Engagement of Temporal Lobe Regions Predicts Response to Educational Interventions in Adolescent Struggling Readers

Roozbeh Rezaie\textsuperscript{a}, Panagiotis G. Simos\textsuperscript{b}, Jack M. Fletcher\textsuperscript{c}, Paul T. Cirino\textsuperscript{c}, Sharon Vaughn\textsuperscript{d} & Andrew C. Papanicolaou\textsuperscript{a}

\textsuperscript{a} Department of Pediatrics, Children's Learning Institute, University of Texas Health Science Center, Houston, Texas, USA
\textsuperscript{b} Department of Psychology, University of Crete, Rethymno, Greece
\textsuperscript{c} Department of Psychology, University of Houston, Houston, Texas, USA
\textsuperscript{d} The Meadows Center for Preventing Educational Risk, University of Texas, Austin, Texas, USA

Available online: 06 Oct 2011

To cite this article: Roozbeh Rezaie, Panagiotis G. Simos, Jack M. Fletcher, Paul T. Cirino, Sharon Vaughn & Andrew C. Papanicolaou (2011): Engagement of Temporal Lobe Regions Predicts Response to Educational Interventions in Adolescent Struggling Readers, Developmental Neuropsychology, 36:7, 869-888

To link to this article: http://dx.doi.org/10.1080/87565641.2011.606404

Please scroll down for article
demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Engagement of Temporal Lobe Regions Predicts Response to Educational Interventions in Adolescent Struggling Readers

Roozbeh Rezaie

Department of Pediatrics, Children’s Learning Institute, University of Texas Health Science Center at Houston, Houston, Texas

Panagiotis G. Simos

Department of Psychology, University of Crete, Rethymno, Greece

Jack M. Fletcher and Paul T. Cirino

Department of Psychology, University of Houston, Houston, Texas

Sharon Vaughn

The Meadows Center for Preventing Educational Risk, University of Texas at Austin, Austin, Texas

Andrew C. Papanicolaou

Department of Pediatrics, Children’s Learning Institute, University of Texas Health Science Center at Houston, Houston, Texas

Brain activation profiles obtained using magnetoencephalography were compared between middle-school students experiencing reading difficulties and non-reading-impaired students during performance of a continuous printed word recognition task. Struggling readers underwent small-group remedial instruction, and students who showed significant gains in word reading efficiency at a one-year follow-up assessment were classified as Adequate Responders whereas those not demonstrating such gains as Inadequate Responders. At baseline, compared to Inadequate Responders, the activation profiles of Adequate Responders featured increased activity in the left middle, superior temporal, and ventral occipitotemporal regions, as well as in the right mesial temporal cortex. The degree of activity in these regions was a significant predictor of improvement in word reading efficiency beyond the prediction afforded by baseline reading accuracy or fluency measures. The engagement of brain areas that typically serve as key components of the brain circuit for reading

This research was supported in part by grant P50 HD052117 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD or the National Institutes of Health.

Correspondence should be addressed to Roozbeh Rezaie, Center for Clinical Neurosciences, Division of Children’s Learning Institute, University of Texas Medical School, Department of Pediatrics, 1333 Moursund Street, Suite H 114, Houston, TX 77030. E-mail: Roozbeh.Rezaie@uth.tmc.edu
may be an important factor in predicting response to intervention in older students who experience reading difficulties.

With the advent of functional neuroimaging, the neural correlates of proficient reading ability have been documented across multiple modalities, including positron emission tomography (PET; Rumsey et al., 1992, 1997), functional magnetic resonance imaging (fMRI; Hoefn et al., 2007; Paulesu et al., 2001; Shaywitz et al., 2002), and magnetoencephalography or magnetic source imaging (MSI; Simos et al., 2001). Although there are inconsistent results across studies that may reflect task, participant, and modality differences (Poeppel, 1996; Price & McCrory, 2005), there is also convergence in identifying the principal features of a neural network supporting reading. These features include activation of ventral occipitotemporal regions, the posterior temporoparietal region, and the inferior frontal lobe, predominantly in the left hemisphere (Eden & Zeffiro, 1998; Gabrieli, 2009; Papanicolaou et al., 2004; Rumsey et al., 1997; Shaywitz et al., 2000).

In struggling readers, there is evidence for underactivation of the left posterior temporoparietal, left inferior parietal, and left occipitotemporal areas, as well as possible hyperactivation of the left inferior frontal regions. Two recent meta-analyses of PET and fMRI studies support these interpretations (Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Richlan, Kronbicher, & Wimmer, 2009; see also Gabrieli, 2009). The review by Maisog and colleagues (2008) was restricted to adults identified with dyslexia and controls, reporting reduced activation in left hemisphere regions involving ventral occipitotemporal, temporoparietal (inferior parietal cortex, superior temporal gyrus), thalamus, and inferior frontal gyrus. In the right hemisphere, reduced activation was apparent in fusiform, postcentral, and superior temporal gyri. In their review Richlan, Kronbicher, and Wimmer (2009) included studies on children with dyslexia and found reduced activation in left hemisphere regions involving the temporoparietal (inferior parietal, superior temporal, middle and inferior temporal regions), and occipitotemporal regions. Furthermore, underactivation of the inferior frontal gyrus and hyperactivity of the primary motor cortex and anterior insula were also found. Thus, there is convergence in results for the left posterior superior temporal, inferior parietal, and occipitotemporal regions, and more inconsistency in frontal regions and right hemisphere regions which may assume a compensatory role into the brain circuit for reading in struggling readers.

