used to determine GG and NC *condition* and *training* effects by collapsing over fMRI subgroups (*Results*). Random-effect t tests based on the individual contrast images were used to compare *conditions*, *groups* and/or *test times*. For illustration, statistical parametric maps of t values for *condition* vs. baseline contrasts of the pooled groups were thresholded at $P < 0.001$, $k \geq 30$, uncorrected (Fig. 2*A*, Table S3 for fMRI subgroups, and Fig. S3*A*). An ANOVA with the factors *training type* (GG, NC) and *time* (pre- and posttraining) was computed for the W–FF contrast to determine which brain areas showed an increase in print-sensitive and GG- vs. NC-related activity from pre- to posttraining ($P < 0.001$, $k \geq 15$, uncorrected, Fig. 2*B*).

ROI analyses were computed for the percent signal change on smoothed data in five spherical ROIs (radius = 6 mm) placed contiguously along the occipito-temporal cortex of the left and right hemispheres (Montreal Neurological Institute coordinates *x*, *y*, *z*: R1: ± 52, −42, −18; R2: ±50, −54, −16; R3: ±48, −66, −14; R4: ±46, −78, −12; R5: ±38, −90, −10) (*SI Text*). Percent signal change was extracted (MARSBAR V0.41 toolbox, provided by M. Brett; http://marsbar.sourceforge.net) from these ROIs for each condition and used for MANOVA analysis and posthoc *t* tests. Because of group differences in the W–FF contrast at T1 in the three posterior ROIs (3–5), baseline measures (T1 means over both hemispheres for each of the three ROIs) were covaried in posthoc MANCOVAs to verify our core analysis. Correspondingly, the training effect in the left R4 ROI also was checked using the *condition* difference of the left R4 at T1 as the covariate.

ERP Recording and Processing. Participants were seated in front of a computer display, and sounds were presented through headphones. The ERPs were recorded from 64 channels (500 Hz, filters 0.1–70 Hz; *SI Text*). After downsampling (256 Hz), the ocular artifacts were removed using an independent component analysis from the 0.1–30 Hz filtered data. Before averaging, the corrected files were bandpass filtered digitally (0.3–30 Hz, 24 dB), and epoched (−125 ms to 1125 ms). Artifacts exceeding ±100 μV (for five children, ±125 μV) in any channel were rejected. ERPs were transformed to the average reference for all subsequent analyses. For each group and test time, separate W and FF averages were computed.

The N1 was determined by means of the interval between two subsequent global field power (GFP) sinks (195–289 ms) in the waveform, defined by averaging the group grand means for each test time and each condition.

ERP Statistics. For the N1 amplitude analyses, the individual mean amplitude for both *conditions* and all *test times* in the N1 interval for the average of LOT (PPO9h, O1′, OI1, PO9) and ROT (PPO10h, O2′, OI2, PO10) electrodes was entered in a MANOVA with the between-subject factor *group* (GG-first, NC-first) and within-subject factors *hemisphere*, *condition*, and *test time* (T1, T2, T3).

Topographic (map) analyses. Training-induced differences in N1 map topography were investigated using TANOVA on normalized (scaled to unity) GFP maps, which detects only purely topographic differences not explained by overall amplitude (GFP) differences. Within the N1 interval, the topographic difference in *condition* (between W stimuli vs. FF stimuli) was determined for both *groups* and all three *test times* (Fig. 4*B*).

N1 source computations. Analogous to the whole-brain fMRI analyses, we computed the statistical difference between Post- vs. Pre-GG training for the print-sensitive N1 sources by collapsing over all 32 children. We used the local autoregressive average (LAURA) model (36) (http://brainmapping.unige.ch/Cartool.htm), a distributed linear inverse source estimation (*SI Text*). For each subject, individual sources relating to W and FF stimuli were first estimated. Print-sensitive pre- and posttraining sources were determined as the difference between the W and FF sources at the relevant time interval. Paired *t* tests were performed on a node-by-node basis to compare the mean Pre- and Post-GG training print-sensitive LAURA-estimated activities.

ACKNOWLEDGMENTS. We thank Dr. Paul Summers for language editing and support. This study was supported by the European Commission FP6, Marie Curie Excellence Grant (MEXT-CT-2004-014203), and Hartmann Müller-Stiftung (project 1252).

