

Published in final edited form as:

*Dev Cogn Neurosci*. 2012 February 15; 2(Suppl 1): S114–S128. doi:10.1016/j.dcn.2011.10.002.

## Differences in neural activation between preterm and full term born adolescents on a sentence comprehension task: Implications for educational accommodations

Laura H.F. Barde<sup>1</sup>, Jason D. Yeatman<sup>2</sup>, Eliana S. Lee<sup>1</sup>, Gary Glover<sup>3</sup>, and Heidi M. Feldman<sup>1</sup>

<sup>1</sup>Department of Pediatrics, Stanford University School of Medicine

<sup>2</sup>Department of Psychology, Stanford University

<sup>3</sup>Departments of Radiology, Stanford University School of Medicine

### Abstract

Adolescent survivors of preterm birth experience persistent functional problems that negatively impact academic outcomes, even when standardized measures of cognition and language suggest normal ability. In this fMRI study, we compared the neural activation supporting auditory sentence comprehension in two groups of adolescents (ages 9–16 years); sentences varied in length and syntactic difficulty. Preterms ( $n = 18$ , mean gestational age 28.8 weeks) and full terms ( $n = 14$ ) had scores on verbal IQ, receptive vocabulary, and receptive language tests that were within or above normal limits and similar between groups. In early and late phases of the trial, we found interactions by group and length; in the late phase, we also found a group by syntactic difficulty interaction. *Post hoc* tests revealed that preterms demonstrated significant activation in the left and right middle frontal gyri as syntactic difficulty increased. ANCOVA showed that the interactions could not be attributed to differences in age, receptive language skill, or reaction time. Results are consistent with the hypothesis that preterm birth modulates brain-behavior relations in sentence comprehension as task demands increase. We suggest preterms' differences in neural processing may indicate a need for educational accommodations, even when formal test scores indicate normal academic achievement.

### Keywords

language; comprehension; prematurity; preterm; fMRI; education

## 1. Introduction

We have long known that extrinsic and intrinsic factors contribute to individual differences in educational achievement. Extrinsic factors include socioeconomic status (Hackman et al., 2010; Kishiyama et al., 2009), size and quality of support systems (Somers et al., 2008), and teacher effectiveness (Stronge et al, 2007). Intrinsic factors include sex (Buchmann et al.,

© 2011 Elsevier Ltd. All rights reserved.

Corresponding author: Laura H.F. Barde, Ph.D., Department of Pediatrics, Division of Neonatal and Developmental Medicine, Stanford University School of Medicine, 750 Welch Road, Suite 315, Palo Alto CA 94304, lbarde@stanford.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

2008; Kraft & Nickel, 1995), motivation (Duckworth et al., 2011; Li-Grining et al., 2010), and intelligence (Deary et al., 2007; Sternberg et al., 2001).

We have only recently begun to learn via neuroscience research that even when academic ability is *equated*, extrinsic or intrinsic factors modulate where cognitive information is processed in the brain. For example, Noble and colleagues (2006) obtained functional magnetic resonance imaging (fMRI) on children from diverse socioeconomic backgrounds (SES) and with equivalent below average phonological skill. They found the children in the lower SES group showed a positive correlation between phonological awareness and activity in the left fusiform gyrus, but children from the higher SES group did not show this correlation. The area of difference is one that is particularly responsive to visually presented words relative to other types of visual or verbal stimuli. Thus, socioeconomic background factors were found to modulate brain-behavior relationships in reading skills. As another example, Burman and colleagues (2008) obtained fMRI on both sexes with high verbal skills during two linguistic tasks, using visual and auditory word presentations. After accounting for differences in accuracy, bilateral activation in the inferior frontal and superior temporal gyri and activation in the left fusiform gyrus was greater in girls than in boys. Activation in the left inferior frontal and fusiform areas of girls was also correlated with linguistic accuracy regardless of stimulus modality, whereas correlation with performance accuracy in boys was found in distinct brain areas when words were presented in the visual or auditory modality. Thus, sex was also found to modulate brain-behavior relationships in language and reading skills. Studies such as these may inform education insofar as effective education may need to tailor teaching strategies, adjust testing methods, or provide other accommodations based on characteristics of the child to improve their level of achievement or performance.

We propose that another important source of intrinsic variation, relevant to academic achievement, is the child's medical history. In this study, we focused on premature birth as an example of a pertinent aspect of a child's medical history because the condition remains prevalent in the U.S. and has long-lasting effects on language and cognitive outcomes. The specific aims of this fMRI study were to (1) compare patterns of neural activation related to auditory language comprehension in preterm and full term born adolescents, particularly when task demands increase, (2) determine if differences in the patterns of activation covary with individual differences in age, offline receptive language skill, or performance on the task, and (3) discuss the implications for educational policy and accommodations for school-age children born prematurely.

### 1.1 Prematurity and long term cognitive outcomes

Prematurity is a condition of utmost public health importance. A normal gestation lasts 40 weeks; each year, over 500,000 children in the U.S. are born before 37 weeks gestation, qualifying as preterm. The health and functional consequences for these children persist far beyond the newborn period. Approximately 20% of survivors born before 32 weeks gestation have a major disability, including sensory and motor impairments and severe developmental delays heralding intellectual disability (Marlow, 2004; Vohr et al., 2000). More importantly, fully 50% of all preterm children have mild to moderate disabilities that are highly pertinent to education.

These mild to moderate disabilities include cognitive, linguistic, and socio-emotional problems that persist into adulthood (Hack, 2009; Lee et al., 2010; Saigal & Doyle, 2008; Saigal & Rosenbaum, 2007). Mean IQ scores of children born preterm can be up to 20 points below those of full term controls matched for sex, SES, and race/ethnicity, even though many survivors score within the normal range (Aylward, 2002; Bhutta, et al., 2002; Hack et al., 2002; Marlow, 2004). Prematurity also contributes significantly to the variance

in scores on language and reading measures after controlling for sex, SES, and IQ (Lee et al., 2010). Particularly relevant to the aims of this paper is the finding that degree of prematurity has shown to be a direct predictor of linguistic processing speed for syntactic comprehension, verbal memory, and reading comprehension (Lee et al., 2010). Other outcomes of prematurity associated with poor educational achievement include weaknesses in executive function skills and high rates of inattention and anxiety symptoms (Aarnoudse-Moens, et al., 2009; Aylward et al., 2002; Loe et al., 2011). Extremely preterm children use special education services at a much higher rate than full term children (Johnson et al., 2009), and while “late preterm” children born between 32 to 36 weeks gestation tend to use less special education services than extremely preterm children, even they experience more grade retention than full term peers. However, even preterm children whose IQ and related test scores are in the normal range and are in mainstream classrooms demonstrate lower academic achievement relative to full term peers (Ross et al., 1991).

Adverse outcomes after prematurity have been attributed to neural injuries associated with hypoxia, ischemia, and inflammation. Neuroimaging studies find reduced brain volumes following preterm birth (Nosarti, 2002; Reiss, 2004), and these volume changes are associated with cognitive outcomes (Reiss, 2004). Prematurity is also associated with a distinctive pattern of neural injury to periventricular cerebral white matter. These injuries can be assessed *in vivo* using Diffusion Tensor Imaging (Back, et al., 2007; Soria-Pastor et al., 2008; Vangberg et al., 2006; Volpe, 2009). Microstructural properties of the white matter, such as fractional anisotropy, are correlated with many aspects of cognitive, language, and academic skills (Andrews et al., 2010; Feldman et al., 2010; Frye et al., 2008; Mullen et al., 2010). In addition, preterm birth begins an accumulation of early life experiences that may be distinctive from the early life experiences of children born at term. This early experience may also affect how children process information.

**1.1.2 Functional imaging in children born preterm**—Relatively few studies have used fMRI to understand neural processing in children born preterm. Peterson (2003) compared 8-year old preterm and full term participants in a passive language listening task and found that the preterm group activated areas of the brain during meaningful speech that the term children activated during phonologically elemental sounds. The findings were most pronounced among the children with the lowest verbal IQ. Ment and colleagues (2006) found similar results in 12-year old children born preterm and full term. Gozzo and colleagues (2009) reported that functional connectivity in the language network was abnormal in children born preterm; they had stronger connections between Wernicke’s area and the right inferior frontal gyrus, and between bilateral superior marginal gyri than did the full term controls. The performance of preterms on the language batteries in these studies was below that of the controls, leaving open the possibility that the results were due to language function rather than prematurity. What is needed is a study where patterns of brain activation are assessed in preterm and full term children with comparable levels of performance on the relevant tasks.

## 1.2 Aims and Hypotheses

Language comprehension is pivotal to social and academic functioning. This fMRI study describes the neural activation patterns of auditory sentence comprehension in preterm and full term adolescents. We have previously reported on patterns of activation on the same task in a group of full term peers in relation to individual differences in offline measures of language function (Yeatman et al., 2010). We selected preterm adolescents on the basis of verbal IQ, receptive vocabulary, and general receptive language test scores within the normal range and relatively well matched to the full term group. These criteria served to improve upon prior fMRI studies of language outcomes following preterm birth by

controlling for offline language and cognitive ability; any differences in neural activation patterns found between groups could then plausibly be attributed to preterm birth. Moreover, approximately half of all children born preterm have mild to moderate impairments in language and/or other abilities that do not meet criteria for special education but nevertheless preclude levels of academic achievement equal to that of their full term peers. Demonstrating that preterm birth significantly affects where the brain supports language comprehension, even when standardized test scores reflect performance in the normal range, may inform and impact further research on educational interventions or accommodations for this prevalent group of school-aged children.

The auditory sentence verification task used here manipulated sentence length and syntactic difficulty in a  $2 \times 2$  factorial design, permitting analyses of the neural response to increasing demands. We extended the analyses by (a) including the preterm group, and (b) by considering not only the initial auditory (i.e., listening) phase of the task, but also the subsequent phase of the task when participants verified and responded to each sentence. In prior work (Yeatman et al., 2010), we found that in full term adolescents, increased syntactic difficulty in sentences led to increased activation in the left temporal-parietal junction and right superior temporal gyrus. Activation in frontal areas was positively correlated with individual differences in standard scores on a receptive vocabulary test.