More recent studies have addressed the neural correlates of intervention response using fMRI and MSI (Aylward et al., 2003; Eden et al., 2004; Odegard, Ring, Smith, Biggan, & Black, 2008; Meyler, Keller, Cherkassky, Gabrieli, & Jus, 2008; Richards et al., 2000, 2002, 2006; Shaywitz et al., 2004; Simos et al., 2002, 2005, 2007a, 2007b; Temple et al., 2003). Across modalities, these studies suggest that effective intervention results in increased activation, or functional normalization, of the left temporoparietal region, with less convergence regarding changes in the occipitotemporal and frontal regions. With the exception of Eden et al. (2004), which included adults, all of these studies have involved children in elementary school. In addition, as Gabrieli (2009) noted, these studies have largely not addressed possible differences in brains of children who respond adequately and inadequately to reading interventions. Of the few available studies making this comparison, a MSI study by Simos and colleagues (2007b) found no changes in brain activity for three children who did not respond adequately to reading intervention. Odegard...
and colleagues (2008) conducted a post intervention comparison of a small number of adequate and inadequate responders and found increased fMRI activation in the left inferior parietal region for responders to the reading intervention. Davis et al. (2011) used fMRI to compare five adequate and five inadequate responders to a reading intervention. Responders showed increased activation of the left hemisphere posterior superior temporal and middle temporal gyri. Thus, all these studies highlight increased activation of the posterior temporoparietal region of the brain circuit supporting reading.

In the present study, we specifically compare activation profiles of adequate and inadequate responders obtained prior to intervention. In addition, we conduct this comparison in middle school children who are largely unstudied in the functional neuroimaging literature. Finally, following the findings of Hoeft et al. (2007), who reported that fMRI activation data predicted reading development beyond information by reading and cognitive tasks, we investigated whether signs of adequate engagement of the neurophysiological processes normally supporting fluent word recognition predict individual capacity to benefit from remedial instruction among students with diverse instructional needs. The study is part of a larger educational intervention project recruiting urban, middle-school students who experience reading difficulties (RD) at the word and/or text level (Vaughn et al., 2010; 2011).

The present study explores features of brain activation during performance of a continuous word recognition task in 20 struggling readers recruited from a larger grade 6–8 reading intervention study (Vaughn et al., 2010; 2011). Following their baseline MEG scans, these children received group remedial instruction focusing on reading fluency, vocabulary and reading comprehension skills, with some responding adequately and others inadequately to the intervention. Neuromagnetic recordings were obtained prior to intervention, and analyzed using a distributed source-modeling approach to derive estimates of the degree and temporal course of task-related regional activation. Individual differences on baseline spatiotemporal activation profiles were then examined as a function of the students’ degree of response to educational remediation which was determined based on one-year follow-up reading assessments. We hypothesized that the activation profiles of students who later demonstrated positive response to intervention would reveal signs of adequate engagement primarily of the left middle temporal and ventral occipitotemporal regions in the left hemisphere, resembling those of typically achieving readers. In contrast, pre-intervention profiles for inadequate responders would closely resemble those typically observed in previous imaging studies in children with RD. In addition, it was predicted that, mainly, degree of activity in these regions would correlate strongly with the amount of gains made by individual students post-intervention. Given the nature of the task, we predicted that these associations would be stronger for achievement measures which mainly assess the availability and retrieval capacity of printed lexical representations of words (sight word reading fluency as compared to pseudoword reading fluency tests).

**METHODS**

 Participants

The sample included 20 adolescent struggling readers who did not meet criteria on the Texas Assessment of Knowledge and Skills (TAKS), a criterion referenced reading comprehension
assessment that represents the state accountability test. These adolescents were volunteers who met the selection criteria specified below from the treatment groups in an intervention study designed for students in grades 6–8 (Vaughn et al., 2010; 2011). Although all of the struggling readers were invited and actively recruited to participate through letters to parents and follow-up phone calls, response rates for imaging studies in school-identified samples are typically very small (e.g., Meyler et al., 2008; Simos et al., 2007b). For the present study, we also required scores below the 25th percentile (standard score of 90) on the Test of Word Reading Efficiency (TOWRE) Sight Word scale (Form A, Torgesen, Wagner, & Rashotte, 1999) because the TAKS cannot be used as a pre-post measure and relatively few of these children had reading difficulties restricted to reading comprehension. This measure was adopted for its close similarity to the activation task used in the present study with respect to the component operations involved.

