

The Association between Gray Matter Volume and Reading Proficiency: A Longitudinal Study of Beginning Readers

Janosch Linkersdörfer^{1,2}, Alina Jurcoane^{1,3}, Sven Lindberg^{1,2},
Jochen Kaiser³, Marcus Hasselhorn^{1,2,3}, Christian J. Fiebach^{1,3,4},
and Jan Lonnemann^{1,2}

Abstract

■ Neural systems involved in the processing of written language have been identified by a number of functional imaging studies. Structural changes in cortical anatomy that occur in the course of literacy acquisition, however, remain largely unknown. Here, we follow elementary school children over their first 2 years of formal reading instruction and use tensor-based morphometry to relate reading proficiency to cortical volume at baseline and follow-up measurement as well as to intraindividual longitudinal volume development between the two measurement time points. A positive relationship was found between

baseline gray matter volume in the left superior temporal gyrus and subsequent changes in reading proficiency. Furthermore, a negative relationship was found between reading proficiency at the second measurement time point and intraindividual cortical volume development in the inferior parietal lobule and the precentral and postcentral gyri of the left hemisphere. These results are interpreted as evidence that reading acquisition is associated with preexisting structural differences as well as with experience-dependent structural changes involving dendritic and synaptic pruning. ■

INTRODUCTION

Reading is a relatively recent human cultural invention that has to be taught explicitly and practiced intensively to be mastered. As a consequence, no genetically predisposed neural systems for reading exist. Instead, learning to read is thought to involve the use of a distributed set of brain regions originally serving other purposes (Dehaene, 2004).

The neural basis of reading has been investigated extensively in functional imaging studies (for meta-analyses, see, e.g., Bolger, Perfetti, & Schneider, 2005; Jobard, Crivello, & Tzourio-Mazoyer, 2003; Turkeltaub, Eden, Jones, & Zeffiro, 2002). Most consistently, these studies identified three neural systems in the left hemisphere to play an important role in the processing of written language. The first system includes parts of the posterior superior temporal gyrus and the inferior parietal lobule and is assumed to support the mapping of orthographic elements of written words (graphemes) onto corresponding phonological elements (phonemes);

a second system in the inferior frontal gyrus has been associated with articulatory processes and the active analysis of phonological elements (e.g., Schlaggar & McCandliss, 2007); and finally, a system in occipitotemporal regions, including parts of the extrastriate fusiform and inferior temporal gyri, is considered to facilitate the direct mapping between familiar visual words or frequent letter combinations within words and corresponding semantic and phonological representations (e.g., McCandliss, Cohen, & Dehaene, 2003; Cohen et al., 2000).

A common developmental hypothesis is that, in the beginning of formal reading instruction, children primarily make use of the temporoparietal system to decode written words in a laborious letter-by-letter (or syllable-by-syllable) manner. Through continuous practice and in tight coupling with the temporoparietal system, the occipitotemporal system is increasingly incorporated, eventually resulting in automatic and fluent reading performance (Pugh et al., 2001). These assumptions have been supported by functional imaging studies showing greater activity in children relative to adults in the temporoparietal system (Church, Coalson, Lugar, Petersen, & Schlaggar, 2008; Booth et al., 2001) and longitudinal increases in activity over the first years of reading instruction in the occipitotemporal system in reading or reading-related tasks (Ben-Shachar, Dougherty, Deutsch, & Wandell, 2011).

¹Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany, ²German Institute for International Educational Research, Frankfurt am Main, Germany, ³Goethe University, Frankfurt am Main, Germany, ⁴Radboud University, Nijmegen, The Netherlands

Despite the large body of functional imaging research, only a few studies have examined neurostructural correlates of reading, that is, how different levels of reading skill relate to differences in cortical gray matter structure. These studies have focused mainly on comparisons between control participants and persons suffering from developmental dyslexia, a severe difficulty in learning to read accurately and fluently. Although both reductions and increases of gray matter properties have been reported in dyslexic readers, the results of two recent coordinate-based meta-analyses of voxel-based morphometry (VBM) studies point to gray matter reductions in bilateral temporoparietal (Richlan, Kronbichler, & Wimmer, 2013; Linkersdörfer, Lonnemann, Lindberg, Hasselhorn, & Fiebach, 2012), left occipitotemporal, and bilateral cerebellar (Linkersdörfer et al., 2012) regions as the most consistent neurostructural correlates of dyslexia.

Studies investigating neurostructural correlates of reading in nonimpaired populations are even more sparse and so far yielded mixed results. More specifically, a positive association between reading performance and measures of gray matter volume, density, or thickness has been demonstrated in bilateral (Blackmon et al., 2010) or left (He et al., 2013; Steinbrink et al., 2008) temporoparietal and left (Zhang et al., 2013) or right (Hoeft et al., 2007) occipitotemporal regions. In addition, negative correlations have been reported for left occipitotemporal and bilateral inferior frontal regions (Blackmon et al., 2010) and the left posterior cingulate cortex (He et al., 2013). Furthermore, as all of these studies employed a cross-sectional design (i.e., participants were studied at only one point in time), they cannot answer the question of whether reading-related differences in cortical anatomy arise in the course of reading instruction and represent the result of different levels of reading proficiency or whether they can be regarded as a pre-existing condition possibly causing differences in reading proficiency. One VBM study (Carreiras et al., 2009) found higher bilateral temporoparietal gray matter density in late-literates (persons who acquired literacy as adults) compared with illiterates, which supports the hypothesis of experience-dependent structural changes. It remains unclear, however, whether structural changes associated with reading acquisition in adulthood resemble those that occur when reading is learned in childhood. Cortical gray matter structure has been shown to undergo massive changes throughout childhood and adolescence, dominated mainly by neuronal and synaptic pruning processes, which result in cortical thinning (e.g., Huttenlocher & Dabholkar, 1997). Pruning is thought to reflect experience-dependent plastic changes, which is supported by studies reporting an association between cortical thinning and better cognitive performance (e.g., Squeglia, Jacobus, Sorg, Jernigan, & Tapert, 2013; Sowell et al., 2004).