We hypothesized that prematurity is an intrinsic factor that differentially modulates the relationship between a child's neural activations and their performance on linguistic tasks such as sentence comprehension. If correct, then the preterm participants, selected for normal-range cognitive and receptive language skill, should produce different patterns of activation than full term participants. If not a main effect of group, we at least expected group by condition interactions because we reasoned that the neural responses might diverge during challenging tasks, similar to the previous findings with full term adolescents that activations differed as a function of individual differences in language abilities. Moreover, in both phases, we expected these interactions particularly when syntactic difficulty increased, as prematurity has been shown to impair syntactic processing (Lee et al., 2010).

## 2. Method

### 2.1. Participants

**2.1.1. Preterm participants**—Eighteen children between the ages of 9.7 and 14.4 years of age ( $M = 12$ ,  $SD = 1.5$ ) with a history of preterm birth participated in this experiment (see Table 1). Nine were male, one was African-American, four were Asian-American, and one was Hispanic. Prematurity was defined as gestational age below 36 weeks to assure separation from the full term group (group  $M = 28.8$ ,  $SD = 2.6$ ). As a group, their mean birth weight was 1210 grams ( $SD = 460$ ). Three children had mild neurological damage as reported on their medical records; however, no children had a history of seizures or profound neurocognitive impairment. Though the group had normal intelligence and receptive language scores, four children were receiving special education services at the time of test. All parents reported the home language environment was dominant for English.

**2.1.2. Full term comparison participants**—Fourteen healthy children between 10 and 16 years of age (mean age = 12.5, standard deviation 1.9), some of whom took part in a larger study of language processing in adolescence served as comparison participants in this study (see Table 1). This group of full term born individuals consisted of six boys, one African-American, three Asian-American, and three Hispanic participants. Mean gestational age (39.6 weeks) and mean birth weight (3266g) were consistent with full term birth. All parents reported the home language environment was dominant for English. None of the full term comparison participants had a history of neurological, developmental, medical, or

psychological disorder, none reported impairments in vision or hearing, and none reported use of special education accommodations.

Preterm participants were enrolled in the Palo Alto, CA site of a larger multi-site study of prematurity outcomes (Lee et al., 2010; Loe et al., 2011). Full term comparison participants were recruited from the local public and private schools (Yeatman et al., 2010). Maternal education level, as a proxy for socioeconomic status (SES), was coded on an ordinal scale where 1 = less than high school, 2 = completed high school but only some college education, and 3 = completed at minimum a college education. Given our location in Palo Alto, CA, both groups were weighted toward high SES (preterm group  $M = 2.6$ , full term group  $M = 2.5$ ; Mann-Whitney U test revealed no differences,  $z = -0.15$ ,  $p > .05$ ). The Institutional Review Board at Stanford University approved this protocol. Parents gave informed consent and study participants gave assent to participate, and families were paid for their participation.

## 2.2. Behavioral measures of language and cognitive ability

We assessed participants' language and cognitive abilities across two sessions, lasting approximately 60–90 minutes each. The Wechsler Abbreviated Scales of Intelligence (WASI; The Psychological Corporation, 1999), a widely used standardized test of intellectual ability yielded both verbal and nonverbal cognitive ability indices as well as a full scale IQ composite score. The Peabody Picture Vocabulary Test, Third edition (PPVT-III; Dunn & Dunn, 1997), a standardized test of receptive vocabulary, asked participants to identify which of four line drawings best represents the stimulus word presented orally by the examiner. The Comprehensive Evaluation of Language Fundamentals, Fourth edition (CELF-IV; Semel et al., 2005) is a norm-referenced test of expressive and receptive language skills. We used the CELF-IV receptive language index to measure the participants' listening and comprehension skills.

## 2.3. fMRI task – Auditory sentence verification

Participants also completed a functional neuroimaging scan protocol that lasted approximately 60 minutes wherein the experimental task was auditory sentence verification. The task is discussed in detail elsewhere (Yeatman et al., 2010), but consisted of auditory presentation of sentences (“Auditory” phase) followed by a picture and button-box response from the participant (“Verification/Response” phase). In the Verification/Response Phase, participants pressed a button to verify YES/NO whether a picture accurately depicted the sentence or not. Sentences and pictures were derived from the Test for Reception of Grammar, Second edition (TROG-2; Bishop, 2003). Sentences were divided into four conditions, each consisting of three different length/syntactic constructions: short/easier, short/difficult, long/easier long/difficult (see Table 2).

Auditory and visual stimuli were presented using E-Prime (Psychology Software Tools, PA) in an event-related design. Each trial lasted 8.75 seconds (3.5 s for the Auditory phase and 5.25 s for the Verification/Response phase). Longer sentences were presented with shorter pauses between words to fit within the 3.5 s duration, which still reflected a natural rate of conversation. Trials were preceded and followed by a fixation cross that lasted a variable amount of time, which permitted reliable estimates of baseline BOLD signals which were then subtracted away in the general analysis of task related activity. The main contrasts of interest were Long minus Short sentences, and syntactically Difficult minus Easy sentences. In this way, comparisons were largely invariant to the sensory components, as those existed in all four conditions.

A total of 96 presentations were divided into 4 runs of 24 stimuli, each run lasting 4 min and 40 s. Ten 4-run task lists were generated with the sentences in random order and separated by randomized time intervals of fixation jittered in 1.75 s (from 0 to 5.25 s). The order of event types was counterbalanced using “optseq2” (<http://surfer.nmr.mgh.harvard.edu/optseq/>), to minimize bias that might result from carry over effects between successive events (Dale, 1999). Each subject was randomly assigned a stimulus list.

Within the scanner, the auditory stimulus was presented through an Avotec pneumatic audio presentation system. Pneumatic tubes were sealed directly onto the earplugs, successfully reducing scanner noise while allowing the subject to clearly hear the sentences. The pictures were projected onto a screen in back of the subject from an MR compatible projector and viewed by a mirror. Button press responses were recorded for later analysis. The entire task was preceded and followed by 10.5 s of fixation to allow the signal to reach equilibrium and to obtain an accurate fixation baseline.

## 2.4. fMRI Data acquisition

We obtained the functional data from a 3T, GE Signa 750 scanner (GE Healthcare Systems). For each participant, the session began with a T1 weighted, 3-plane localizer image that was used for subsequent prescriptions. We then acquired a T1 weighted inplane anatomy with 26, 4mm thick slices (GRE, TR = 34 ms, TE = 2 ms, flip angle 30°, FOV = 22 cm, Matrix 256 × 192), aligned to the plane of the anterior and posterior commissures. Prior to collecting functional data, we ran a high order shim (Kim et al., 2002) to correct for inhomogeneities in the magnetic field. T2\*-weighted functional images were then collected using a spiral in/out pulse sequence (Glover & Law, 2001) (TR = 1.75 s; TE = 30 ms; flip angle = 70°; FOV = 22 cm; Matrix = 64 × 64; slice thickness = 4 mm) and the same prescription as the inplane anatomy. In total, 160 volumes were collected in each sequence, and the sequence was repeated four times. We discarded the first six volumes (10.5 s). After the functional scans, we collected high resolution T1 weighted IR-prep 3D FSPGR scans (FOV = 24 × 16.8 cm, Matrix = 256 × 192 × 130, 1.2mm slices, TI = 300 ms, flip angle = 15°, 1 NEX).

**2.4.1. fMRI data pre-processing**—Data were preprocessed and analyzed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>). For all participants, images were corrected for slice acquisition timing then realigned to the mean functional image to correct for within-scan and between-scan motion. The inplane image was coregistered to the mean functional image to correct for any movement between collection of anatomical and functional data. The high resolution T1 weighted image was coregistered to the inplane and segmented into gray matter, white matter, and cerebral spinal fluid. These images were smoothed and used to generate a brain mask. The gray matter segmentation was used to estimate the non-linear transformation to the MNI template. This transformation was applied to the functional images and they were resampled to 3 × 3 × 3mm voxels. The images then were smoothed with an 8 mm Gaussian kernel to reduce spatial noise. We used the same approach for the preterm and full term group, recognizing that other studies have also used similar transformations across groups (Ment, et.al, 2006). While this method has a precedent, it has not yet been fully validated.

## 3. Results

### 3.1. Behavioral measures

Behavioral data were analyzed in MATLAB (The Mathworks, MA) and SPSS (SPSS Inc., IL). For each subject, we computed the number of correct responses in total and for each sentence type. We also computed mean reaction time to respond on correct trials only. Three

preterm participants were excluded from all analyses based on performance <70% correct in the scanner.

Table 3 includes the mean standardized scores on the WASI, PPVT-III and CELF-IV for preterm and full term participants. As a group, the preterm children were slightly above average on these measures. Though the preterm group had a lower mean IQ than the full term group, the difference did not reach statistical significance. The other scores were very similar and also did not differ significantly from the full term group. To examine correlations among these measures within the preterm group, we used Spearman correlations because of small sample size. Examining correlations among receptive language tests, we found that scores on the PPVT-III strongly correlated with CELF-IV scores ( $r = 0.77, p < .000$ ), and both measures strongly correlated with WASI Full IQ (PPVT:  $r = .66, p < .000$ ; CELF-IV:  $r = .80, p < .000$ ).

Table 3 also includes the number of correct responses and mean reaction times for each sentence type in the fMRI task. A repeated measures ANOVA for the preterm group only revealed that accuracy did not vary as a function of sentence length ( $F(1, 14) = 2.87, p = 0.11$ ). However, within the group, as expected, the preterms' reaction times were significantly longer ( $F(1, 14) = 181.94, p < 0.001$ ) for long sentences than for short sentences. For syntactically difficult sentences, accuracy was lower ( $F(1, 14) = 15.63, p = 0.001$ ) and reaction time longer ( $F(1, 14) = 70.69, p < 0.001$ ). Reaction time showed a significant interaction between length and difficulty ( $F(1, 14) = 38.73, p < 0.001$ ) but accuracy did not ( $F(1, 14) = 0.06, p = 0.82$ ). Using Spearman correlations, reaction times among all trial types were highly correlated ( $r > 0.8, p < 0.001$ ). Age was significantly and positively correlated with task accuracy on the shortest, easiest sentences ( $r = 0.69, p = 0.004$ ) and the longest, hardest sentences ( $r = .63, p = .009$ ). Age was also significantly and negatively correlated with reaction time on all sentence types ( $r = -0.578$  to  $-.646, p = 0.024$ – $0.009$ ). Neither standardized measure of receptive language skills showed a significant correlation with task accuracy ( $r = 0.304$ – $0.451, p = 0.091$ – $0.270$ ) or reaction time ( $r = -0.113$  to  $-0.332, p = 0.226$ – $0.689$ ). For details regarding the full term group's performance on this task, we refer the reader to Yeatman and colleagues (Yeatman et al., 2010).