Individual differences in response to intervention were assessed on a group basis using changes in TOWRE Sight Word reading scores between baseline and a one-year follow-up visit representing one-half of a standard deviation (a 7-point difference cutoff). Ten children were classified as Adequate Responders (AR group), demonstrating a mean improvement on the TOWRE Word Reading Efficiency of 11 ± 3 points (range: +7 to +14 points), and ten as Inadequate Responders (IR), improving on average by only 0.7 ± 5 points (range: –3 to +3 points). Other studies of elementary school children have demonstrated that the TOWRE reliably detects changes in responder status (Fletcher et al., 2011). At follow-up, three students in the AR group still read at or below the 25th percentile on the TOWRE Sight Word reading scale, but all demonstrated significant changes in word fluency performance. The majority of AR students also showed strong gains on pseudoword reading fluency as indicated by difference scores on the corresponding TOWRE subtest averaging points. Number of points gained on this subtest ranged between 1 and 21 for this group, and only one student still scored below the 25th percentile at follow-up. While at baseline the IR group actually performed significantly higher than the AR group on both TOWRE subtests, this trend was reversed at follow-up.

A third group of 20 children who had never experienced difficulties in reading (NI group) served as comparisons, having standard scores >92 on the TOWRE composite index. Table 1 displays demographic and psychoeducational information for each of the three groups of participants, which were comparable on age and handedness. On average, NI students scored higher on the IQ measures than each of the two RD groups, although these differences did not reach statistical significance (p > .05). There was also a non-significant tendency for a higher proportion of boys than girls in the NI group, and a significantly higher proportion of minority students in both RD groups as compared to the NI group. As expected, both RD groups scored significantly lower than the NI group on other measures of reading and spelling at baseline, but did not differ from each other on these measures.

The presence of symptoms indicative of attention deficit hyperactivity disorder (ADHD) was assessed via the Child Behavior Checklist Parent Form (CBCL; Achenbach, 1991) or the Inattention and Hyperactivity-Impulsivity scales of the teacher-completed Strengths and Weaknesses of ADHD symptoms and Normal Behavior (SWAN; Swanson et al., 2005). All participants had T scores <55 on the former scale or a mean score lower than 1.67 on the latter scale indicating low risk for ADHD (Chen, Faraone, Biederman, & Tsuang, 1994). In addition to having scores on the CBCL or SWAN below the clinically significant range, none of the participants had a clinical diagnosis of ADHD.
TABLE 1
Demographic Data and Performance on Standardized Tests for the Three Groups of Participants (mean ± SD)

<table>
<thead>
<tr>
<th></th>
<th>NI (n = 20)</th>
<th>AR (n = 10)</th>
<th>IR (n = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (b/g)¹</td>
<td>13/7</td>
<td>6/4</td>
<td>5/5</td>
</tr>
<tr>
<td>Age (mo.)</td>
<td>151 ± 11</td>
<td>158 ± 7</td>
<td>153 ± 11</td>
</tr>
<tr>
<td>Ethnicity (Caucasian/Hispanic/African American)³</td>
<td>10/3/7</td>
<td>0/4/6</td>
<td>0/7/3</td>
</tr>
<tr>
<td>Hand. (R/L)³</td>
<td>18/2</td>
<td>9/1</td>
<td>9/1</td>
</tr>
<tr>
<td>VIQ</td>
<td>103 ± 12</td>
<td>93 ± 14</td>
<td>95 ± 13</td>
</tr>
<tr>
<td>PIQ</td>
<td>102 ± 12</td>
<td>101 ± 16</td>
<td>93 ± 14</td>
</tr>
<tr>
<td>FSIQ</td>
<td>103 ± 9</td>
<td>96 ± 12</td>
<td>93 ± 11</td>
</tr>
<tr>
<td>WJ–III RC</td>
<td>105 ± 6 †§</td>
<td>83 ± 9 †</td>
<td>84 ± 8 †§</td>
</tr>
<tr>
<td>WJ–III WA</td>
<td>106 ± 11 †§</td>
<td>86 ± 7 †</td>
<td>87 ± 6 †§</td>
</tr>
<tr>
<td>WJ–III LWID</td>
<td>103 ± 10 †§</td>
<td>81 ± 12 †</td>
<td>82 ± 9 †§</td>
</tr>
<tr>
<td>WJ–III Spelling</td>
<td>110 ± 11 †§</td>
<td>83 ± 16 †</td>
<td>79 ± 10 †§</td>
</tr>
<tr>
<td>TOWRE–Words-Time 1</td>
<td>98 ± 2 †#@</td>
<td>82 ± 9 †#</td>
<td>86 ± 4 †#@</td>
</tr>
<tr>
<td>Time 2</td>
<td>—</td>
<td>91 ± 3</td>
<td>86 ± 5</td>
</tr>
<tr>
<td>TOWRE–Pseudowords-Time 1</td>
<td>97 ± 3 †#@</td>
<td>81 ± 5 †#e</td>
<td>86 ± 4 †#e</td>
</tr>
<tr>
<td>Time 2</td>
<td>—</td>
<td>90 ± 7</td>
<td>84 ± 5</td>
</tr>
</tbody>
</table>

¹Φ = .22, p = .3; ²Φ = .66, p < .002; ³Φ = .09, p > .75, ⁴Φ = .001, ⁵Φ < .01.
WJ–III RC, WJ–III WA, and WJ–III LWID refer to the Woodcock-Johnson III, Reading Comprehension, Word Attack, and Letter-Word Identification subtests, respectively. TOWRE Words and TOWRE Pseudowords refer to the Test of Word Reading Efficiency Sight Word and Pseudoword subtests, respectively. Time 1 and Time 2 indicate assessments at baseline (concurrent with MEG testing) and at the one year follow up (post intervention). AR = Adequate Responders; FSIQ = Full-scale IQ; IR = Inadequate Responders; MEG = magnetoencephalography; NI = Non-impaired; VIQ = Verbal IQ; PIQ = Performance IQ.