The present study examined possible gray matter changes associated with the acquisition of written language by following elementary school children longitudinally in their

first 2 years of reading instruction. For this longitudinal sample, we collected both standardized reading tests and structural magnetic resonance imaging (MRI) data. To relate children's reading proficiency to cortical volume differences at the baseline and follow-up measurements as well as to differences in the intraindividual cortical volume development between the two measurement time points, we used tensor-based morphometry (TBM; see, e.g., Ashburner & Friston, 2003). In contrast to VBM, which identifies regional neurostructural differences by segmentation of tissue classes and statistical analysis of the resulting tissue probability maps after nonlinear transformation to a common space, TBM derives information from high-resolution deformation fields obtained from these nonlinear transformations. In cross-sectional TBM studies, participants' structural images are nonlinearly transformed to a common anatomical template; in longitudinal studies, structural images of the same participant are nonlinearly transformed to each other. From the resulting deformation tensor fields, the Jacobian matrix is computed for each voxel, whose determinant quantifies the fractional volume expansion and contraction needed to align the two images. These relative local volume differences between a participant's image and the template, that is, the group average, in a cross-sectional analysis or between images from the same participant at different time points in a longitudinal analysis are then, after transformation to a common space, used in the statistical analysis. TBM is considered to offer more precise and robust results than VBM, especially in longitudinal analyses (Anderson et al., 2012; Hua et al., 2009; Lee et al., 2007).

METHODS

Participants

Twenty-eight normally developing children participated in the MRI procedure at both measurement time points. The study was approved by the ethics committee of the chamber of physicians of the state of Hessen, Germany, and informed consent was obtained from all participating children and their parents. Six children were excluded from the analysis because of movement artifacts (see below), resulting in a final sample of 22 children (10 girls). All children were in their first year of elementary school at the first measurement time point (mean age = 7.5 years, $SD = 0.4$ years) and in their second year of elementary school at the second measurement time point (mean age = 8.4 years, $SD = 0.4$ years), with approximately 1 year in between (mean time between measurements = 349 days, $SD = 45$ days).

Cognitive Assessment

At both measurement time points, reading proficiency was assessed by means of two standardized reading tests targeting comprehension and fluency at the single word

level. Comprehension ability was measured with the word comprehension subtest of a German reading test with norms for grades 1–6 (ELFE 1–6; Lenhard & Schneider, 2006). In this test, children see a picture and select the appropriate written word out of four alternatives. Distractor words are phonologically and orthographically similar to the target word and matched for length in syllables. The test consists of 72 items, and children solve as many of them as possible in 3 min. Reading fluency was measured with the fluency subtest of a further German reading test with norms for grades 1–6 (SLRT-II; Moll & Landerl, 2010). In this test, children read out loud as many words as possible from a list of 156 existing words and a second list of 156 pseudowords, with a time limit of 1 min per list. Standardized test scores (*T* scores) for the two reading tests were averaged to yield a composite score representing a robust measure of the child's reading proficiency at the respective measurement time point. In addition, to assess intraindividual changes in reading proficiency between the two measurement time points, a change score was computed for each child by subtracting the composite score at the first measurement time point from the composite score at the second measurement time point. As a measure of general intelligence, Raven's Colored Progressive Matrices (Bulheller & Häcker, 2002) were administered at the second measurement time point. This test is an untimed power test designed for children aged 4–11 years and consists of 36 colored diagrammatic puzzles, each with a missing part that has to be chosen from six alternatives. The number of correctly solved items was used as a measure of intelligence in the analysis. All cognitive measures were taken on a separate day preceding the MRI scanning session.

Image Acquisition

Children were accustomed to the scanning procedure and trained to lie still using a custom-built MRI simulator (mock scanner) on the day of the cognitive assessment. Structural MRI data were acquired with a 3-T Siemens Magnetom Tim Trio scanner using a standard transmit–receive head coil. The same protocol was used for each participant and both measurement time points. T1-weighted images were acquired using an optimized Modified Driven Equilibrium Fourier Transform (Deichmann, Schwarzbauer, & Turner, 2004) sequence (repetition time = 7.92 sec, echo time = 2.48 sec, flip angle = 16°, FOV = 256 × 224 mm, 176 sagittal slices, voxel size = 1 × 1 × 1 mm). During the scanning procedure, the children were instructed to lie as still as possible and were able to watch a movie.

Image Processing and TBM

Individual structural images from both measurement time points were visually inspected with regards to signal-to-noise ratio and screened for movement artifacts by two independent researchers. On the basis of these

inspections, six children had to be excluded from the analysis because of excessive movement artifacts. Image preprocessing included removal of non-brain tissue using the hybrid watershed/surface deformation procedure (Ségonne et al., 2004) implemented in the Freesurfer software package (Fischl, 2012) and correction for field inhomogeneity with the N4ITK algorithm (Tustison et al., 2010). After these steps, the images were submitted to two processing pipelines: a cross-sectional pipeline assessing interindividual differences in cortical volume at the first and second measurement time points, respectively, and a longitudinal pipeline assessing interindividual differences in intraindividual cortical volume development between the two measurement time points. Processing was carried out using the Advanced Normalisation Tools software package (Avants et al., 2011).

Cross-sectional processing involved the creation of group-specific optimal average templates for each of the two measurement time points by iteratively refining nonlinear coregistrations of all images from the respective measurement time point (see Klein et al., 2010). After a linear transformation, each individual image from one measurement time point was then nonlinearly transformed to the template from the same measurement time point using the Symmetric Normalization algorithm (Avants, Epstein, Grossman, & Gee, 2008) implemented in the Advanced Normalisation Tools. For the resulting deformation tensor field, the determinant of the Jacobian matrix was computed, which quantifies the fractional volume expansion and contraction required to match the template at each voxel. The Jacobian maps were subjected to a log transformation to render the values more normally distributed (Avants & Gee, 2004). To allow for group analysis of the Jacobian maps in a common reference space, the linear transformation between the respective group template and the Montreal Neurological Institute (MNI) space was computed and used to transform the Jacobian map of each participant into MNI space. Before statistical analysis, the Jacobian maps were smoothed with a 4-mm FWHM Gaussian filter.