Repeated-measures ANOVA for the within (sentence length, syntactic difficulty) and between (group) subjects factors revealed that preterms' performance was extremely similar to that of the full term controls on the task. Interaction terms for accuracy were not significant, indicating preterms were just as accurate as full terms on longer and shorter sentences ( $F(1, 26) = 0.10, p = .76$ ) and syntactically difficult and easier sentences ( $F(1, 26) = 0.002, p = .96$ ). Reaction was not different between the groups as a function of syntactic difficulty ( $F(1, 26) = 0.05, p = .82$ ), but the length  $\times$  group interaction was significant ( $F(1, 26) = 6.68, p = .02$ ); surprisingly the preterm group was faster on long sentences.

### 3.2. Functional neuroimaging measures

Statistical analysis was performed for each subject by calculating a general linear model for voxels within the cortical areas of the brain mask. The designs for the four sentence types were each convolved with a canonical hemodynamic response function (HRF) and their temporal and dispersion derivatives were estimated. Regressors were included for the six motion parameters, for each run and for the overall mean. A high pass filter with a 128 s cutoff was used to remove low frequency drift and an AR(1) correction was used for serial autocorrelation. For each participant, first-level analyses were performed in which two contrasts of interest (longshort sentences, syntactically difficult-easy sentences) were estimated at two time points in the trial, similar to other time-course investigations of sentence processing (Carpenter et al., 1999). We assessed (a) the first 3.5 seconds

corresponding to listening to a sentence (i.e., the Auditory phase), and (b) the next 3.5 seconds of the trial that corresponded to viewing the picture and making a button press decision (i.e., the Verification/Response phase). While the portion of the trial in which participants were allowed to verify the picture-to-sentence match and make the appropriate button press technically lasted 5.25 s, all response times were under 3.5 s. Therefore, we modeled activation for these two phases similarly.

Group level statistics were then calculated using a random effects model with the contrast images from each subject. Gestational age was not covaried in the models for three reasons. First, the distribution of gestational age was bimodal. Second, there was limited variation in the preterm group ( $n = 15$ ) as 10 participants had gestational ages of 27 or 28 weeks. Third, we expected no variation as a function of gestational age in the full term group. Also, due to data loss, analyses of the Auditory phase and Verification/Response phase had different sample sizes (Auditory phase: preterm  $n = 15$ , full term  $n = 14$ ; Verification/Response phase: preterm  $n = 14$ , full term  $n = 12$ ). We set the statistical threshold for omnibus ANOVA at  $p < 0.05$  (corrected for multiple comparisons using Familywise Error or FWE). To reduce the risk of accumulating alpha errors, we analyzed clusters with a minimal cluster size of 10 voxels. For post-hoc analyses, we used a less stringent statistical criterion of  $p < .01$ .

### 3.3. fMRI analyses

Below we present the results of two omnibus ANOVAs, separated by phase of the experimental task, i.e., the Auditory and Verification/Response phases. We describe the patterns of activation corresponding to main effects and interactions, followed by appropriate post-hoc tests to describe any differences between groups. We then describe activation that was responsive to individual variation in general receptive language ability, again for both groups. For all analyses, figures illustrate activation for relevant contrasts on selected T1 image slices, and accompanying graphs describe the difference in activation for a given contrast (e.g., longer minus shorter sentences) by group per cluster. Tables list the areas of activation, peak  $Z$  scores, number of active voxels per cluster, and MNI coordinates for the peak activations.

**3.3.1. Auditory phase ANOVA results**—There was no main effect of group in the Auditory phase of the trial. There was a main effect of condition and two significant two-way interactions: difficulty  $\times$  length and group  $\times$  length. Activation in the right superior temporal gyrus (MNI coordinates 63, -21, 6) corresponded to the interaction of difficulty and length ( $Z = 5.43$ ,  $p < 0.05$  FWE). Table 4 describes the activation corresponding to the significant interactions and posthoc tests in this phase.

Figure 1 depicts multiple results related to the group  $\times$  length interaction in the Auditory Phase. Panel A shows areas of significant group  $\times$  length interaction ( $p < 0.05$  FWE) on relevant axial slices of a group T1 image. Panel B illustrates the differences in percent activation associated with increased sentence length for the preterm (green) and full term (blue) groups at each significant cluster shown in Panel A. Panel C depicts *post hoc* test results on the group  $\times$  length interaction ( $p < 0.01$  uncorrected) for the whole brain, revealing areas in which the difference in activation between Long and Short sentences was greater for full terms relative to preterms: small areas within left occipital lobe, right occipital lobe, and left medial frontal gyrus. In this analysis, there were no areas in which the difference on this contrast was greater for preterms relative to full terms. The three-way interaction of group  $\times$  length  $\times$  difficulty was nonsignificant.



**3.3.2. Verification/Response phase ANOVA results**—There was no main effect of group in the Verification/Response phase of the trial. There was a main effect of condition and several significant two-way interactions: difficulty  $\times$  length, group  $\times$  length, and group  $\times$  difficulty. Activation in the right superior temporal gyrus, right and left occipital lobes, precuneus, left inferior frontal gyrus, right parahippocampal gyrus, and cerebellum areas corresponded to the interaction of difficulty  $\times$  length (see Table 5).

Figure 2 depicts multiple results related to the group  $\times$  length interaction in the Verification/Response Phase. Panel A shows areas of significant group  $\times$  length interaction ( $p < 0.05$  FWE) on relevant axial slices of a group T1 image. Panel B illustrates the differences in percent activation associated with increased sentence length for the preterm (green) and full term (blue) groups at each significant cluster shown in Panel A. Panel C depicts *post hoc* test results on the group  $\times$  length interaction ( $p < 0.01$  uncorrected) for the whole brain, revealing areas in which the difference in activation between Long and Short sentences was greater for preterms relative to full terms: right inferior parietal lobe, left middle frontal gyrus, left and right insula, and left middle temporal gyrus. In this analysis, there were no areas in which the difference on this contrast was greater in full terms relative to preterms. The three-way interaction of group  $\times$  length  $\times$  difficulty was nonsignificant.

Figure 3 depicts multiple results related to the group  $\times$  difficulty interaction in the Verification/Response Phase. Panel A shows areas of significant group  $\times$  difficulty interaction ( $p < 0.05$  FWE) on relevant slices of a group T1 image. Panel B illustrates the differences in percent activation associated with increased syntactic difficulty for the preterm (green) and full term (blue) groups at each significant cluster shown in Panel A. Panel C depicts *post hoc* test results on the group  $\times$  difficulty interaction ( $p < 0.01$  uncorrected) for the whole brain, revealing areas in which the difference in activation between syntactically Difficult and Easy sentences was greater for preterms relative to full terms: the left middle frontal gyrus, left inferior frontal gyrus, and left parietal lobe. Again, there were no areas in which the difference on this contrast was greater for full terms relative to preterms.

**3.3.3. Effects of individual differences in receptive language ability**—We recognized that the patterns of neural activation on this task described above may be related to other factors inherent in these two groups of adolescents, such as age, receptive language ability, and differences in response time. To explore this possibility, we repeated the above posthoc analyses covarying for these factors. We constructed random effects models for each posthoc test including the following covariates separately: age, CELF-IV receptive language index standard score, and global response time. We found that even with these covariates, the areas of activation and direction of effects in the group  $\times$  condition interactions (i.e., group  $\times$  length, group  $\times$  difficulty) did not change. (These results are summarized in Supplementary Table 1.)

## 4. Discussion

Findings from this neuroimaging study are consistent with our hypothesis that prematurity is a modulator of the neural response to increased task demands (here, increased sentence length or syntactic difficulty) during auditory sentence comprehension. Adolescent participants completed an auditory sentence verification task in the fMRI scanner, a task that entailed listening to a sentence, viewing a picture, and pressing a button to indicate whether or not the picture accurately depicted the meaning of the sentence. Two factors, sentence length and syntactic difficulty, were manipulated in a factorial design and we analyzed two phases within each trial. The Auditory phase reflected participants' ability to passively listen to and comprehend the sentences, while the Verification/Response phase necessitated

higher-order cognition, i.e., comparing the meaning of the sentence to a visual image and making a decision as to its validity. Accuracy and response time during the task were reasonably similar between groups. Critically, ANOVA tests of the areas of neural activation demonstrated no main effects of group, but rather several interesting group by condition interactions. Children born preterm showed important differences in response to increased task demands compared to children born at term, even when performance on the task was comparable.

A particular strength of the design was that both preterm and full term adolescents had scores on standardized assessments of IQ, receptive vocabulary, and receptive language skill that were within normal limits, and scores on these measures were also reasonably well matched between groups prior to scanning. In this way, differences between groups can reliably be attributed to prematurity and not to differential performance.

### 4.3 Interactions in the Auditory phase

During the Auditory phase of the task, the significant group  $\times$  length interaction activated left and right superior temporal gyrus and the occipital lobe. Longer sentences also activated similar regions in our previous study (Yeatman et al., 2010); likely reflecting increased auditory, linguistic, and semantic processing. Examination of image data extracted from the three significant clusters in the interaction (see Figure 1, Panel B) indicates that preterms showed numerically larger differences between longer and shorter sentences relative to full terms; the greater activation perhaps suggesting greater cognitive output to meet task demands. *Post hoc* tests revealed that the length manipulation elicited a significantly greater activation in full terms relative to preterms in left and right occipital lobes, and left medial frontal gyrus. Medial frontal gyrus activation has previously been observed in neuroimaging studies of sentence comprehension, particularly when sentences included subject- and object-relative clauses (Caplan et al., 1999; Caplan et al., 2000) as did our sentences. Caplan and colleagues tentatively attributed activation in medial frontal gyri to roles in attentional and control processes (Posner et al., 1987, 1988). Thus, the *post hoc* findings for the full term group suggest that they appropriately recruit areas of the brain in the initial phase of the task that support other cognitive mechanisms (i.e., attention, cognitive control) to handle increased task demands.