Procedures

Intervention. Students were randomly assigned to receive intervention in either a large group of 10–15 students or a small group of about 5 students. Since both groups received the same intervention and the results did not indicate an effect of group size (Vaughn et al., 2010), we collapsed across these interventions for this study to maximize the sample size. For the interventions, the instruction took place for 45–50 minutes per day (regular class period) throughout the school year (September to May). Three phases of instruction composed the yearlong intervention. In Phase I (approximately 7–8 weeks), word study and fluency were emphasized, with additional instruction provided in vocabulary and comprehension. Phase II (approximately 17–18 weeks) emphasized vocabulary and comprehension, with additional instruction and practice in the application of the word study and fluency skills and strategies learned in Phase I. Phase III (8–10 weeks) continued the instructional emphasis on vocabulary and comprehension, with more time spent on independent student application of skills and strategies. The results of the study demonstrated relatively small effects compared to a business as usual comparison group (Vaughn et al., 2010; 2011), which is consistent with most large scale studies of secondary students (Vaughn &
Fletcher, in press). However, many students improved relative to the baseline, which is the focus of the present study.

**Imaging task.** Each participant was tested on a printed word recognition task, adapted from Simos et al. (2000). Changes in the selection of the stimuli and the number of targets were made to reduce overall task difficulty and memory load. The task was selected in order to tap into the skills of rapid printed word recognition, learning, and discrimination, which were explicitly taught as part of the intervention program. There were only five target stimuli which were repeated in a different random order, mixed with a different set of 30 distractors (non-repeating words) in each of four blocks of stimuli (for a total of 140 stimuli). Target words (*age, enjoy, hope, jar,* and *road*) included four monosyllabic and one disyllabic word, and had a mean frequency in the Zeno et al. G6-7 corpus (Zeno, Ivens, Millard, & Duvvuri, 1995) of 158 occurrences per million (range: 32-194 occurrences). A slightly higher proportion of distractors were disyllabic (40%) and the remaining monosyllabic, with a mean frequency of occurrence of 150 words per million in the same corpus (range: 18–820). Stimuli were presented for 1 sec, one at a time (with a randomly varied interstimulus interval of 3–4 sec), through a Sony LCD projector (Model VPL-PX21) on a back-projection screen located approximately 60 cm in front of the participant, and subtended 1.0°–3.0° and 0.5° of horizontal and vertical visual angle, respectively. Participants were presented the target words prior to the scan and instructed to read them aloud (to ensure that they could) and to “try to remember the words.” During the scan they were presented with the target stimuli mixed with the distractors and asked to raise their index finger “as soon as they saw one of the words they had read before the scan.” Responding hand was counterbalanced across blocks of trials and responses were manually recorded by the experimenter. Scan duration was approximately 15 min (including short breaks between blocks of stimuli).

**Imaging procedures.** MEG recordings were obtained with a whole-head neuromagnetometer array (4-D Neuroimaging, Magnes WH3600), consisting of 248 first-order axial gradiometer coils. The magnetic flux measurements were digitized at 250 Hz, filtered with a bandpass filter between 0.1 and 20 Hz and subjected to baseline adjustment (using the 150 ms prestimulus recording) and to a noise reduction algorithm. The single-trial event-related field segments (ERFs) in response to 80–100 stimulus presentations, were averaged after excluding those containing eye movement or other myogenic or mechanical artifacts.

To identify the intracranial origin of ERFs, the magnetic flux distribution at successive points (4 msec apart) was analyzed using a minimum norm model to obtain estimates of the time-varying strength of intracranial currents (MNE Software, v. 2.5, http://www.nmr.mgh.harvard.edu/martinos/userInfo/data/sofMNE.php; Hämäläinen & Ilmoniemi, 1994). Estimated current sources were anatomically constrained by an MRI-derived surface model of each participant’s brain using FreeSurfer software (Dale, Fischl, & Sereno, 1999). Minimum norm estimates attempt to reconstruct the intracranial sources of activity by identifying the smallest distribution of dipoles (e.g., minimum norm) that can account for the magnetic flux distribution recorded simultaneously over the entire head surface at successive time points. In contrast to the spherical head model employed by single equivalent current dipole approach, the MNE technique affords greater spatial resolution by assuming a continuous distribution of dipolar sources along the cortical surface, which are anatomically constrained using a realistic model of the head constructed from each participants’ high-resolution MRI. Solving the inverse problem using the MNE method initially required the construction of a cortical surface model based on individual
brain anatomy. Specifically, the surface model was created using automated extraction techniques that generate a detailed geometric description of the gray–white-matter boundary of the neocortical mantle. For each of the cerebral hemispheres, a regular tessellation of the cortical surface consisting of approximately 150,000 (depending on the individuals’ cortical surface area) equilateral triangles known as vertices was created. Actual estimation of the activity sources was derived by defining a solution source space, using a grid-spacing of several millimeters, to model each vertex (cortical patch) as a potential current dipole perpendicular to the cortical surface. The inverse solution was subsequently reduced to obtaining an estimate of the scalar distribution of dipole strength across current sources within orientation-specific cortical patches of approximately 3,000 vertices covering the entire cortical surface (Dale et al., 1999). Co-registration of each MEG dataset with its corresponding MRI dataset was performed using an automated co-registration routine within the MNE Software suite.