Longitudinal processing involved registration of each participant's image from the second measurement time point (follow-up scan) to his or her image from the first measurement time point (baseline scan). To avoid potential bias resulting from asymmetric interpolation (see, e.g., Fox, Ridgway, & Schott, 2011), the linear transformation between the follow-up and the baseline image was split halfway between the two images according to the approach described by Yushkevich et al. (2010). The symmetric linear transformations from the respective images to the halfway space were then used to initialize the nonlinear registration of the follow-up to the baseline image. Jacobian determinants of the resulting deformation field, in this case indicating the voxel-wise fractional volume expansion and contraction between the two measurement time points in the respective participant, were computed and log transformed. The Jacobian map of

each participant was finally transformed into MNI space by combining (a) the linear transformation from the halfway space to the baseline image space with (b) the linear and nonlinear transformations from the baseline image space to the baseline group template space and (c) the linear transformation from the baseline group template space to the MNI space, which had been computed in the cross-sectional processing pipeline. Finally, the Jacobian maps were smoothed with a 4-mm FWHM Gaussian filter.

Statistical Analysis

Whole-brain voxel-wise regression analysis was performed on the smoothed Jacobian images using the SPM8 software package (Wellcome Trust Centre for Neuroimaging, University College London, London, United Kingdom). First, cortical volume differences at the first measurement time point (T1) were related to (a) the composite score for reading proficiency at T1, (b) the composite score for reading proficiency at the second measurement time point (T2), and (c) the change in reading proficiency between T1 and T2 as variable of interest. Participant's intelligence score, gender, and age at T1 in days were included in the model as nuisance variables. Second, cortical volume differences at T2 were related to the composite score for reading proficiency at T2, and participant's intelligence score, gender, and age at T2 in days were included in the model as nuisance variables. Finally, differences in the intraindividual cortical volume development between T1 and T2 were related to the composite score for reading proficiency at T2, and participant's intelligence score, gender, age at T1, and time interval between T1 and T2, both measured in days, were included in the model as nuisance variables.

For all group-level statistics, a gray matter tissue probability map, adjusted for the age and gender distribution of the sample, was generated using the Template-O-Matic toolbox (Wilke, Holland, Altaye, & Gaser, 2008), thresholded at $p = .25$ and binarized to yield an explicit gray matter mask that was applied to the statistical analysis. Clusters of voxels were considered significant and reported when exceeding a voxel level threshold of $p < .001$ (uncorrected) and a cluster size threshold of $p < .05$, corrected for multiple comparison using Gaussian random field theory (Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994).

Analysis of behavioral performance and post hoc analyses of the TBM results were carried out using the statistical package R, version 2.15.1 (R Project for Statistical Computing, Vienna, Austria; R Core Team, 2012).

RESULTS

Behavioral Performance

Reading proficiency in our sample, measured by composite scores combining comprehension and fluency perfor-

mance, varied between T scores of 42.25 and 69 (mean = 57.10, $SD = 7.13$) at T1 and between 39 and 66.5 (mean = 53.92, $SD = 6.79$) at T2, indicating that children's reading proficiency varied over the whole range of normal competencies, with a tendency toward above-average performance. We observed high correlations between the different reading (sub)tests (i.e., comprehension, fluency-subtest words, fluency-subtest pseudowords) at both measurement time points (T1: $r = .7-.83$; T2: $r = .6-.66$; all $ps < .005$), supporting our decision to integrate them into a single composite score. Moreover, reading performance over the two measurement time points was highly correlated (composite score, $r = .8$; comprehension, $r = .73$; fluency-subtest words, $r = .69$; fluency-subtest pseudowords, $r = .75$; all $ps < .001$). Reading proficiency change scores varied between T scores of -11.75 and 5 (mean = -3.18 , $SD = 4.44$). The mean decrease in reading proficiency scores between T1 and T2 might, at least in part, be attributable to the fact that reading proficiency scores in our sample were relatively high at T1; thus, subsequent measures were more likely to result in lower values, a statistical phenomenon known as "regression toward the mean" (e.g., Bland & Altman, 1994). A moderate negative relationship between the reading composite score at T1 and the reading proficiency change score ($r = -.39$, $p = .08$) supports this assumption. Intelligence scores (IQ) ranged from 88 to 135 (mean = 107.23, $SD = 11.33$), indicating, again, normal to above-average abilities in our sample. At both measurement time points, reading proficiency did not correlate significantly with intelligence (measurement time point 1: $r = .08$, $p = .726$; measurement time point 2: $r = .05$, $p = .837$) or sex (coded as 0 for girls and 1 for boys; measurement time point 1: point biserial $r = -.4$, $p = .07$; measurement time point 2: point biserial $r = -.31$, $p = .165$).

TBM

A cluster showing a positive relationship between gray matter volume at T1 and change in reading proficiency scores between T1 and T2 could be found in the left superior temporal gyrus (cluster extent = 1469 voxels, cluster level $p = .034$ corrected, peak t in cluster = 5.93; cf. Table 1 and Figure 1). The analyses relating volume differences at T1 to reading proficiency at T1 or T2 and the analysis relating volume differences at T2 to reading proficiency at T2 did not result in significant findings. To exclude the possibility that the significant cluster indicating a larger volume of left superior temporal areas in children with higher reading proficiency change scores was solely based on differences in the baseline level of reading proficiency, we computed the partial correlation between the mean Jacobian value, averaged over all voxels contained in the cluster, and the reading change score, while controlling for reading proficiency at T1 and the same variables as in the whole-brain analysis

Table 1. Significant Cluster Showing a Positive Association between Gray Matter Volume at the First Measurement Time Point and Change in Reading Proficiency

Region	MNI Coordinates of Local Maxima			<i>t</i>	Cluster Size (Voxels)	Cluster <i>p</i> Value
	<i>x</i>	<i>y</i>	<i>z</i>			
L superior temporal gyrus	-64	-38	10	5.93	1469	.034
	-64	-27	13	5.55	^a	^a

^aSubpeak within cluster.