### 4.4. Interactions in the Verification/Response phase

Turning to the Verification/Response phase of the trial, the significant interactions showed wider activation patterns than earlier in the trial, reflecting the change in cognitive demands from passive listening to active verification and formulation of a response. We again found a group  $\times$  length interaction in this later phase, and most of this activation was left-lateralized. Beyond the right and left superior and middle temporal gyri and left occipital lobe activation related to increased sentence length that we found in the Auditory phase, we also observed activation in left middle frontal gyrus, left inferior frontal gyrus, the left precuneus, and the postcentral gyrus in the left hemisphere. Moreover, examination of the image data for each of these significant clusters in the interaction, as well as *post hoc* tests, point to large differences between preterm and full term groups in the left frontal regions and the left precuneus. Activation in these clusters is consistent with other fMRI studies of sentence comprehension/verification tasks that reported activation in Broca's area and left prefrontal cortex more generally (Bunge et al., 2000; Hashimoto & Sakai, 2002; Rogalsky & Hickok, 2011), as well as the left temporal lobe (Friederici, 2002). It appears that preterms require greater recruitment of prefrontal cortex relative to their full term peers when verifying and responding to longer sentences. The precuneus has been implicated in tasks requiring visuospatial imagery and episodic retrieval (Cavanna et al., 2006; Krause et al., 1999;

Lundstrom et al., 2005), perhaps suggesting that preterms rely more on a visual imagery strategy than full term children when listening to longer sentences.

Importantly, we found a significant group  $\times$  difficulty interaction in the second phase of the trial in left and right middle temporal gyrus, left precuneus, left insula, right cerebellum, and right cingulate gyrus. Interestingly, there was overlap between areas related to the interaction of group  $\times$  length and group  $\times$  difficulty; the common activation in bilateral superior temporal gyri and left precuneus areas suggests a shared system for accommodating increased task demands in sentence comprehension, though the two groups differentially recruited these areas of the brain. Last and most relevant to the aims of the study, *post hoc* tests revealed that the difference in activation related to Difficult versus Easy sentences was greater for preterms relative to full terms in the left and right middle frontal gyri (a part of the dorsolateral prefrontal cortex). This finding suggests that in comparison to full term peers, preterm adolescents recruited areas of the brain associated with cognitive control (MacDonald et al., 2000) when processing difficult material. We know that cognitive control is a domain that is weak in preterms as a group (Loe et al., 2011). We carefully note that though speculative at this point, it is plausible that preterm children have to engage cognitive control mechanisms more than full term children in order to attain comparable performance in the scanner. Thus, underlying apparently normal performance may be different neural mechanisms, pointing to the continued need for more neuroimaging research to understand the functional differences in preterms on linguistic and cognitive tasks.

Taken together, these findings are notable for an absence of main effects of group collapsing across both task conditions (increased sentence length or syntactic difficulty). Instead, we found group  $\times$  condition interactions: in both phases of the trial we found a group  $\times$  length interaction, and in the second phase of the trial we found a group  $\times$  difficulty interaction. Moreover, *post hoc* tests of both interactions in the second phase of the trial, where participants verified the relationship between sentence and picture and made a YES/NO response, revealed that the preterm group had greater activation in response to the experimental contrasts relative to the full terms, but not vice versa. The single most important take-home message from these findings is that group status (i.e., presence/absence of preterm birth) modulated the functional neuroanatomy related to increased task demands during auditory sentence comprehension. We verified that through ANCOVA analyses that the group differences could not be attributed to any subtle between-group differences in age, receptive language skill, or reaction time in the scanner.

#### 4.5 Educational implications

There has been an intense debate regarding the integration of cognitive neuroscience and education (Blakemore & Frith, 2005; Gabrieli, 2009; Goswami, 2006; Howard-Jones, 2010; Varma et al., 2008). Translating brain-behavior findings directly into meaningful changes in curricula or into best practices within the educational system is an endeavor that some believe at best should be attempted cautiously (Bruer, 1997; Goswami, 2009; Stern, 2005). We agree that neuroscience at this point in time cannot readily inform the development of educational curricular or teaching methods. For such purposes, behavioral and educational research provides a more direct approach for assessment of whether a particular curriculum or teaching method is effective.

Yet cognitive neuroscience has provided numerous findings from atypical populations that afforded points of reference between the disciplines and in turn shaped educational approaches, including reading/dyslexia (Frey & Fisher, 2010; Gabrieli, 2009; Hoeft et al., 2010; Kevan & Pammer, 2009), mathematical learning/dyscalculia (Butterworth et al., 2011), emotion regulation/behavioral disorders (Blair, 2002; MacLeod, 2010; Ochsner et al., 2010), and attention/ADHD (Loe et al., 2011; Vaidya et al., 2005). The contribution of

cognitive neuroscience has focused on identifying both core deficits and underlying phenotypes of a particular disorder or deficit; these findings in turn have guided the creation of targeted interventions and development of evidence-based practices in education (Fletcher et al., 2005; Räsänen et al., 2009; Shaywitz et al., 2004; Wilson et al., 2006).

We believe that studies, such as this one, are similarly relevant to education because they provide important insights about the students, though not students with obvious impairments in an academic domain. Rather, this study informs us about intrinsic, background factors pertinent to education and optimal academic achievement. In a similar way, for example, drawing from research in literacy development, we know that poor readers from families with low SES show different patterns of activation than poor readers with high SES (Noble et al., 2006), equating for performance on a reading-related task. This finding does not in and of itself determine which is the best curriculum for teaching reading to everyone, or even to poor children. However, it suggests that the way children process information for reading is the cumulative effect of their social history and circumstance. The educational implications of these findings are that literacy programs may need to consider the child's SES in developing and assessing educational methods or in providing accommodations for children from poor homes. Also, intervention studies prompted by research like that of Noble and colleagues (2006) may account for variation in extrinsic factors such as SES in the efficacy of their training programs (Keller & Just, 2009; Meyler et al., 2008).

In this study, we show that children born preterm who have performance comparable to full term peers have different patterns of neural activation on a language comprehension task. These effects persisted when variance in age, receptive language, and response time were taken into account. The engagement of regions of frontal lobe suggests that preterm children may be utilizing cognitive control areas to a greater degree in the late phase of the task than do their full term peers. If we agree that these findings suggest greater recruitment of cognitive control mechanisms among the preterm group, then we may need to consider strategies to reduce the need for effortful control processes when teaching and assessing children born preterm. Accommodations, such as more trials for learning, use of visual and other supports for understanding, and/or more time during testing, might be appropriate for this population.

We recognize that we have studied a small sample of children born preterm who may not be fully representative of the entire population. We recognize that we must validate methods, such as co-registration and normalization of the clinical and typical population together. We further acknowledge that the differences in neural activations that we found are modest in the Auditory Phase, where full terms showed greater activation on *post hoc* tests, and greater in the second phase of the task, where preterms showed greater activation on *post hocs*. However, we believe these observations open the door for a new type of inquiry into intrinsic variation among students on the basis of medical history, even students whose performance appears typical at the behavioral level.

We hope the findings in this paper inspire efforts toward better understanding of the neural basis of cognitive and linguistic ability for atypical populations of children. We emphasize this understanding of distinctive neural activation patterns is particularly important in cases where children with atypical medical histories and/or neurological status have comparable performance on standardized measures and do not meet eligibility for special education. How researchers in the fields of cognitive neuroscience and education should intervene differently is a matter for future study. Certainly the basic findings herein need to be replicated and extended to other linguistic, cognitive, and academic outcomes. In cases where students are not eligible for special education, future findings may guide mainstream classroom accommodations that collectively provide relief from the increased effort required

to attain normal performance. Educational policy may also need to allow for variations in eligibility criteria as a function of medical and/or neurological history.

## 5. Conclusions

We know that preterm born children, even those that have been spared profound sensorimotor and cognitive deficits, still experience difficulties in linguistic and academic domains that are not well described by formal testing. Prior studies have described specific kinds of mild to moderate deficits common following preterm birth that have relevance to educational achievement, including decreased speed of processing, impaired comprehension, and attentional and executive function problems. This study adds to this existing knowledge base and tells us that (a) preterm born children use different brain areas even while verbal IQ, receptive language skills, and task performance may match those of their full term peers, and (b) they use different brain areas primarily when the linguistic task becomes more difficult. In the future, these findings may indicate a need to prepare teachers for the possibility of targeted accommodations in the classroom for this group of students. While we defer making concrete suggestions from this study, we believe that educators and policymakers should consider the educational implications of the unique linguistic and cognitive processing styles of preterm born children and adolescents.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We extend our thanks and appreciation to the patients and families who participated in our study. We also thank the members of the Developmental and Behavioral Pediatrics laboratory group for their critiques and suggestions, and to Fumiko Hoeft for her invaluable expertise and generosity.

Sources of Support: This work was supported by a grant from the National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development, RO1 HD046500 to Heidi M. Feldman and by the Clinical and Translational Science Award 1UL1 RR025744 for the Stanford Center for Clinical and Translational Education and Research (Spectrum) from the National Center for Research Resources, National Institutes of Health. We also acknowledge support from the Lucile Packard Children's Hospital Foundation and a Pediatric Research Fellowship grant to Laura H.F. Barde.

## References

- Aarnoudse-Moens CSH, Weisglas-Kuperus N, van Goudoever JB, Oosterlaan J. Meta-Analysis of Neurobehavioral Outcomes in Very Preterm and/or Very Low Birth Weight Children. *Pediatrics*. 2009; 124(2):717–728. [PubMed: 19651588]
- Andrews JA, Ben-Shachar M, Yeatman JD, Flom LL, Luna B, Feldman HM. Reading performance correlates with white-matter properties in preterm and term children. *Developmental and Child Neurology*. 2010; 52(6):94–100.
- Aylward GP. Cognitive and neuropsychological outcomes: More than IQ scores. *Mental Retardation and Developmental Disabilities Research Reviews*. 2002; 8:234–240. [PubMed: 12454899]
- Back SA, Riddle A, McClure MM. Maturation-dependent vulnerability of perinatal white matter in premature birth. *Stroke*. 2007; 38:724–730. [PubMed: 17261726]
- Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJS. Cognitive and behavioral outcomes of school-aged children who were born pre-term: A meta-analysis. *Journal of the American Medical Association*. 2002; 288(6):728–737. [PubMed: 12169077]
- Bishop, DV. Test for reception of grammar (TROG-2). Oxford, United Kingdom: Pearson Assessment; 2003.

