ORIGINAL ARTICLE

Brain Functional and Structural Predictors of Language Performance

Michael A. Skeide, Jens Brauer, and Angela D. Friederici

Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Address correspondence to Michael A. Skeide, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1A, 04103 Leipzig, Germany. Email: skeide@cbs.mpg.de

Abstract

The relation between brain function and behavior on the one hand and the relation between structural changes and behavior on the other as well as the link between the 2 aspects are core issues in cognitive neuroscience. It is an open question, however, whether brain function or brain structure is the better predictor for age-specific cognitive performance. Here, in a comprehensive set of analyses, we investigated the direct relation between hemodynamic activity in 2 pairs of frontal and temporal cortical areas, 2 long-distance white matter fiber tracts connecting each pair and sentence comprehension performance of 4 age groups, including 3 groups of children between 3 and 10 years as well as young adults. We show that the increasing accuracy of processing complex sentences throughout development is correlated with the blood–oxygen-level-dependent activation of 2 core language processing regions in Broca's area and the posterior portion of the superior temporal gyrus. Moreover, both accuracy and speed of processing are correlated with the maturational status of the arcuate fasciculus, that is, the dorsal white matter fiber bundle connecting these 2 regions. The present data provide compelling evidence for the view that brain function and white matter structure together best predict developing cognitive performance.

Key words: development, DTI, fMRI, language, prediction

Introduction

The relation between brain structure and behavior has long been discussed in the context of lesion studies. More recently, imaging studies made it possible to look at the relation between brain structure and behavior on the one hand and brain functional activation and performance on the other. A core question in developmental neuroscience has been which of these brain aspects primarily constitutes emerging complex behavior, is it structural maturation before functional specialization or vice versa. When turning to brain maturation, both the gray matter (Hoeft et al. 2007; Steinbeis et al. 2012) as well as white matter (Yeatman et al. 2012) may be considered. In particular, the white matter fiber bundles connecting different brain regions are of interest as they are essential for the transmission of electrical impulses between these. Their surrounding myelin is crucial for the speed of transmission as shown by in vitro and animal studies

(Fields 2010; Wake et al. 2011). Given these basic neurophysiological findings, one would hypothesize the degree of white matter myelination to have an impact on the behavioral performance at distinct stages of development, in particular with respect to processing speed. One possibility to approach this issue in humans is to measure white matter fractional anisotropy (FA), gray matter blood–oxygen-level-dependent (BOLD) activity, and behavior, both in terms of accuracy and speed, in the very same individuals and to analyse the predictive value of the different brain measures for behavioral performance across development.

Here we approach this question by taking a higher cognitive function in humans, namely language, as a test case. We chose sentence comprehension as the critical behavioral language paradigm as the ability to process syntactically complex sentences develops slowly, but can be evaluated systematically in different age groups during development. Although children

acquire language with ease and show basic skills by the age of 3 (Markman et al. 2003; Nazzi and Bertoncini 2003; Oberecker et al. 2005), it takes a long time before they are adept with complex grammar. Even 7-year-old healthy children make striking errors when it comes to processing sentences which do not follow a simple subject-first word order, but a more complex object-first word order (Dittmar et al. 2008; Schipke et al. 2012). Although linguists and psychologists have characterized object-first sentences as syntactically more complex (Chomsky 1995; Fanselow 2001; Friedmann et al. 2009), there is no convincing explanation for this developmental phenomenon. Our current knowledge about the neural basis of developmental language disorders specifically affecting syntax processing, such as specific language impairment, is too coarse to shed light on this topic (van der Lely and Pinker 2014).

In the neurolinguistic adult literature, it was suggested that object-first sentences are more difficult to process than subject-first sentences due to increased working memory load and increased syntax-specific computational demands, particularly reordering processes in the posterior part of Broca's area within the left inferior frontal gyrus (IFG opercularis) (Caplan et al. 2000, 2002, 2008; Friederici et al. 2006; Meyer et al. 2012). The notion that this brain region plays a decisive role for complex syntax was repeatedly confirmed by patient studies, particularly by studies on agrammatic primary progressive aphasia patients (Wilson et al. 2011; Mesulam et al. 2014).

In a previous study, we have shown that the increasing performance on syntactically complex sentences goes along with an increasing functional selectivity and autonomy of the left IFG opercularis (Skeide et al. 2014). Here we further analyse the fMRI data from this study and in addition diffusion-weighted magnetic resonance imaging (dMRI) data from the same children as well as fMRI and dMRI data from an adult group to investigate which neural measure, that is, gray matter BOLD activity or white matter FA, is the better predictor for sentence comprehension performance. We propose that the late emergence of complex grammar abilities is not only subject to a particular brain functional activation pattern, but moreover, to a neurostructural constraint, namely the maturational state of a particular fiber tract that transmits information between 2 language-relevant brain regions in the left hemisphere. This is the dorsal language pathway consisting of the arcuate fasciculus (AF) that connects the IFG opercularis and the posterior part of the superior temporal gyrus (posterior STG) (Frey et al. 2008; Wilson et al. 2011). Our hypotheses arise out of 2 key considerations. The IFG opercularis and the posterior STG are known as crucial for processing complex sentences in adults (Sakai 2005; Friederici 2011, 2012) and the AF has been discussed to have 2 termination points in the frontal cortex, one that terminates in the prefrontal cortex which myelinizes early and one that terminates in Broca's area with the latter only developing late during childhood (Brauer et al. 2011, 2013; Perani et al. 2011).