Each region of interest (ROI) was selected from an atlas of regions that were generated during the head model construction phase using the method of Desikan and colleagues (Desikan et al., 2006), in FreeSurfer software (Dale et al., 1999). In summary, participants’ MRI scans were initially pre-processed using a fully automated approach which, following correction for magnetic field inhomogeneity, was used to skull-strip and segment the brain into gray and white matter. The gray matter segmentations for each subject were “inflated” to optimize visualization of anatomy across the cortical surface, by taking into account individual sulcal and gyral complexities. Subsequently, inflated gray matter segmentations were deformed against a standard template image consisting of 34 ROIs (per cerebral hemisphere), the boundaries for which were previously defined using manual delineation on a set of 40 MRI scans (see Desikan et al., 2004 for details). Previous MEG studies using Equivalent Current Dipole modeling of magnetic activity as well as MNE (Rezaie et al., in press; Simos et al., 2011), using an identical anatomical labeling system, indicated that ROIs likely to show activity that varies systematically as a function of reading ability include the following areas in both hemispheres: superior (STG; BA 22) and middle temporal gyri (MTG; BA 21) excluding cortex along the banks of the superior temporal sulcus (STS); the cortical surface within STS; supramarginal gyrus (SMG; BA 40); angular gyrus (ANG; BA 39); pars opercularis of the inferior frontal gyrus (IFG; BA 44); rostral inferior (BA 47) and rostral middle frontal cortices (BA 46/9); motor cortex (BA 4); fusiform gyrus (BA 37); and lateral occipito-temporal cortex (BA 19). In addition, activity in mesial temporal cortex (parahippocampal region; MTL) was examined given the episodic memory demands posed by the activation task. The program outputs a current estimate value for each voxel and each 4 msec time point. This value is then used to compute the two dependent measures used in the analyses outlined in the following paragraphs: (1) the average current across all voxels defining each of the ROIs listed above and across all of the 4 msec time points comprising 14 successive 50 msec time bins (100–150, 150–200 msec, and so on up to 800 msec); and (2) the latency (in msec after stimulus onset) when the averaged current in a given ROI reached peak amplitude within the entire recording epoch.

Analytic approach. In order to identify potential predictors of intervention response, the average current for each 50 msec time bin and each ROI, obtained prior to intervention, were initially submitted to an ANOVA with Area (11), Hemisphere (2), and Time bin (14) as the within subjects variables and Group (3) as the between subjects variable. Significant four-way interactions involving Group were further evaluated by examining three-way (e.g.,
Time × Hemisphere × Group in each ROI) or two way interactions (e.g., Time × Group) which, if significant, were explored by testing simple main effects of Group. All ANOVA results were evaluated using the Huynh-Feldt method as a precaution against inhomogeneity of variance problems.

To address the (longitudinal) predictive value of the brain activation measures, a second set of analyses consisted of partial correlations between reading fluency difference scores (the difference between baseline and follow up on TOWRE Sight Word Reading and TOWRE Pseudoword Reading standard scores) and peak latency/average current in each ROI (and time bins).

RESULTS

In-Scanner Task Performance

The three groups performed at comparable levels at baseline on the in-scanner printed word reading task as indicated by the absence of a significant Group main effect (controlling for age) on the word recognition task \(p > .7\). Performance was slightly higher for the NI (\(M = 83.6, SD = 12\%\)) as compared to the AR (\(M = 81.0, SD = 12\%\)) and IR groups (\(M = 78.1, SD = 16\%\)).

MEG Data-Degree of Activity

The omnibus ANOVA revealed a significant Time × Area × Group interaction, \(F(234, 4212) = 1.51, p < .01\). These effects were further explored by performing three-way ANOVAs (Time × Hemisphere × Group) separately for each ROI. As hypothesized, the IR group displayed reduced degree of activity relative to both AR and NI groups in three key areas of the brain circuit for reading: the MTG, STG, and the ventral occipitotemporal (VOT) region, primarily in the left hemisphere. In addition, significantly reduced activity in the IR group was found in the right parahippocampal gyrus. Results are described in more detail below.

Group differences in MTG, STG, and the VOT region varied between hemispheres as indicated by Hemisphere × Time × Group interactions, \(F(26,507) = 2.20, p < .01, F(26,507) = 2.55, p < .001, F(26,507) = 2.90, p < .0001\), respectively.

In MTG, follow-up ANOVAs revealed that the Group × Hemisphere interaction was restricted to the 400–500 msec latency range \(p < .025\). Two way ANOVAs performed separately at each hemisphere, indicated that group differences were significant only in the left hemisphere in this latency range, \(F(2,39) = 7.81, p < .001\). Bonferroni-corrected pairwise comparisons showed that the degree of activity in the left MTG was higher for the NI group than for the IR group \(p < .001\). A trend for higher degree of activity in the AR group relative to the IR group did not reach significance (adjusted \(p < 0.1\)).