- Blair C. School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*. 2002; 57(2):111–127. [PubMed: 11899554]
- Blakemore, S-J.; Frith, U. *The learning brain: Lessons for education*. Oxford, UK: Blackwell; 2005.
- Bruer JT. Education and the brain: A bridge too far. *Educational Researcher*. 1997; 26(8):4–16.
- Buchmann C, DiPrete TA, McDaniel A. Gender inequalities in education. *Annual Review of Sociology*. 2008; 34:319–337.
- Bunge SA, Klingberg T, Jacobsen RB, Gabrieli JDE. A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Sciences*. 2000; 97:3573–3578.
- Burman DD, Bitan T, Booth JR. Sex differences in neural processing of language among children. *Neuropsychologia*. 2008; 46:1349–1362. [PubMed: 18262207]
- Butterworth B, Varma S, Laurillard D. Dyscalculia: From brain to education. *Science*. 2011; 332:1049–1053. [PubMed: 21617068]
- Caplan D, Alpert N, Waters G. PET studies of sentence processing with auditory sentence presentation. *Neuroimage*. 1999; 9:343–351. [PubMed: 10075904]
- Caplan D, Alpert N, Waters G, Olivieri A. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping*. 2000; 9:65–71. [PubMed: 10680763]
- Carpenter PA, Just MA, Keller TA, Eddy WF, Thulborn KR. Time course of fMRI-activation in language and spatial networks during sentence comprehension. *Neuroimage*. 1999; 10:216–224. [PubMed: 10417254]
- Cavanna AE, Trimble MR. The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*. 2006; 129(3):564–583. [PubMed: 16399806]
- Dale AM. Optimal experimental design for event-related fMRI. *Human Brain Mapping*. 1999; 8(2–3): 109–114. [PubMed: 10524601]
- Deary I, Strand S, Smith P, Fernandes C. Intelligence and educational achievement. *Intelligence*. 2007; 35(1):13–21.
- Duckworth AL, Quinn PD, Lynam DR, Loeber R, Stouthamer-Loeber M. Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences*. 2011; 108(19):7716–7720.
- Dunn, LM.; Dunn, LM. *Peabody Picture Vocabulary Scale- Third Edition (PPVT-III)*. Circle Pines, Minnesota: American Guidance Service, Inc; 1997.
- Feldman HM, Yeatman JD, Lee ES, Barde LHF, Gaman-Bean S. Diffusion tensor imaging: A review for pediatrics researchers and clinicians. *Journal of Developmental and Behavioral Pediatrics*. 2010; 31(4):346–356. [PubMed: 20453582]
- Fletcher JM, Francis DJ, Morris RD, Lyon GR. Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*. 2005; 34:506–522. [PubMed: 16083392]
- Friederici AD. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*. 2002; 6(2):78–84. [PubMed: 15866191]
- Frey N, Fisher D. Reading and the brain: What early childhood educators need to know. *Early Childhood Education Journal*. 2010; 38(2):103–110.
- Frye RE, Hasan K, Xue L, Strickland D, Malmberg B, Liederman J, Papanicolaou A. Splenium microstructure is related to two dimensions of reading skill. *Neuroreport*. 2008; 19:1627–1631. [PubMed: 18806688]
- Gabrieli JDE. Dyslexia: A new synergy between education and cognitive neuroscience. *Science*. 2009; 325:280–283. [PubMed: 19608907]
- Glover GH, Law CS. Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magnetic Resonance in Medicine*. 2001; 46(3):515–522. [PubMed: 11550244]
- Goswami U. Neuroscience and education: from research to practice? *Nature Reviews Neuroscience*. 2006; 7:406–413.
- Goswami U. Mind, brain, and literacy: Biomarkers as usable knowledge for education. *Mind, Brain, and Education*. 2009; 3(3):176–184.

- Gozzo Y, Vohr B, Lacadie C, Hampson M, Katz KH, Maller-Kesselman J, Schneider KC, Peterson BS, Rajeevan N, Makuch RW, Constable RT, Ment LR. Alterations in neural connectivity in preterm children at school age. *Neuroimage*. 2009; 48:458–463. [PubMed: 19560547]
- Hack M. Adult outcomes of preterm children. *Journal of Developmental and Behavioral Pediatrics*. 2009; 30(5):460–470. [PubMed: 19823140]
- Hack M, Flannery DJ, Schlucter M, Cartar L, Borawski E, Klein N. Outcomes in young adulthood for very low birth weight infants. *New England Journal of Medicine*. 2002; 346(3):149–157. [PubMed: 11796848]
- Hackman DA, Farah MJ, Meaney MJ. Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*. 2010; 11(9):651–659.
- Hashimoto R, Sakai K. Specialization in the left prefrontal cortex for sentence comprehension. *Neuron*. 2002; 35(3):589–597. [PubMed: 12165479]
- Hoefl F, Ueno T, Reiss AL, Meyler A, Whitfield-Gabrieli S, Glover GH, Keller TA, Kobayashi N, Mazaika P, Jo B, Just MA, Gabrieli JDE. Prediction of children's reading skills using behavioral, functional, and structural imaging measures. *Behavioral Neuroscience*. 2007; 121(3):602–613. [PubMed: 17592952]
- Howard-Jones, P. *Introducing neuroeducational research: Neuroscience, education, and the brain*. New York: Routledge; 2010.
- Johnson SJ, Hennessy EM, Smith R, Trikić R, Wolke D, Marlow N. Academic attainment and special education needs in extremely preterm children at 11 years of age: The EPICure Study. *Archives of Disease in Childhood*. 2009; 94(4):F283–F289. [PubMed: 19282336]
- Keller TA, Just MA. Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers. *Neuron*. 2009; 64:624–631. [PubMed: 20005820]
- Kevan A, Pammer K. Predicting early reading skills from pre-reading measures of dorsal stream functioning. *Neuropsychologia*. 2009; 47(14):3174–3181. [PubMed: 19651148]
- Kim D-H, Adalsteinsson E, Glover GH, Spielman DM. Regularized higher-order in vivo shimming. *Magnetic Resonance in Medicine*. 2002; 48(4):715–722. [PubMed: 12353290]
- Kishiyama MM, Boyce WT, Jimenez AM, Perry LM, Knight RT. Socioeconomic disparities affect prefrontal function in children. *Journal of Cognitive Neuroscience*. 2009; 21(6):1106–1115. [PubMed: 18752394]
- Kraft RH, Nickel LD. Sex-related differences in cognition: Development during early childhood. *Learning and Individual Differences*. 1995; 7(3):249–271.
- Krause BJ, Schmidt D, Mottaghy FM, Taylor J, Halsband U, Herzog H, Tellmann L, Müller-Gärtner H-W. Episodic retrieval activates the precuneus irrespective of the imagery content of word-pair associates: A PET study. *Brain*. 1999; 122(2):255–263. [PubMed: 10071054]
- Lee ES, Yeatman JD, Luna B, Feldman HM. Specific language and reading skills in school-aged children and adolescents are associated with prematurity after controlling for IQ. *Neuropsychologia*. 2010; 49(5):906–913. [PubMed: 21195100]
- Li-Grining CP, Votruba-Drzal E, Maldonado-Carreño C, Haas K. Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*. 2010; 46(5):1062–1077. [PubMed: 20822223]
- Loe IM, Lee ES, Luna B, Feldman HM. Behavior problems of 9–16 year old preterm children: Biological, sociodemographic, and intellectual contributions. *Early Human Development*. 2011; 87(4):247–252. [PubMed: 21316875]
- Lundstrom BN, Ingvar M, Petersson KM. The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. *Neuroimage*. 2005; 27(4):824–834. [PubMed: 15982902]
- MacDonald AW, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*. 2000; 288:1835–1838. [PubMed: 10846167]
- MacLeod G. Identifying obstacles to a multidisciplinary understanding of 'disruptive' behavior. *Emotional and Behavioural Difficulties*. 2010; 15(2):95–109.
- Marlow N. Neurocognitive outcome after very pre-term birth. *Archives of Disease in Childhood: Fetal and Neonatal Edition*. 2004; 89(3):149–157.