We examined 80 healthy participants from 4 different age groups, 3–4 years, 6–7 years, 9–10 years, and adults that were matched for gender, handedness, intelligence quotient (IQ) and parental education. Language performance was tested in a well-established child-appropriate sentence–picture-matching task (Schipke et al. 2012). Participants were required to listen to syntactically simple subject relative clauses (SR) and complex object relative clauses (OR) and match their meaning to the correct picture from a pair. During this procedure, functional scans (functional magnetic resonance imaging [fMRI]) were acquired to measure the functional responses of language-sensitive regions. Additionally, dMRI scans were taken to specify the FA

of the fiber bundles connecting these regions within the language network. We then conducted correlational analyses to specify the relation between BOLD signal change, FA, and sentence comprehension (response accuracy and reaction times [RTs]). The factors age and verbal working memory were covaried out in these analyses to segregate the specific effects of the neural measures from more general maturational and cognitive effects.

Materials and Methods

Participants

Here we chose to test children of 3 age groups between 3 and 10 years. The motivation for this age range was mainly based on previous behavioral and electrophysiological studies. Behavioral studies indicate that by the age of 3 years children demonstrate the main aspects of syntactic knowledge such as word order and inflections in simple sentences (Hirsh-Pasek and Golinkoff 1996; Guasti 2002; Clark 2003; Höhle et al. 2006). This view is supported by electrophysiological studies (Silva-Pereyra et al. 2005). More complex sentences such as passive constructions and object-first sentences, however, are only mastered much later during development, that is between the age of 4 and 7 years (Fox and Grodzinsky 1998; Oberecker and Friederici 2006) compatible with electrophysiological evidence (Schipke et al. 2011). By the age of 10 years children still make syntactic errors during language production (Chomsky 1969) and unlike adults do not process syntax independent of semantics during language perception (Friederici 1983). Also the electrophysiological and hemodynamic responses only gradually develop toward an adult-like pattern during late childhood (Hahne et al. 2004; Skeide et al. 2014). The general finding that full syntactic abilities develop late has already been discussed in the context of neuroanatomical data indicating that BA 44 (pars opercularis) as the syntax-relevant region reaches an adult-like cytoarchitectonic asymmetry toward the left hemisphere only at the age of 11 years (Amunts et al. 2003). Thus it appears that the acquisition of syntactic knowledge and its adult-like use spans a long developmental period that we are covering here in a combined neurofunctional and neurostructural approach for the first time.

Initially, 89 children were recruited on a voluntary basis from kindergartens and primary schools in Leipzig. The children underwent a training session in a mock scanner in order to familiarize them with the experimental procedure. Handedness (modified version of the Edinburgh Handedness Inventory) (Oldfield 1971), IQ (Kaufman ABC Sequential Processing Scale), verbal working memory (Kaufman ABC Number Recall Test) (Melchers and Preuss 2009), and language comprehension skills (German version of the Test for Reception of Grammar TROG-D) (Fox 2011) were assessed for all participants. Mean IQ scores were: 3–4 year-olds: 107.53 ± 11.79, range: 88–132; 6–7 year-olds: 109.47 ± 14, range: 82–131; 9–10 year-olds: 104.55 ± 8.79, range: 92–122. No significant differences in mean IQ were detected: 3–4 versus 6–7: $t_{1,37} = 0.46$; 3–4 versus 9–10: $t_{1,37} = 0.89$; 6–7 versus 9–10: $t_{1,37} = 1.3$, $P = 0.21$. A broad range of parental education levels was represented in the sample comprising parents with a 10-year-high school degree (German Realschulabschluss) but no further professional qualification as well as parents with the highest German academic degree (German Habilitation). The majority of the parents (52.5%) had a university degree. A Kruskal–Wallis H test showed that there was no statistically significant difference in parental education between the 3 child samples ($\chi^2(2) = 0.143$, $P = 0.931$). Note that parental education data were not available for 4 children, three 6–7-year-olds and one 9–10-year-old. No participant had a history

of neurologically relevant diseases or psychiatric disorders. Of note, 83 of 89 children (handedness laterality quotient greater than or equal to +60, K-ABC and TROG-D test results within normal range as well as no history of relevant diseases) were invited for fMRI scanning. Twenty participants had to be excluded from further data analysis because fMRI data could not be collected or could not be analyzed due to too much movement during scanning. Datasets were disregarded if head motion exceeded 3 mm/TR. Three more datasets were discarded because of radiofrequency noise artifacts of unknown origin.