A similar pattern of results was noted for STG, where the Group × Hemisphere interaction was restricted to the 300–500 msec latency range \(p < .03\), and the Group main effect was also restricted to the left hemisphere, \(F(2,39) = 5.41, p < .008\). Pairwise comparisons showed that the degree of activity in the left STG was higher for the NI group than for IR group \(p < .008\).

The pattern of group differences was also similar in the VOT region, although group effects were noted much earlier in the neuromagnetic response (the Group × Hemisphere interaction was restricted to the 100–250 msec latency range; \(p < .03\)). As in the case of MTG and STG,
the group main effect was restricted to the left hemisphere \((p < .005)\), reflecting higher degree of activity for the AR than the IR group (Bonferroni-corrected \(p < .003\)). Group differences did not reach significance in the right hemisphere in any of the three ROIs \((p > .3\) in all cases). Significant hemisphere asymmetries \((L > R)\) were found for MTG and STG in the latency ranges where significant group effects were also found in each ROI, only for the NI \((p < .02)\) and IR groups \((p < .03)\).

**Peak latency.** No significant differences in peak latency were found between the three groups (see Figures 1 and 2).

### Correlations Between Neurophysiological Activity and Reading Ability

In order to preserve the temporal information that is inherent in the MNE time series data, Pearson correlation coefficients were computed between estimates of the magnitude of regional neurophysiological activity for each time window in each of the 11 ROIs and achievement standard scores (TOWRE Sight Word reading efficiency and TOWRE Pseudoword reading efficiency). In particular, correlations were computed for changes in standard scores between baseline and one-year follow-up, for each TOWRE subscale separately, in order to address process-specific associations between degree of regional activity and achievement-response to intervention. It was predicted that given the nature of the activation task, which places increased demands for rapid word recognition, regional activation would mainly predict post-intervention gain on word reading fluency. Given the relatively small group size, these analyses were conducted on the entire sample of RD students regardless of RTI status \((N = 20)\).

Table 2 reports significant coefficients \((p < .02)\). Results of the correlational analyses were generally consistent with the findings of group differences in the degree of activity. Specifically, degree of activity in two left hemisphere regions (MTG, VOT) and degree of activity in the right parahippocampal gyrus emerged as significant, positive predictors of improvement in word reading efficiency.

Subsequently, each of the MNE measures associated with significant correlations listed above for the entire subgroup of RD students, were entered into hierarchical linear multiple regression analyses as predictors of the degree of change in TOWRE word reading efficiency scores (excluding variables which showed high degree of collinearity). Degree of activity measures accounted for 40% of the variance in TOWRE Word Reading difference scores \((\text{Adj } R^2 = .40, F(3,16) = 5.28, p < .01)\). The optimal predictor model included three independent variables: left hemisphere MTG activity (between 450 and 500 ms, \(\beta = .19\)), left hemisphere VOT (150–200 msec, \(\beta = .26\)) and right hemisphere MTL (550–600 msec, \(\beta = .46\)). The regression plots in Figure 3 show a rather linear relationship between each of the three MNE measures and TOWRE word reading efficiency difference scores across RD subgroups. However, none of the degree of activity measures was associated with significant regression coefficients in this group, probably due to considerable shared variance among them.

In additional analyses, baseline reading accuracy (WJ–III Word Attack, WJ–III Word Identification) or fluency standard scores (TOWRE Word Reading Efficiency and TOWRE Pseudoword Reading Efficiency) were entered in the first step of hierarchical linear regression models. Only reading fluency measures correlated significantly with the degree of change in word
### TABLE 2

Pearson Correlation Coefficients Between Degree of Activity in Three Regions of Interest (ROIs) and Amount of Longitudinal Change in TOWRE Sight Word Reading Efficiency Standard Scores

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MTG (LH)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOT (LH)</td>
<td>.54</td>
<td>.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTL (RH)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.60</td>
<td>.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only significant coefficients with Test of Word Reading Efficiency (TOWRE) Word reading efficiency are shown ($r > .50, p < .02$; there were no significant correlations with degree of activity in STG). Corresponding coefficients with TOWRE Pseudoword Reading Efficiency were $< .2$ in all cases. Data for the entire group of reading difficulties students ($N = 20$) are displayed. MTG = middle temporal; MTL = medial temporal lobe; STG = superior temporal; VOT = ventral occipitotemporal.
FIGURE 1 Time course of estimated neurophysiological activity (in nanoAmpere meters) associated with word recognition in the middle temporal (MTG), superior temporal (STG), and ventral occipitotemporal (VOT) regions of interest (ROI) in the left (LH) and right hemispheres (RH). Stimulus onset was at 0 msec (displayed activity starts at 100 msec post-stimulus onset). Significant group differences were only found in the left hemisphere during specific latency windows (bound by rectangles).
FIGURE 2  Brain activation map snapshots for the word recognition task from three representative participants: a typically developing reader (left-hand pair of columns), a student who later showed adequate response to intervention (middle pair of columns), and a student who did not show adequate response (right-hand pair of columns). When these recordings were made, this Adequate Responder student scored 79 points on the Test of Word Reading Efficiency (TOWRE). Sight Word reading efficiency subtest, and 88 points when tested again one year later. Corresponding scores for the Inadequate Responders (IR) student shown here were 84 and 81 points, respectively. The relative intensity of activated voxels is shown at the bottom of the figure. Each set of images presents activity in regions of interest (ROI) (demarcated by green tracings) where significant group differences were observed, at a different latency after stimulus onset (at 100–250, 400–500, and 300–500 msec, from top to bottom, respectively), in the left (LH) and right (RH) hemispheres. (color figure available online)