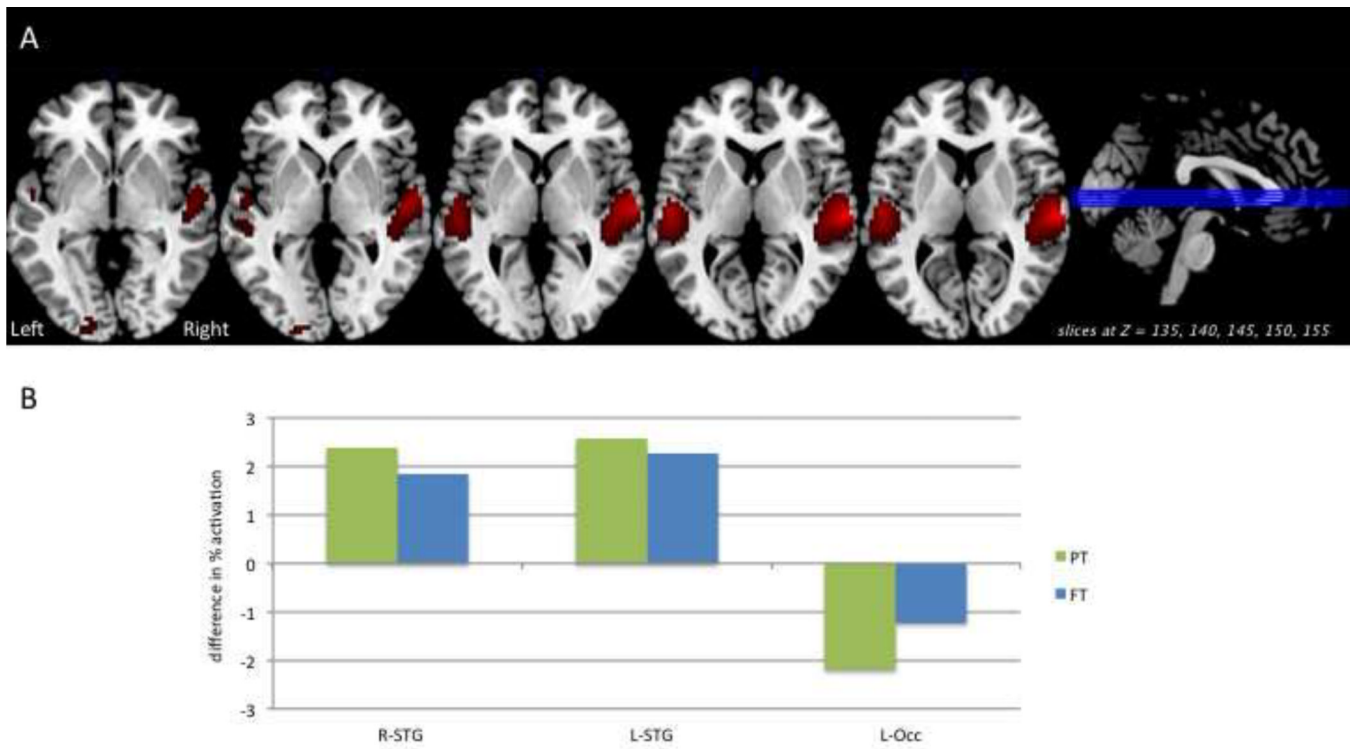
- Ment LR, Peterson BS, Meltzer JA, Vohr B, Allan W, Katz KH, Lacadie C, Schneider KC, Duncan CC, Makuch RW, Constable RT. A functional magnetic resonance imaging study of the long-term influences of early indomethacin exposure on language processing in the brains of prematurely born children. *Pediatrics*. 2006; 118:961–970. [PubMed: 16950986]
- Meyler A, Keller TA, Cherkassky VL, Gabrieli JDE, Just MA. Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia*. 2008; 46(10):2580–2592. [PubMed: 18495180]
- Mullen KM, Vohr BR, Katz KH, Scheider KC, Lacadie C, Makuch M, Reiss AL, Constable RT, Ment LR. Preterm birth results in alterations in neural connectivity at age 16 years. *Neuroimage*. 2010; 54(4):2563–2570. [PubMed: 21073965]
- Noble KG, Farah MJ, McCandliss BD. Socioeconomic background modulates cognition-achievement relationships in reading. *Cognitive Development*. 2006; 21(3):349–368. [PubMed: 19789717]
- Nosarti C, Al-Asady MHS, Frangou S, Stewart AL, Rifkin L, Murray RM. Adolescents who were born very *preterm* have decreased brain volumes. *Brain*. 2002; 125(Pt 7):1616–1623. [PubMed: 12077010]
- Ochsner KN, Bunge SA, Gross JJ, Gabrieli JDE. Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*. 2002; 14(8):1215–1229. [PubMed: 12495527]
- Peterson BS. Brain imaging studies of the anatomical and functional consequences of preterm birth for human brain development. *Annals of the New York Academy of Sciences*. 2003; 1008:219–237. [PubMed: 14998887]
- Posner MI, Inhoff AW, Friedrich FJ, Cohen A. Isolating attentional systems: a cognitive-anatomical analysis. *Psychobiology*. 1987; 15:107–121.
- Posner MI, Peterson SE, Fox PT, Raichle ME. Localization of cognitive operations in the human brain. *Science*. 1988; 240:1627–1631. [PubMed: 3289116]
- Räsänen P, Salminen J, Wilson AJ, Aunio P, Dehaene S. Computer-assisted intervention for children with low numeracy skills. *Cognitive Development*. 2009; 24:450–472.
- Reiss AL, Kesler SR, Vohr B, Duncan CC, Katz KH, Pajot S, Schneider KC, Makuch RW, Ment LR. Sex differences in cerebral volumes of 8-year-olds born preterm. *Journal of Pediatrics*. 2004; 145(2):242–249. [PubMed: 15289777]
- Rogalsky C, Hickok G. The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience*. 2011; 23:1664–1680. [PubMed: 20617890]
- Ross G, Lipper EG, Auld PA. Educational status and school-related abilities of very low birth weight premature children. *Pediatrics*. 1991; 88(6):1125–1134. [PubMed: 1956729]
- Saigal S, Doyle LW. The lancet series on preterm birth (3): An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Child: Care, Health, and Development*. 2008; 34(3):407.
- Saigal S, Rosenbaum P. What matters in the long term: reflections on the context of adult outcomes versus detailed measures in childhood. [Review]. *Seminars In Fetal & Neonatal Medicine*. 2007; 12(5):415–422. [PubMed: 17707702]
- Semel, E.; Wiig, EH.; Secord, WA. *Clinical Evaluation of Language Fundamentals (CELF-Preschool, 2nd Edition)*. San Antonio, TX: Harcourt Assessment; 2002.
- Shaywitz BA, Shaywitz SE, Blachman B, Pugh KR, Fulbright RK, Skudlarski P, Mencil WE, Constable RT, Holahan JM, Marchione KE, Fletcher JM, Lyon GR, Gore JC. Development of left occipito-temporal systems for skilled reading following phonologically-based reading intervention in children. *Biological Psychiatry*. 2004; 55:926–933. [PubMed: 15110736]
- Somers CL, Owens D, Piliawsky M. Individual and social factors related to urban African-American adolescents' school performance. *The High School Journal*. 2008; 91(3):1–11.
- Soria-Pastor S, Gimenez M, Narberhaus A, Falcon C, Botet F, Bargallo N, Mercader JM, Junque C. Patterns of cerebral white matter damage and cognitive impairment in adolescents born very preterm. *International Journal of Developmental Neuroscience*. 2008; 26:647–654. [PubMed: 18765280]
- Stern E. Pedagogy meets neuroscience. *Science*. 2005; 310(5749):745. [PubMed: 16272083]
- Sternberg RJ, Grigorenko EL, Bundy DA. The predictive value of IQ. *Merrill-Palmer Quarterly*. 2001; 47:1–41.



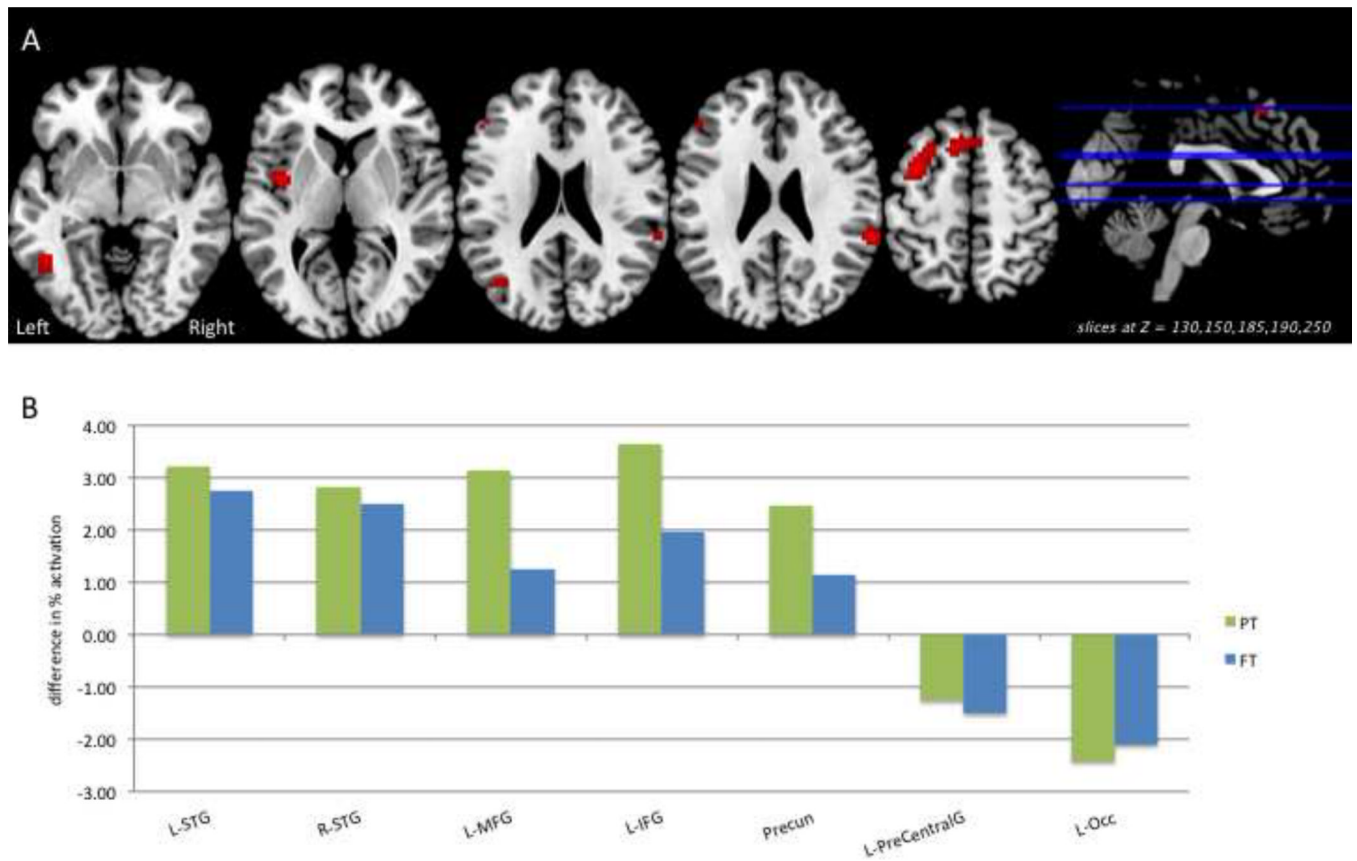
- Stronge JH, Ward TJ, Tucker PD, Hindman JL. What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education*. 2007; 20(3–4):165–184.
- The Psychological Corporation. Wechsler Abbreviated Scale of Intelligence (WASI) Manual. San Antonio, TX: The Psychological Corporation; 1999.
- Vaidya CJ, Bunge SA, Dudukovic NM, Zalecki CA, Elliott GR, Gabrieli JDE. Altered neural substrates of cognitive control in childhood ADHD: Evidence from functional magnetic resonance imaging. *American Journal of Psychiatry*. 2005; 162:1605–1613. [PubMed: 16135618]
- Vangberg TR, Skranes J, Dale AM, Martinussen M, Brubakk A-M, Haraldseth O. Changes in white matter diffusion anisotropy in adolescents born prematurely. *Neuroimage*. 2006; 32:1538–1548. [PubMed: 16843682]
- Varma S, McCandliss BD, Schwartz DL. Scientific and pragmatic challenges for bridging education and neuroscience. *Educational Researcher*. 2008; 37(3):140–152.
- Vohr BR, Wright LL, Dusick AM, Mele L, Verter J, Steichen JJ, et al. Neurodevelopmental and Functional Outcomes of Extremely Low Birth Weight Infants in the National Institute of Child Health and Human Development Neonatal Research Network, 1993–1994. *Pediatrics*. 2000; 105(6):1216–1226. [PubMed: 10835060]
- Volpe J. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. *Lancet Neurology*. 2009; 8:110–124. [PubMed: 19081519]
- Wilson A, Revkin SK, Cohen D, Cohen L, Dehaene S. An open trial assessment of “The Number Race,” an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*. 2006; 2:1–16. [PubMed: 16393341]
- Yeatman JD, Ben-Shachar M, Glover GH, Feldman HM. Individual differences in auditory sentence comprehension in children: An event-related functional magnetic resonance imaging investigation. *Brain and Language*. 2010; 114:72–79. [PubMed: 20053431]