(i.e., intelligence score, sex, age at T1). The results of this post hoc analysis were significant (r partial = 0.83, $p < .001$), indicating that gray matter volume in the superior temporal cluster at T1 was indeed related to changes in reading proficiency, even when accounting for baseline level in reading proficiency. Furthermore, to investigate whether cortical volume in the superior temporal cluster was related differently to developmental changes in the

separate reading (sub)tests (i.e., comprehension, fluency-subtest words, fluency-subtest pseudowords) summarized in the reading change score, partial correlations between change scores for the separate (sub)tests and the mean Jacobian value, averaged over all voxels contained in the cluster, were computed, while, again, controlling for the same variables as in the whole-brain analysis. The results show significant positive correlations for all reading (sub)tests

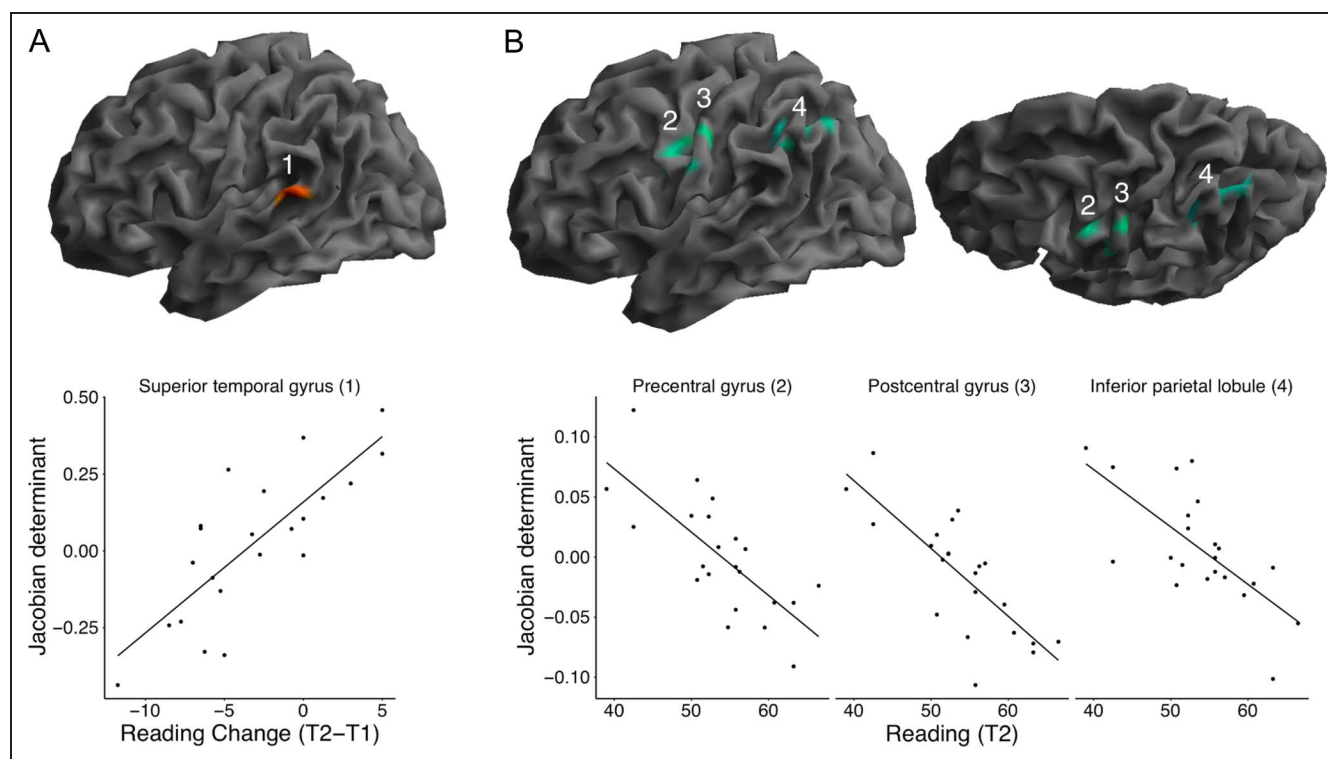


Figure 1. Associations between cortical volume and reading proficiency. (A, top) Surface rendering of the left brain hemisphere displaying the cluster (red) in the superior temporal gyrus showing a positive association between gray matter volume at the first measurement time point (T1) and change in reading proficiency between the first and second measurement time points (T2). (A, bottom) Relationship between reading proficiency change (change score subtracting the composite score of standardized reading tests at T1 from the composite score at T2) and the gray matter volume at T1 (log-transformed Jacobian determinants, averaged over all voxels in the cluster; positive values indicate relatively larger volume, and negative values indicate relatively smaller volume compared with the group average). (B, top) Surface renderings showing the three left-hemispheric clusters (green) that display a negative association between gray matter volume change between T1 and T2 and reading proficiency at T2, in the precentral (2) and postcentral (3) gyri and in the inferior parietal lobule (3). (B, bottom) Scatterplots illustrating the relationship between reading proficiency (composite score of standardized reading tests at T2) and gray matter volume development between T1 and T2 (log-transformed Jacobian determinants, averaged over all voxels in the respective clusters; positive values indicate relative volume expansion, and negative values represent relative volume contraction). For purposes of visualization, significant effects (cf. Tables 1 and 2) are displayed at $p < .0025$.

Table 2. Significant Clusters Showing a Negative Association between Gray Matter Volume Development and Reading Proficiency at the Second Measurement Time Point

Region	MNI Coordinates of Local Maxima			<i>t</i>	Cluster Size (Voxels)	Cluster <i>p</i> Value
	<i>x</i>	<i>y</i>	<i>z</i>			
L precentral gyrus	-47	-7	37	-5.60	743	.018
	-45	-2	28	-4.94	a	a
	-52	0	24	-4.02	a	a
L postcentral gyrus	-56	-18	42	-5.61	603	.049
	-59	-11	29	-4.82	a	a
L angular gyrus	-36	-62	40	-5.15	978	.004
	-44	-55	37	-4.83	a	a
L supramarginal gyrus	-49	-48	35	-5.08	a	a

^aSubpeak within cluster.

(*r* partial = .59–.7, all *ps* < .005), indicating that the different subtests were equally associated with gray matter volume in the cluster.

Relating differences in the intraindividual cortical volume development between the two measurement time points to reading proficiency at T2 resulted in three significant clusters showing a negative relationship (Table 2; Figure 1). The largest cluster (cluster extent = 978 voxels, cluster level *p* = .004 corrected, peak *t* in cluster = -5.15) was located in the left inferior parietal lobule, including mainly the supramarginal and angular gyri. The second cluster was found in the left precentral gyrus (cluster extent = 743 voxels, cluster level *p* = .018 corrected, peak *t* in cluster = -5.6), and the third cluster was located in the left postcentral gyrus (cluster extent = 603 voxels, cluster level *p* = .049 corrected, peak *t* in cluster = -5.61). No significant clusters showing a positive relationship between reading proficiency and volume development were found. Possible differences in the relationship between cortical development in the three clusters and performance in the separate reading (sub)tests (i.e., comprehension, fluency-subtest words, fluency-subtest pseudowords), which were summarized in the composite score for reading proficiency, were examined by computing partial correlations between *T* scores for the separate reading (sub)tests at T2 and mean Jacobian values for the separate clusters, averaged over all voxels contained in the respective cluster, including the same nuisance variables as in the whole-brain analysis (i.e., intelligence score, gender, age at T1, and time interval between T1 and T2). The results show significant negative correlations for all reading (sub)tests and clusters (*r* partial = -.62 to -.75, all *ps* < .005), indicating that the different (sub)tests were equally associated with volume development in the three clusters.