reading efficiency (Pearson $r = -0.54$, $p < 0.005$ for TOWRE words and $r = -0.56$, $p < 0.005$ for TOWRE pseudowords). In both cases, however, the combined neuromagnetic measures introduced in the second step of the analysis, resulted in a significant increase in the amount of variance accounted for by the changes in TOWRE difference scores. In the model including
FIGURE 3 Regression plots of Test of Word Reading Efficiency (TOWRE) Sight Word Reading Efficiency standard scores (difference between baseline and one-year follow-up) over degree of time-dependent neurophysiological activity in left hemisphere middle temporal (MTG) and ventral occipitotemporal (VOT) and degree of right hemisphere medial temporal lobe (MTL) activation. Higher levels of activity at baseline predicted greater improvement in reading efficiency for the entire group of struggling readers, although a stronger association was noted for students who showed adequate response to intervention (solid circles).

TOWRE word reading efficiency standard score in the first step, the final Adj $R^2$ was .49, $R_{change}^2 = .31$, $F_{change}(3,15) = 3.83$, $p < .03$. Corresponding model fit results when TOWRE word reading efficiency standard score was entered into the model first were: Adj $R^2 = .58$, $R_{change}^2 = .36$, $F_{change}(3,15) = 5.32$, $p < .01$. 
DISCUSSION

The objective of the current study was to investigate the neural correlates of intervention response in adolescent struggling readers, initially identified due to their failure on a state-wide reading comprehension assessment, accompanied by poor reading fluency and comprehension on standardized achievement tests. Distributed source modeling of MEG data were obtained during performance of a continuous printed word recognition task. Compared to non-reading impaired students, roughly comparable in age and IQ, we observed marked underactivation of the MTG, STG and VOT among struggling readers who failed to show adequate response to reading intervention. Moreover, brain activation profiles of AR students prior to intervention featured increased activity in the left MTG, STG, and VOT, as well as in the right mesial temporal cortex in comparison to Inadequate Responders to intervention. The degree of activity in these regions was a significant predictor of students’ subsequent improvement in word reading efficiency, beyond the prediction afforded by reading accuracy or fluency measures at baseline. The higher the degree of activity in these regions at particular latency windows (characterizing early stages of word processing in VOT and late stages in MTG and STG), the greater the subsequent gains in word reading automaticity. Some unique contributions of the present data to addressing research questions regarding the neurobiological correlates of response to intervention include the following:

Sample Characteristics

Although relatively small, the present sample features broad diversity in terms of: (a) ethnic background and socioeconomic status (SES), and (b) achievement and general ability profiles (students were initially selected if they had failed a state-wide test focusing on reading comprehension). These rather unique features of the present sample, in relation to earlier imaging studies, support the generalizability of the findings.

Participant Age

The majority of previous brain imaging-based intervention studies included either younger samples (students in primary grades, e.g., Hoeft et al., 2007; Odegard et al., 2008; Shaywitz et al., 2004; Simos et al., 2005, 2007a; Temple et al., 2003), or adults with a history of reading disability (Eden et al., 2004). Generally, these studies have shown that completion of an intensive instruction program was accompanied by increased neurophysiological activity in temporo-parietal cortices in the left hemisphere. Moreover, several of these studies (Odegard et al., 2008; Simos et al., 2007b) have demonstrated that subsequent inadequate responders to intervention exhibit profiles typical of RD children at baseline, relative to adequate responders. The present study extends these findings by including adolescent struggling readers attending middle school. The results across studies, however, are similar in observing increased activation in the left temporal parietal regions of the brain circuit supporting reading. The differences in which components of the temporal parietal regions are engaged may reflect sample, intervention, modality, and task differences.
Activation Profiles and Systematic Inter-Subject Variability in the Degree of Regional Activity Demonstrated Task- and Process-Specificity

We found that the degree of activity in brain regions that are normally critical for word recognition (see detailed discussion below) when engaged in a task that poses such demands, correlated strongly with improvement in word reading fluency measured outside the scanner. Conversely, degree of activity in these regions at baseline failed to map onto individual differences in gains on decoding automaticity (pseudoword reading fluency). These differences are not attributed to limited variance on scores on the latter measure, as indicated by inspection of the distribution of these scores. This simple dissociation is complemented by the results of analyses of baseline MEG data obtained in the context of a pseudoword naming task, posing minimal demands for word recognition while maximizing sublexical processing load. Utilizing a partially overlapping sample of struggling readers ($N = 27$) with the present sample, we found that the degree of activity in superior temporal and inferior parietal regions (known to be normally involved in phonological decoding processes) at baseline, predicted amount of individual improvement in pseudoword reading fluency at the 1-year follow-up (Rezaie et al., unpublished data).