### Highlights

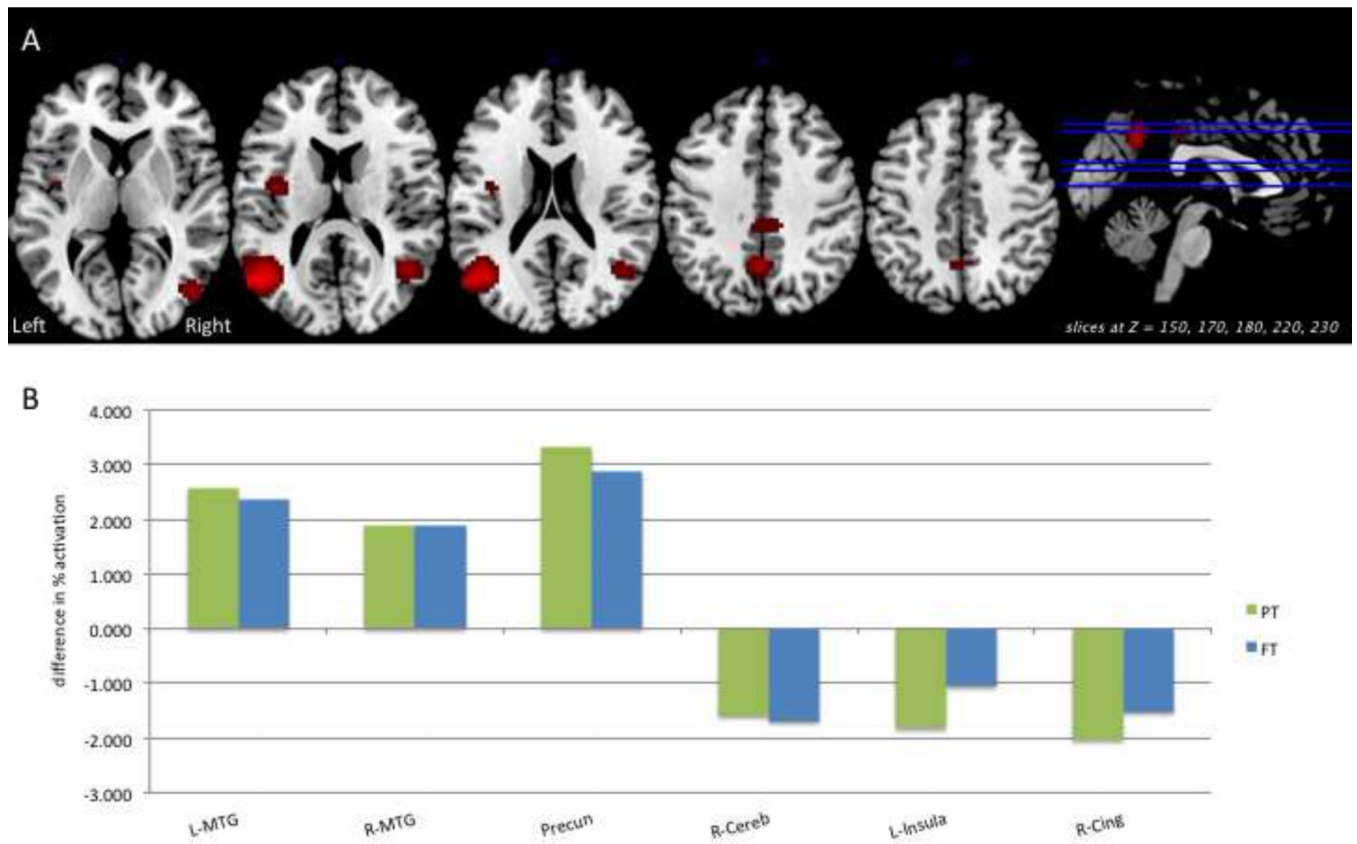
- We ask if a child's medical history accounts for variance in academic achievement.
- Preterm birth is a medical event that increases risk of poor educational outcomes.
- Preterm and full term adolescents completed an fMRI sentence comprehension task.
- Preterms activated bilateral MFG more than full terms as syntactic demands increased.
- Differences in preterms' functional activation may have educational implications.



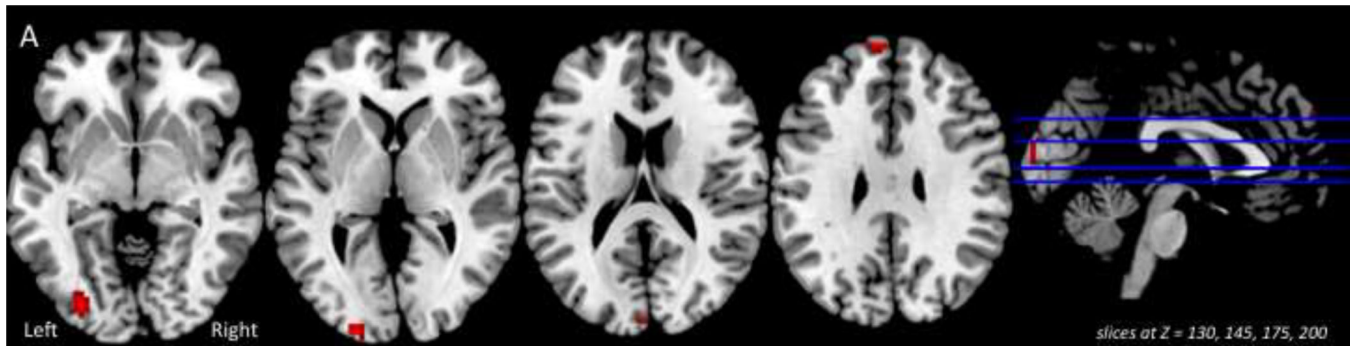
**Figure 1.** Significant Auditory phase activation related to the group  $\times$  length interaction ( $p < 0.05$  FWE) is shown on relevant axial slices of a group T1 image (Panel A). The differences in percent activation associated with increased sentence length for the preterm (green) and full term (blue) groups is shown (Panel B) at each significant cluster from the interaction. Minimum cluster size  $> 10$  voxels.



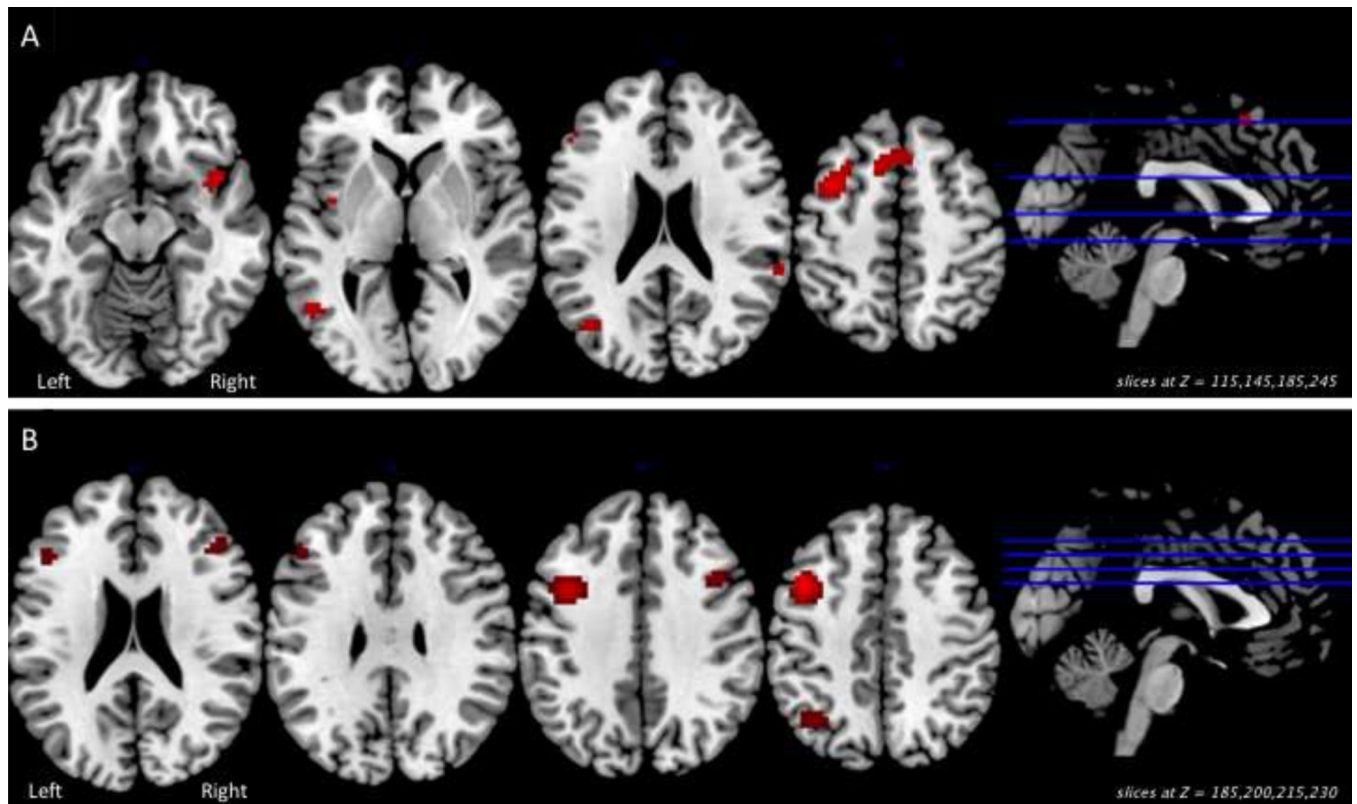
**Figure 2.** Significant Verification/Response phase activation related to the group  $\times$  length interaction ( $p < 0.05$  FWE) is shown on relevant axial slices of a group T1 image (Panel A). The difference in percent activation (Long – Short) associated with increased sentence length for the preterm (green) and full term (blue) groups is shown (Panel B) at each significant cluster from the interaction. Minimum cluster size  $>10$  voxels.