DISCUSSION

In this study, we used TBM to investigate the relationship between reading proficiency and cortical gray matter volume in elementary school children during their first 2 years of formal reading instruction. Higher baseline gray matter volume in the left superior temporal cortex was positively related to subsequent changes in reading proficiency. Moreover, longitudinal volume reductions in the inferior parietal lobule and the precentral and postcentral gyri of the left hemisphere were related to higher reading proficiency at the follow-up measurement.

The significant cluster showing a positive relationship between gray matter volume at the first measurement time point and change in reading proficiency between the two measurement time points was located in the left posterior superior temporal gyrus, including parts of the planum temporale. These areas, which lie directly posterior to the primary auditory cortex, are considered as part of “Wernicke’s area” and have been implicated in different language functions (for an overview, see, e.g., Price, 2012). The planum temporale, because of its marked leftward asymmetry, has traditionally been considered as an auditory association area specific to speech perception (Geschwind & Levitsky, 1968). More recent research, however, has shown that this region is equally activated by speech and tone stimuli (Binder et al., 2000; Binder, Frost, Hammeke, Rao, & Cox, 1996). Furthermore, the planum temporale, especially its posterior portions near the TPJ, seems to be involved in both speech perception and speech production (Hickok, Okada, & Serences, 2009; Wise et al., 2001) but also in the audio-motor interaction between nonspeech sounds and motor actions (Hickok, Buchsbaum, Humphries, & Muftuler, 2003). Therefore, current models of speech processing view the planum temporale as a sensorimotor integration area,

which relays information between auditory cortical regions and anterior systems involved in motor control of the vocal tract and which therefore plays an important role in the acquisition of spoken language (Hickok, Houde, & Rong, 2011). The association between the gray matter volume of this area in beginning readers and subsequent changes in their reading proficiency, as demonstrated by our study, suggests that the planum temporale might also be involved in the acquisition of written language. This assumption is supported by functional imaging studies reporting overlapping left planum temporale activation for listening to and reading words (Buchsbaum et al., 2005; Nakada, Fujii, Yoneoka, & Kwee, 2001). Therefore, this region might facilitate the multisensory integration of visual input and auditory–motor representations.

The left inferior parietal cluster, which exhibited a negative association between cortical volume development and reading proficiency at the second measurement time point, was located in the supramarginal and angular gyri, areas that correspond well to the anatomical characterization of the temporoparietal part of the phonological reading system (e.g., Démonet, Taylor, & Chaix, 2004; Rumsey et al., 1997). The temporoparietal system is assumed to support the transformation of orthographic elements of visual words (graphemes) into associated phonological elements (phonemes; Booth et al., 2002; Shaywitz et al., 2002). Thus, structural changes in this region over the first years of reading instruction, as demonstrated by our study, are in line with developmental accounts of reading acquisition, suggesting that children rely mainly on phonological recoding strategies in the beginning stages of learning to read (Ehri, 1999; Frith, 1985; Chall, 1983). Furthermore, our results are in agreement with the neuroanatomical model of reading development by Pugh et al. (2001), which suggests a predominant role of the temporoparietal system in beginning readers, a claim that has been supported by findings from functional imaging studies showing greater activity in children relative to adults in these regions in response to reading or reading-related tasks (Church et al., 2008; Booth et al., 2001).

Two additional clusters showing negative associations between volume development and reading proficiency were found in the precentral and postcentral gyri of the left hemisphere. Neural populations in proximity to the peaks of these clusters have been attributed to the sensorimotor control of movements of the lips, tongue, jaw, and larynx (e.g., Grabski et al., 2012; Bohland & Guenther, 2006; Pulvermüller et al., 2006) and have been shown to produce speech through highly coordinated sequences of spatial patterns of activity that hierarchically represent the articulatory features of phonemes (Bouchard, Mesgarani, Johnson, & Chang, 2013). Furthermore, activity in these speech production areas, especially in the left hemisphere, has also been demonstrated for the passive perception of speech sounds (Pulvermüller et al., 2006;

Wilson, Saygin, Sereno, & Iacoboni, 2004), supporting theories postulating a motor involvement in speech perception (Galantucci, Fowler, & Turvey, 2006; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). These findings have been interpreted as evidence for the existence of cell assemblies integrating articulatory and auditory neurons to represent the phonological properties of speech sounds, which might develop in response to an increased tuning to the phonemic structure of the native language in the early stages of learning to speak (Pulvermüller & Fadiga, 2010). The formation of such an auditory–motor network might be facilitated by the multisensory integration function of the planum temporale discussed above. The left precentral gyrus has further been implicated as part of the neural reading network by meta-analyses of functional imaging studies investigating nonimpaired adults (Bolger et al., 2005; Turkeltaub et al., 2002) and children (Houdé, Rossi, Lubin, & Joliot, 2010). Thus, neural representations of articulatory features of phonemes also seem to play an important role in the processing of written language.

Indeed, learning to read an alphabetic script seems to increase children's awareness of the phonological structure of spoken language through the incorporation of orthographic information, ultimately leading to a reorganization of existing phonological representations (Nation & Hulme, 2011; Pattamadilok, Knierim, Kawabata Duncan, & Devlin, 2010; Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998). The observed structural changes in the precentral and postcentral gyri of the left hemisphere, which co-occur with similar changes in the left inferior parietal lobule, might therefore be interpreted as the neural correlate of such reorganizational processes. In beginning readers, the intensive use of phonological recoding strategies implemented by neural circuits in the left inferior parietal cortex may lead to the mapping of visual (i.e., orthographic) features of words onto preexisting phonological representations in left hemisphere circuits subserving speech perception and production (Pattamadilok et al., 2010). The formation of a higher-order neural assembly that integrates orthographic and phonological information may promote the reciprocal refinement of grapheme–phoneme conversion knowledge and articulatory–auditory representations of phonemes and might possibly be further strengthened by the coactivation of inferior parietal with precentral and postcentral neurons because of lip movements and subvocal articulation, typically found in children in the beginning stages of learning to read (e.g., Huey, 1968).