The increased engagement of the left MTG and STG in AR relative to IR prior to intervention has important implications for assessing individual capacity to benefit from remedial instruction among students with diverse instructional needs, given the importance of these regions in the development of fluent word recognition. In particular, the role of the posterior portion of MTG (BA 21) in lexical/semantic processing has been supported by numerous studies (Booth et al., 2002; Damasio & Damasio, 1983; Fiebach, Friederici, Müller, & von Cramon, 2002; Gaillard et al., 2001; Halgren et al., 2002; McCandliss, Cohen, & Dehaene, 2003; Pugh et al., 1996; Simos et al., 2002; Turkeltaub, Eden, Jones, & Zeffiro, 2002; Tyler, Marslen-Wilson, & Stamatakis, 2005; Wehner, Ahlfors, & Mody, 2007). Among the functions hosted by this region may be storage and access to the to lexical/semantic information for printed stimuli (Fiebach et al., 2002; Pugh et al., 1996), as demonstrated by lesion studies that have shown that damage to this region underlies deficits in a variety of tasks tapping into lexical/semantic processing, including semantic priming (Tyler et al., 2005). Similarly, while much of the empirical evidence indicates the cortex lying on the lateral and superior aspects of the STG (BA 22) hosts neurophysiological processes supporting phonological processing of spoken and written language (Beauvois & Derouesne, 1979; Caplan, Gow, & Makris, 1995; Majerus et al., 2005; Papanicolaou et al., 2003; Specht et al., 2003), imaging and lesion studies have also implicated this region in lexical/semantic processing of spoken language. For example, studies of experienced adult readers have demonstrated the involvement of the posterior portion of BA 22 in lexical/semantic processing during silent reading tasks (Cannestra et al., 2000; Fiebach et al., 2002; Fiez, Balota, Raichle, & Petersen, 1999; Haist et al., 2001; Okada & Hickok, 2006; Simos et al., 2009). Moreover, patients with a clinical profile of fluent aphasia, often resulting from damage to BA22 (Benson, 1985; Vignolo, 1988), consistently exhibit deficits in a numerous experimental tasks drawing on lexical/semantic processing (Hagoort, 1993; Janse, 2006; Rodd, Davis, & Johnsrude, 2005).

In addition to MTG and STG, increased strength of activation in AR, relative to IR, at baseline, was also found in VOT (BA 37, anatomically corresponding to the fusiform gyrus). Crucially, this region is believed to integrate orthographic, phonological, and morphological information of printed stimuli, as well as to host neurophysiological processes responsible for graphemic processing (Flowers et al., 2004; McCandliss et al., 2003; Pammer et al., 2004; Vigneau, Jobard,
Mazoyer, & Tzourio-Mazoyer, 2005). Furthermore, a growing body of evidence indicates that the VOT is involved in the storage of, and access to, familiar orthographic representations, which may take the shape of “word forms” (Kronbichler et al., 2004; Simos et al., 2009), and play an important role in both word recognition and spelling (Mani et al., 2008; Philipose et al., 2007). Thus, one possible interpretation of the increase in VOT activity among AR students is that processing and retrieval of stored orthographic representations of target words is enhanced in these individuals.

The importance of differences in the engagement of MTG, STG, and VOT between AR and IR is substantiated by significant associations between achievement measures, which mainly assess the availability and retrieval capacity of printed lexical representations of words, and the degree of activity in these regions. Rezaie et al. (unpublished data) employed an identical continuous word recognition task in a large, representative sample of typical and struggling readers. They reported that the degree of activation in the left MTG and STG, as well as the VOT, were significant correlates of word reading and spelling achievement scores among NI children, particularly during late processing of the printed words. Conversely, this latter study also found that in struggling reading, enhanced activity in corresponding right hemisphere regions was associated with lower achievement scores, a trend unaffected by age, being evident across the age span of the study sample (7 to 15 years).

Finally, increased activity observed in the MTL may reflect the coordination required for the storage and retrieval of active traces for target words which are normally stored in MTG, STG, and VOT. Indeed, there is evidence from lesion and functional imaging studies highlighting the critical role of the mesial temporal cortices (Brown & Aggleton, 2001; Turriziani, Fadda, Caltagirone, & Carlesimo, 2004; Wixted & Squire, 2004) for recognition of previously encountered stimuli.

Depending on task demands, different sets of brain areas may become predominantly active during reading of particular stimuli in the scanner. However, it is unlikely that the entire set of brain areas that comprise the brain circuit for reading are active, and in the same temporal sequence, during decoding of unfamiliar letter strings, recognition of words in isolation, and reading words in context. Accordingly, multiple measures of brain activity, in various regions and latencies for the same task, and the same region—at the same or different latencies—across tasks, are necessary in order to optimize the estimated likelihood of successful interventions at the level of the brain. Naturally, factors pertaining to the features of the intervention itself (content, structure, group size, etc.), as well as individual differences in cognitive abilities (working memory, problem-solving ability, morphosyntactic and lexical knowledge) and psychoemotional variables (motivation, goal orientation; Morgan & Sideridis, 2006) are important determinants of response to intervention. The precise manner in which these variables impact on the function of the brain circuit available to benefit from the intervention should be addressed systematically in future larger scale studies. A much larger sample is needed in order to address the potential impact of individual differences in educational history.

REFERENCES


The timing of regional brain activation associated with word recognition in children with reading difficulties.