**Figure 3.** Significant Verification/Response phase activation related to the group  $\times$  difficulty interaction ( $p < 0.05$  FWE) is shown on relevant axial slices of a group T1 image (Panel A). The difference in percent activation (Difficult – Easy) associated with increased syntactic difficulty for the preterm (green) and full term (blue) groups is shown (Panel B) at each significant cluster from the interaction. Minimum cluster size  $> 10$  voxels.



**Figure 4.**  
*Post hoc* test results on the Auditory phase group  $\times$  length interaction ( $p < 0.01$  uncorrected) for the whole brain, revealing areas in which the difference in percent activation between Long and Short sentences was greater for full terms relative to preterms.



**Figure 5.**

We present *post hoc* test results from significant interactions in the Verification/Response phase: (Panel A) Group  $\times$  length interaction ( $p < 0.01$  uncorrected) for the whole brain, revealing areas in which the difference in activation between Long and Short sentences was greater for preterms relative to full terms. (Panel B) Group  $\times$  difficulty interaction ( $p < 0.01$  uncorrected) for the whole brain, revealing areas in which the difference in activation between Difficult and Easy sentences was greater for preterms relative to full terms.

Table 1

Study participant demographics.

Subject	Group	Gender	Age	GA	BW	SES	Spec Ed
20013	P	F	14.4	28	1191	3	N
20016	P	F	12.5	26	630	1	Y
20015	P	M	14.4	28	936	3	N
20054	P	M	11.6	34.5	2410	3	N
20032	P	F	10.7	34	2200	3	N
20017	P	M	11.9	28.5	1624	3	N
20008	P	M	12	27	1027	2	Y
20028	P	F	11.9	28.5	1192	1	N
20029	P	M	12.4	29	1139	3	N
20014	P	M	14.4	28	1028	3	N
20037	P	M	11.1	31	1020	2	Y
20041	P	F	10.1	27	879	3	N
20001	P	F	14.3	27	775	3	N
20042	P	F	10.1	27	879	3	N
20052	P	M	9.7	27	1134	3	N
20033*	P	M	12	28	1361	3	N
20051*	P	F	10	27	1021	3	N
20007*	P	F	13	34	1327	3	Y
20038	F	F	13	40	3000	1	N
20021	F	F	14	40	3487	3	N
20020	F	M	12	40	3884	3	N
20044	F	M	11	40	2580	1	N
20022	F	F	14	40	3317	3	N
20045	F	M	16	40	2551	1	N
20027	F	M	14	40	3515	3	N
20012	F	M	14	37	2920	2	N
20048	F	M	14	40	3771	3	N
20034	F	F	13	40	3147	3	N



Subject	Group	Gender	Age	GA	BW	SES	Spec Ed
20026	F	M	14	39	3345	3	N
20023	F	F	13	40	3714	3	N
20025	F	F	12	39	3685	3	N
20018	F	F	14	40	2807	3	N

\* = Participant in-scanner performance <70% correct (not included in analyses)

Age = age at time of testing

GA = gestational age, in weeks

BW = birth weight, in grams

SES = socioeconomic status, measured by level of maternal education (rank order scores 1–3, where 3 represents highest SES level)

**Table 2**

Examples of stimuli used in the auditory sentence verification task during functional magnetic resonance imaging (fMRI).

<b>Sentence Condition</b>	<b>Syntactic constructions</b>	<b>Examples</b>
Short/easy	• Reversible in and on	• The cup is in the box.
	• Three elements	• The dog stands on the table.
	• Reversible SVO	• The man is chasing the dog.
Short/hard	• Reversible passive	• The sheep is pushed by the boy.
	• Pronoun gender/number	• The lady is pointing at them.
	• Singular/plural inflection	• The cows are under the tree.
Long/easy	• Four elements	• The horse sees the cup and the book.
	• Relative clause in subject	• The man that is eating looks at the cat.
	• Not only X but Y	• The man is not only running but also pointing.
Long/hard	• Pronoun binding	• The man sees that the boy is pointing at him.
	• Relative clause in object	• The girl chases the dog that is jumping.
	• Center-embedded sentence	• The scarf the book is on is blue.

**Table 3**

All participants' test scores (assessed offline) and performance on the auditory sentence verification task (in the scanner).

Measure	Preterm		Full-Term		t-test	p
	Mean	SD	Mean	SD		
<i>Standardized test score</i>						
WASI Full IQ	109.8	18	121.8	12.8		.34
PPVT-III	116.3	18.5	116.2	9.3		.45
CELF-IV receptive language	109.1	17	114.6	10.2		.52
<i>In-scanner sentence comprehension</i>						
Mean # correct (maximum 24)						
Short/easy	22.5	1.2	22.2	1.5		.21
Short/hard	20.9	2.2	20.5	1.4		.31
Long/easy	21.5	2.1	20.9	2.8		.30
Long/hard	19.7	2.9	19.0	2.0		.30
<i>Mean RT (msec)</i>						
Short/Easy	1738	401	1851	377		.21
Short/hard	1923	428	2003	408		.10
Long/easy	1931	410	2169	422		.01
Long/hard	2521	418	2670	500		.05

Note: Independent-samples t-test *p*-values for in-scanner performance are included at the request of a reviewer, but should be adjusted for multiple comparisons. In the text, we report results from ANOVAs for these measures.

WASI = Wechsler Abbreviated Scale of Intelligence.

PPVT-III = Peabody Picture Vocabulary Test, Third edition.

CELF-IV = Clinical Evaluation of Language Fundamentals, Fourth edition

**Table 4**

Details regarding statistically significant activation in cortical areas during the Auditory phase of the sentence verification task that correspond to the following statistical tests: (a) Interaction of difficulty  $\times$  length, (b) *Post hoc* analysis of difficulty greater than length, (c) *Post hoc* analysis of length greater than difficulty, (d) Interaction of group  $\times$  length, and (e) *Post hoc* analysis, full terms greater than preterms. Cluster size  $>10$  contiguous voxels. Statistical significance for interactions set at  $p < 0.05$  FWE correction for multiple comparisons; Significance level for *post hoc* analyses set at  $p < .01$  uncorrected.

Effect and Cortical area	Peak Z-score	Number of active voxels	MNI Coordinates (x, y, z)
<i>a. Interaction, difficulty <math>\times</math> length</i>			
Right superior temporal gyrus	5.43	68	63 -21 6
<i>b. Post hoc, difficulty greater than length</i>			
Left and right occipital lobes	4.51	458	-12 -87 -3
Left middle frontal gyrus	3.63	474	-33 45 10
Left precuneus	3.50	334	0 -60 27
Left middle temporal gyrus	2.68	24	-45 -66 15
<i>c. Post hoc, length greater than difficulty</i>			
Right superior temporal gyrus	5.53	622	63 -21 6
Left superior temporal gyrus	4.97	502	-54 -15 6
<i>d. Interaction, group <math>\times</math> length</i>			
Right superior temporal gyrus	7.49	413	63 -21 6
Left superior temporal gyrus	6.91	266	-57 -18 6
Left occipital lobe	5.21	29	-18 -96 -6
<i>e. Post hoc, full terms greater than preterms</i>			
Left occipital lobe	3.39	18	-21 -99 0
Right occipital lobe	2.99	25	3 -96 9
Left medial frontal gyrus	2.96	13	-6 60 30

**Table 5**

Details regarding statistically significant activation in cortical areas during the Verification/Response phase of the sentence verification task that correspond to the following statistical analyses: (a) Interaction of difficulty  $\times$  length, (b) *Post hoc* analysis, difficulty greater than length, (c) *Post hoc* analysis, length greater than difficulty, (d) Interaction of group  $\times$  length, (e) *Post hoc* analysis, preterms greater than full terms, (f) Interaction of group  $\times$  difficulty, and (g) *Post hoc* analysis, preterms greater than full terms. Cluster size  $>10$  contiguous voxels. Statistical significance for interactions set at  $p < 0.05$  FWE correction for multiple comparisons; Significance level for *post hoc* analyses set at  $p < .01$  uncorrected.

Phase and Cortical area	Peak Z-score	Number of active voxels	MNI Coordinates (x, y, z)
<i>a. Interaction, difficulty <math>\times</math> length</i>			
Right superior temporal gyrus	7.21	274	66 -22 4
Right occipital lobe	6.31	181	45 -73 4
Right precuneus	5.79	91	3 -58 31
Left inferior frontal gyrus	5.71	236	-51 -28 4
Left occipital lobe	5.55	92	-21 -100 10
<i>b. Post hoc, difficulty greater than length</i>			
Left and Right occipital lobe	6.51	1,707	45 -73 4
Right precuneus	5.85	476	3 -58 31
Left medial frontal gyrus	4.20	407	0 56 31
<i>c. Post hoc, length greater than difficulty</i>			
Right superior temporal gyrus	7.21	2,051	66 -22 4
Left superior temporal gyrus	5.75	2,184	-51 -28 4
Right cerebellum	5.60	377	27 -43 -14
Left cerebellum	5.33	417	-24 -52 -14
Left precentral gyrus	3.53	124	-45 2 28
Right middle frontal gyrus	3.22	42	45 38 13
Left superior temporal gyrus	2.98	36	-63 -58 25
Right medial frontal gyrus	2.93	62	12 -13 7
Left precentral gyrus	2.79	36	15 11 52
Left medial frontal gyrus	2.68	36	-18 14 52
<i>d. Interaction, group <math>\times</math> length</i>			
Right superior temporal gyrus	7.28	432	66 -22 4
Left middle temporal gyrus	7.28	695	-54 -40 4