Previous studies that related reading performance to measures of gray matter volume, gray matter density, or gray matter thickness in normal reading participants in a cross-sectional manner, that is, at only one point in time, reported mostly positive associations in temporoparietal (He et al., 2013; Blackmon et al., 2010; Steinbrink et al., 2008) and occipitotemporal regions (Zhang et al., 2013; Hoeft et al., 2007). In contrast to these studies, no

association between gray matter volume at baseline or follow-up measurement and reading proficiency at the respective (or the following) measurement time point could be found in this study. However, intraindividual changes in reading proficiency between the two measurement time points were positively related to baseline gray matter volume in the superior temporal gyrus, including parts of the planum temporale, that is, similar areas as reported by some of the studies cited above (i.e., He et al., 2013; Blackmon et al., 2010; Steinbrink et al., 2008). Furthermore, it resonates with studies reporting neurostructural differences between normal reading participants and participants suffering from developmental dyslexia, a specific impairment affecting the acquisition of reading skills. Most specifically, a recent meta-analysis of VBM studies (Richlan et al., 2013) found a cluster of reliable gray matter reductions in dyslexic readers in superior temporal regions only slightly more inferior to the superior temporal cluster reported in this study. Moreover, dyslexic readers have been reported to not show the typical leftward asymmetry of the planum temporale (e.g., Hynd, Semrud-Clikeman, Lorys, Novey, & Eliopoulos, 1990; Larsen, Høien, Lundberg, & Odegaard, 1990), although reports have been inconsistent (for a review, see Heim & Keil, 2004). Structural abnormalities of the planum temporale and surrounding areas have further been associated with other developmental speech and language disorders like specific language impairment (e.g., Badcock, Bishop, Hardiman, Barry, & Watkins, 2012) or stuttering (e.g., Foundas et al., 2004). We therefore interpret the significant positive relationship between left superior temporal gray matter volume and subsequent changes in reading proficiency as evidence for preexisting neurostructural differences, which might be associated with differences in general linguistic abilities. These abilities may be of increasing importance in the course of long-term reading development (e.g., Dickinson, Golinkoff, & Hirsh-Pasek, 2010), which might explain why, in contrast to previous studies investigating neurostructural correlates of reading in older participants, no association between interindividual differences in gray matter volume in this area and reading proficiency measured at the respective time point could be found in the beginning readers investigated in this study.

Our results further demonstrated a negative relationship between longitudinal cortical volume development and reading proficiency at the follow-up measurement in inferior parietal and precentral and postcentral regions. Whereas better readers on average exhibited volume reductions, poor readers on average exhibited volume growth in these areas. These findings seem inconsistent with previous findings of positive correlations between measures of cortical volume and reading proficiency in left inferior parietal regions, as reported by some of the studies cited above (i.e., He et al., 2013; Blackmon et al., 2010). Furthermore, the only previous study that reports volume changes directly related to reading acquisition found higher gray matter density in temporoparietal regions in

persons who acquired literacy as adults when compared with illiterate adults (Carreiras et al., 2009). It remains unclear, however, whether the processes of neural plasticity underlying learning in adulthood are the same as those underlying learning in childhood. Although changes in cortical anatomy can be regarded as the product of a mixture of developmental and experiential processes at any point in development, the influence of these processes may vary over time (Galván, 2010). We thus hypothesize that the observed discrepancies may be because of the age and/or the developmental stage of the participants examined in this study: Cortical development is characterized by an excessive increase in brain connectivity through dendritic branching and the formation of new synapses in infancy and early childhood, followed by a phase dominated by dendritic and synaptic pruning in late childhood and adolescence (e.g., Huttenlocher & Dabholkar, 1997). This pruning phase is thought to reflect the experience-dependent strengthening of frequently used and weakening or elimination of sparsely used connections; it might mark a shift from relatively diffuse cortical representations of cognitive functions in early childhood toward a more accurate, efficient, and faster, that is, more adult like, processing in specialized neural networks, which are further refined throughout life (Huttenlocher, 2002). We suggest that the observed cortical volume reductions in better readers over the first 2 years of elementary school represent the formation of a more mature and fine-tuned, left-lateralized cortical network for the processing of phonological features of speech sounds.

This assumption is in line with recent studies that report an association between cortical thinning and cognitive performance. For example, Sowell et al. (2004) found a negative association between longitudinal cortical thickness development and vocabulary knowledge in elementary school children in the left dorsal frontal and parietal lobes. Similarly, thinner parietal cortices were reported to be associated with better performance in verbal learning and memory, visuospatial functioning, and spatial planning and problem solving (Squeglia et al., 2013). Furthermore, thinner, that is, more mature, cortical gray matter seems to be associated with more mature patterns of brain activity and increased functional selectivity in cognitive tasks (Wendelken, O'Hare, Whitaker, Ferrer, & Bunge, 2011; Lu et al., 2009). Thus, pruning processes might be of special importance in younger children at the beginning of reading development and may lay out the basic neural foundation of the reading network. In later stages of reading development, where learning builds on largely defined neural and cognitive systems and overall plasticity is reduced, the reading network may be consolidated through increased network-internal connectivity. This might be accompanied by increased dendritic branching and synapse formation (e.g., Draganski & May, 2008), which may explain the findings of higher cortical volume in better readers in studies examining older participants. This, however, we suggest may be a plasticity-related process that is

secondary to the cortical volume changes we observed in the early stages of reading acquisition. The observed differences in volume development between good and poor readers might be due to interindividual differences in the coordination of pruning and growth processes (cf. Yeatman, Dougherty, Ben-Shachar, & Wandell, 2012). Although in better readers, pruning might be more prominent and result in the faster formation of more mature and efficient neural reading circuits, growth processes might dominate development in the poor readers resulting in more diffuse cortical representations. These experience-dependent neurostructural changes in the beginning of reading instruction might be transient and vanish in the course of development.