1. Dehaene S, Cohen L, Sigman M, Vinckier F (2005) The neural code for written words: A proposal. *Trends Cogn Sci* 9:335–341.
2. Cohen L, et al. (2000) The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123:291–307.
3. Brem S, et al. (2006) Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage* 29:822–837.
4. Vinckier F, et al. (2007) Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron* 55:143–156.
5. Brem S, et al. (2009) Tuning of the visual word processing system: Distinct developmental ERP and fMRI effects. *Hum Brain Mapp* 30:1833–1844.
6. Baker CI, et al. (2007) Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc Natl Acad Sci USA* 104:9087–9092.
7. McCandliss BD, Cohen L, Dehaene S (2003) The visual word form area: Expertise for reading in the fusiform gyrus. *Trends Cogn Sci* 7:293–299.
8. Maurer U, et al. (2007) Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. *Brain* 130:3200–3210.
9. Maurer U, et al. (2006) Coarse neural tuning for print peaks when children learn to read. *Neuroimage* 33:749–758.
10. Bruno JL, Zumberge A, Manis FR, Lu ZL, Goldman JG (2008) Sensitivity to orthographic familiarity in the occipito-temporal region. *Neuroimage* 39:1988–2001.
11. Shaywitz BA, et al. (2007) Age-related changes in reading systems of dyslexic children. *Ann Neurol* 61:363–370.
12. Church JA, Coalson RS, Lugar HM, Petersen SE, Schlaggar BL (2008) A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. *Cereb Cortex* 18:2054–2065.
13. Maurer U, Brem S, Bucher K, Brandeis D (2005) Emerging neurophysiological specialization for letter strings. *J Cogn Neurosci* 17:1532–1552.
14. Turkeltaub PE, Gareau L, Flowers DL, Zeffiro TA, Eden GF (2003) Development of neural mechanisms for reading. *Nat Neurosci* 6:767–773.
15. Helenius P, Tarkiainen A, Cornelissen P, Hansen PC, Salmelin R (1999) Dissociation of normal feature analysis and deficient processing of letter-strings in dyslexic adults. *Cereb Cortex* 9:476–483.
16. Paulesu E, et al. (2001) Dyslexia: Cultural diversity and biological unity. *Science* 291: 2165–2167.
17. Bentin S, Mouchetant-Rostaing Y, Giard MH, Echallier JF, Pernier J (1999) ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *J Cogn Neurosci* 11:235–260.
18. Tarkiainen A, Helenius P, Hansen PC, Cornelissen PL, Salmelin R (1999) Dynamics of letter string perception in the human occipitotemporal cortex. *Brain* 122:2119–2132.
19. Parviainen T, Helenius P, Poskiparta E, Niemi P, Salmelin R (2006) Cortical sequence of word perception in beginning readers. *J Neurosci* 26:6052–6061.
20. Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J (2001) DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychol Rev* 108:204–256.
21. Maurer U, McCandliss BD (2007) The development of visual expertise for words: The contribution of electrophysiology. *Single-Word Reading: Biological and Behavioural Perspectives*, eds Grigorenko EL, Naples AJ (Lawrence Erlbaum Associates, Mahwah, NJ), pp 43–64.
22. Nobre AC, Allison T, McCarthy G (1994) Word recognition in the human inferior temporal lobe. *Nature* 372:260–263.
23. Wagner RK, Torgesen JK (1987) The nature of phonological processing and the acquisition of reading skills. *Psychol Bull* 101:192–212.
24. Xue G, Chen C, Jin Z, Dong Q (2006) Language experience shapes fusiform activation when processing a logographic artificial language: An fMRI training study. *Neuroimage* 31:1315–1326.
25. Xue G, Poldrack RA (2007) The neural substrates of visual perceptual learning of words: Implications for the visual word form area hypothesis. *J Cogn Neurosci* 19: 1643–1655.
26. Lyytinen H, Ronimus M, Alanko A, Poikkeus AM, Taanila M (2007) Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nord Psykol* 59:109–126.
27. Tagamets MA, Novick JM, Chalmers ML, Friedman RB (2000) A parametric approach to orthographic processing in the brain: An fMRI study. *J Cogn Neurosci* 12:281–297.
28. Haxby JV, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430.
29. Frith U (1985) Beneath the surface of developmental dyslexia. *Surface Dyslexia, Neuropsychological and Cognitive Studies of Phonological Reading*, eds Patterson K, Marshall J, Coltheart M , (Erlbaum, London), pp 301–330.
30. Sommer IE, Aleman A, Bouma A, Kahn RS (2004) Do women really have more bilateral language representation than men? A meta-analysis of functional imaging studies. *Brain* 127:1845–1852.
31. Lefly DL, Pennington BF (2000) Reliability and validity of the adult reading history questionnaire. *J Learn Disabil* 33:286–296.
32. Raven JC (2002) *Coloured Progressive Matrices (CPM)* (Harcourt Test Services, Frankfurt).
33. Jansen A, Mannhaupt G, Marx H, Skowronek H (1999) *Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC)* (Hogrefe, Göttingen).
34. Elben CE, Lohaus A (2000) *Marburger Sprachverständnistest für Kinder* (Hogrefe, Göttingen).
35. Landerl K, Wimmer H, Moser E (1997) *Der Salzburger Lese- und Rechtschreibtest (SLRT)* (Huber, Bern).
36. Grave de Peralta Menendez R, Gonzalez Andino S, Lantz G, Michel CM, Landis T (2001) Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topogr* 14:131–137.