To conclude, this study investigated neurostructural correlates of reading in beginning readers by following elementary school children over their first 2 years of reading instruction. A positive relationship was found between left superior temporal gray matter volume at the baseline measurement and subsequent changes in reading proficiency. This finding may represent preexisting neurostructural differences that might be associated with long-term potential for the development of reading and other language skills. In addition, a negative relationship was found between cortical volume development and reading proficiency at the follow-up measurement in cortical areas in the inferior parietal lobule and precentral and postcentral gyri of the left hemisphere. These results might represent experience-dependent neuroanatomical changes due to synaptic and dendritic pruning and, thus, highlight the importance of developmental studies using longitudinal designs, especially in children in the beginning stages of learning to read.

Acknowledgments

This research was funded by the Hessian initiative for the development of scientific and economic excellence (LOEWE). C. J. F. is furthermore supported by an Emmy-Noether Grant from the German Research Foundation (DFG FI-848/3) and by a VIDI Grant from the Netherlands Organization for Scientific Research (45209006). We would like to thank all the participating children and their families for their support.

Reprint requests should be sent to Janosch Linkersdörfer, German Institute for International Educational Research, Schloßstraße 29, 60486 Frankfurt am Main, Germany, or via e-mail: j.linkersdoerfer@idea-frankfurt.eu.

REFERENCES

- Anderson, V. M., Schott, J. M., Bartlett, J. W., Leung, K. K., Miller, D. H., & Fox, N. C. (2012). Gray matter atrophy rate as a marker of disease progression in AD. *Neurobiology of Aging*, *33*, 1194–1202.
- Ashburner, J., & Friston, K. J. (2003). Morphometry. In R. Frackowiak (Ed.), *Human brain function* (pp. 707–722). London: Elsevier Press.
- Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*, 26–41.
- Avants, B. B., & Gee, J. C. (2004). Geodesic estimation for large deformation anatomical shape averaging and interpolation. *Neuroimage*, *23*(Suppl. 1), S139–S150.
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*, *54*, 2033–2044.
- Badcock, N. A., Bishop, D. V. M., Hardiman, M. J., Barry, J. G., & Watkins, K. E. (2012). Co-localisation of abnormal brain structure and function in specific language impairment. *Brain and Language*, *120*, 310–320.
- Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2011). The development of cortical sensitivity to visual word forms. *Journal of Cognitive Neuroscience*, *23*, 2387–2399.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*, 512–528.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, *119*, 1239–1247.
- Blackmon, K., Barr, W. B., Kuzniecky, R., Dubois, J., Carlson, C., Quinn, B. T., et al. (2010). Phonetically irregular word pronunciation and cortical thickness in the adult brain. *Neuroimage*, *51*, 1453–1458.
- Bland, J. M., & Altman, D. G. (1994). Regression towards the mean. *British Medical Journal*, *308*, 1499.
- Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage*, *32*, 821–841.
- Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping*, *25*, 92–104.
- Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2002). Functional anatomy of intra- and cross-modal lexical tasks. *Neuroimage*, *16*, 7–22.
- Booth, J. R., Burman, D. D., Van Santen, F. W., Harasaki, Y., Gitelman, D. R., Parrish, T. B., et al. (2001). The development of specialized brain systems in reading and oral-language. *Child Neuropsychology*, *7*, 119–141.
- Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, *495*, 327–332.
- Buchsbaum, B. R., Olsen, R. K., Koch, P. F., Kohn, P., Kippenhan, J. S., & Berman, K. F. (2005). Reading, hearing, and the planum temporale. *Neuroimage*, *24*, 444–454.
- Bulheller, S., & Häcker, H. O. (2002). *Coloured Progressive Matrices (CPM). Deutsche bearbeitung und normierung nach J. C. Raven*. Frankfurt, Germany: Pearson Assessment.
- Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., et al. (2009). An anatomical signature for literacy. *Nature*, *461*, 983–986.
- Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain*, *121*, 1053–1063.
- Chall, J. S. (1983). *Stages of reading development*. New York: McGraw Hill.
- Church, J. A., Coalson, R. S., Lugar, H. M., Petersen, S. E., & Schlaggar, B. L. (2008). A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. *Cerebral Cortex*, *18*, 2054–2065.
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., 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.
- Dehaene, S. (2004). Evolution of human cortical circuits for reading and arithmetic: The “neuronal recycling” hypothesis. In S. Dehaene, J. R. Duhamel, M. Hauser, & G. Rizzolatti (Eds.), *From monkey brain to human brain* (pp. 133–158). Cambridge, MA: MIT Press.
- Deichmann, R., Schwarzbauer, C., & Turner, R. (2004). Optimisation of the 3D MDEFT sequence for anatomical brain imaging: Technical implications at 1.5 and 3 T. *Neuroimage*, *21*, 757–767.
- Démonet, J. F., Taylor, M. J., & Chaix, Y. (2004). Developmental dyslexia. *Lancet*, *363*, 1451–1460.
- Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, *39*, 305–310.
- Draganski, B., & May, A. (2008). Training-induced structural changes in the adult human brain. *Behavioural Brain Research*, *192*, 137–142.
- Ehri, L. C. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading: A psychological perspective* (pp. 79–108). Oxford: Blackwell Science.
- Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*, 774–781.
- Foundas, A. L., Bollich, A. M., Feldman, J., Corey, D. M., Hurley, M., Lemen, L. C., et al. (2004). Aberrant auditory processing and atypical planum temporale in developmental stuttering. *Neurology*, *63*, 1640–1646.
- Fox, N. C., Ridgway, G. R., & Schott, J. M. (2011). Algorithms, atrophy and Alzheimer’s disease: Cautionary tales for clinical trials. *Neuroimage*, *57*, 15–18.
- Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, *1*, 210–220.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301–330). London: Routledge and Kegan Paul.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*, 361–377.
- Galván, A. (2010). Neural plasticity of development and learning. *Human Brain Mapping*, *31*, 879–890.
- Geschwind, N., & Levitsky, W. (1968). Human brain: Left–right asymmetries in temporal speech region. *Science*, *161*, 186–187.
- Grabski, K., Lamalle, L., Vilain, C., Schwartz, J.-L., Vallée, N., Tropes, I., et al. (2012). Functional MRI assessment of orofacial articulators: Neural correlates of lip, jaw, larynx, and tongue movements. *Human Brain Mapping*, *33*, 2306–2321.
- He, Q., Xue, G., Chen, C., Chen, C., Lu, Z.-L., & Dong, Q. (2013). Decoding the neuroanatomical basis of reading ability: A multivoxel morphometric study. *Journal of Neuroscience*, *33*, 12835–12843.
- Heim, S., & Keil, A. (2004). Large-scale neural correlates of developmental dyslexia. *European Child and Adolescent Psychiatry*, *13*, 125–140.
- Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory–motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*, 673–682.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, *69*, 407–422.
- Hickok, G., Okada, K., & Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory–motor integration for speech processing. *Journal of Neurophysiology*, *101*, 2725–2732.
- Hoefl, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., et al. (2007). Prediction of children’s reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience*, *121*, 602–613.
- Houdé, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: An fMRI meta-analysis of 52 studies including 842 children. *Developmental Science*, *13*, 876–885.
- Hua, X., Leow, A. D., Levitt, J. G., Caplan, R., Thompson, P. M., & Toga, A. W. (2009). Detecting brain growth patterns in normal children using tensor-based morphometry. *Human Brain Mapping*, *30*, 209–219.
- Huey, E. (1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press.
- Huttenlocher, P. R. (2002). *Neural plasticity: The effects of environment on the development of the cerebral cortex*. Cambridge, MA: Harvard University Press.
- Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, *387*, 167–178.
- Hynd, G. W., Semrud-Clikeman, M., Lorys, A. R., Novey, E. S., & Eliopoulos, D. (1990). Brain morphology in developmental dyslexia and attention deficit disorder/hyperactivity. *Archives of Neurology*, *47*, 919–926.
- Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A meta-analysis of 35 neuroimaging studies. *Neuroimage*, *20*, 693–712.
- Klein, A., Ghosh, S. S., Avants, B., Yeo, B. T. T., Fischl, B., Ardekani, B., et al. (2010). Evaluation of volume-based and surface-based brain image registration methods. *Neuroimage*, *51*, 214–220.
- Larsen, J. P., Høien, T., Lundberg, I., & Odegaard, H. (1990). MRI evaluation of the size and symmetry of the planum temporale in adolescents with developmental dyslexia. *Brain and Language*, *39*, 289–301.
- Lee, A. D., Leow, A. D., Lu, A., Reiss, A. L., Hall, S., Chiang, M.-C., et al. (2007). 3D pattern of brain abnormalities in Fragile X syndrome visualized using tensor-based morphometry. *Neuroimage*, *34*, 924–938.
- Lenhard, W., & Schneider, W. (2006). *Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen, Germany: Hogrefe.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.
- Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., & Fiebach, C. J. (2012). Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: An ALE meta-analysis. *PloS One*, *7*, e43122.
- Lu, L. H., Dapretto, M., O’Hare, E. D., Kan, E., McCourt, S. T., Thompson, P. M., et al. (2009). Relationships between brain activation and brain structure in normally developing children. *Cerebral Cortex*, *19*, 2595–2604.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*, 293–299.
- Moll, K., & Landerl, K. (2010). *SLRT-2—Lese- und Rechtschreibtest*. Bern, Switzerland: Hans Huber.
- Nakada, T., Fujii, Y., Yoneoka, Y., & Kwee, I. L. (2001). Planum temporale: Where spoken and written language meet. *European Neurology*, *46*, 121–125.
- Nation, K., & Hulme, C. (2011). Learning to read changes children’s phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science*, *14*, 649–659.
- Pattamadilok, C., Knierim, I. N., Kawabata Duncan, K. J., & Devlin, J. T. (2010). How does learning to read affect

- speech perception? *Journal of Neuroscience*, *30*, 8435–8444.
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, *62*, 816–847.
- Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders*, *34*, 479–492.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, *11*, 351–360.
- Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences, U.S.A.*, *103*, 7865–7870.
- R Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Richlan, F., Kronbichler, M., & Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: A meta-analysis of voxel-based morphometry studies. *Human Brain Mapping*, *34*, 3055–3065.
- Rumsey, J. M., Nace, K., Donohue, B., Wise, D., Maisog, J. M., & Andreason, P. (1997). A positron emission tomographic study of impaired word recognition and phonological processing in dyslexic men. *Archives of Neurology*, *54*, 562–573.
- Schlaggar, B. L., & McCandliss, B. D. (2007). Development of neural systems for reading. *Annual Review of Neuroscience*, *30*, 475–503.
- Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., et al. (2004). A hybrid approach to the skull stripping problem in MRI. *Neuroimage*, *22*, 1060–1075.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*, 101–110.
- Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, *24*, 8223–8231.
- Squeglia, L. M., Jacobus, J., Sorg, S. F., Jernigan, T. L., & Tapert, S. F. (2013). Early adolescent cortical thinning is related to better neuropsychological performance. *Journal of the International Neuropsychological Society*, *19*, 962–970.
- Steinbrink, C., Vogt, K., Kastrup, A., Müller, H.-P., Juengling, F. D., Kassubek, J., et al. (2008). The contribution of white and gray matter differences to developmental dyslexia: Insights from DTI and VBM at 3.0 T. *Neuropsychologia*, *46*, 3170–3178.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage*, *16*, 765–780.
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., et al. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*, 1310–1320.
- Wendelken, C., O’Hare, E. D., Whitaker, K. J., Ferrer, E., & Bunge, S. A. (2011). Increased functional selectivity over development in rostral lateral prefrontal cortex. *Journal of Neuroscience*, *31*, 17260–17268.
- Wilke, M., Holland, S. K., Altaye, M., & Gaser, C. (2008). Template-O-Matic: A toolbox for creating customized pediatric templates. *Neuroimage*, *41*, 903–913.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*, 701–702.
- Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within “Wernicke’s area.” *Brain*, *124*, 83–95.
- Yeatman, J. D., Dougherty, R. F., Ben-Shachar, M., & Wandell, B. A. (2012). Development of white matter and reading skills. *Proceedings of the National Academy of Sciences, U.S.A.*, *109*, E3045–3053.
- Yushkevich, P. A., Avants, B. B., Das, S. R., Pluta, J., Altinay, M., Craige, C., et al. (2010). Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: An illustration in ADNI 3 T MRI data. *Neuroimage*, *50*, 434–445.
- Zhang, M., Li, J., Chen, C., Mei, L., Xue, G., Lu, Z., et al. (2013). The contribution of the left mid-fusiform cortical thickness to Chinese and English reading in a large Chinese sample. *Neuroimage*, *65*, 250–256.