Initially, 26 adult participants were recruited, who were all right-handed, had normal intelligence, memory and language skills, and no history of relevant diseases. Two datasets containing movement artifacts (motion exceeding 3 mm/TR) and 2 further datasets with radiofrequency noise artifacts of unknown origin had to be excluded. Two more datasets were not included in the analysis because their behavioral data could not be recorded properly due to technical problems.

In order to rule out the possible confound that those participants that were excluded due to excessive head motion could have differed significantly with respect to sentence comprehension performance from the participants whose MRI data were of sufficient quality for analysis, we compared the corresponding age groups running Mann–Whitney U tests which did not reveal any significant differences (see Supplementary Table 1). All participants that did not generate analyzable MRI data had normal cognitive and demographic characteristics comparable with the remaining the sample.

Datasets from 4 age groups were subjected to the final fMRI analysis: 20 children 3–4 years of age (12 girls, mean: 4.4, range: 3.9–4.11), 20 children 6–7 years of age (11 girls, mean: 7.5; range: 6.7–7.11) and 20 children 9–10 years of age (8 girls, mean: 10.3, range: 9.7–10.11) and 20 adults (7 females, mean: 26.5, range: 21.8–33.6). During an informative briefing about the study, adults and parents gave written informed consent and children gave verbal assent to participate in the study. All experimental procedures were approved by the University of Leipzig Ethical Review Board.

Behavioral Data: Acquisition

While lying in the scanner, participants listened to questions that incorporated embedded relative clauses (mean duration: 3.81 s, duration range: 3.57–3.99 s) and saw 2 pictures (Fig. 1). When auditory stimulation was over, and the visual stimulus was still present, participants performed a sentence–picture-matching task in which they were given a response window of up to 2.5 s to make a decision, via button press, as to which picture showed the correct agent–patient (who does what to whom) relation. The child-appropriateness of sentence–picture-matching tasks across all investigated age groups is based on the extensive experience with illustrated books that is already established in toddlers.

In order to minimize eye movement, pictures were presented through an MR compatible eyeglass display (VisuaStim XGA, Resonance Technology Inc.). Sentences were spoken by a professional female native speaker in a unisonous and moderately child-directed prosody. Participants were instructed to attend to the pictures, listen carefully to the sentences, and respond as fast as possible as soon as auditory stimulation was over. Responses as well as RTs were recorded with Presentation (http://www.neurobs.com) in order to measure the performance of the participants.

The experiment was 15 min long and comprised 96 target trials (48 trials per condition) as well as 12 null event trials whose order was pseudorandomized for each participant. The 2 conditions of the present experiment were derived from a previously published experiment with 4 conditions (Skeide et al. 2014), crossing the factor syntactic complexity (SR vs. OR) with the factor semantic plausibility (semantically plausible vs. implausible sentences) by collapsing trials across the 2 semantic plausibility conditions to test exclusively for the effect of syntactic complexity. Each target trial started with a jitter (blank screen) of variable length (0, 0.5, 1, or 1.5 s) in order to sample several points of the hemodynamic response function. The position of the correct picture and the position of agent and patient within the picture were counterbalanced. Word frequencies of all nouns and verbs were analyzed and not significantly related neither to response accuracies nor to RTs, according to the Mannheim spoken corpora (http://celex.mpi.nl/).

In order to ensure that all children understood the task, they were carefully familiarized with it during the prescanning training session. Additionally, all responses given during fMRI scanning were supervised online with a button press monitoring device to ensure regular controlled button presses by all participants. Most importantly, the evaluation of the mean response accuracy rates revealed that all 3- to 4-year-old children performed significantly above chance at least in the simple syntax condition.

The task used in the experiment targets the assignment of thematic roles for which it is obligatory to process word order and case marking information whereas semantics provides no additional information that is obligatory to solve the task. Accordingly, it is considered unlikely that the syntactic manipulation used in our experiment could be confounded by semantics.

Subject relative clauses (simple)

Wo ist **der große** Fuchs, **der den kleinen** Käfer trägt?

*Where is **the big** fox, [**who**]_{NOM} [**the**]_{ACC} **small** beetle carries?*
Where is the big fox, who carries the small beetle?

Object relative clauses (complex)

Wo ist **der kleine** Käfer, **den der große** Fuchs trägt?

*Where is **the small** beetle, [**who**]_{ACC} [**the**]_{NOM} **big** fox carries?*
Where is the small beetle, who the big fox carries?

Figure 1. Example sentences and a corresponding set of 2 pictures that were presented simultaneously during the sentence–picture-matching task. The participants were required to select the picture showing the correct agent–patient relation (who does what to whom), which was inverted in the incorrect picture.

Behavioral Data: Analysis

As a first step, we tested separately for each group if the mean response accuracy values (accuracy) and the mean RTs, respectively, were normally distributed. Kolmogorov–Smirnov (K–S) and Shapiro–Wilk (S–W) tests on accuracy carried out using SPSS 19 (IBM Corp.) did reveal significant results for the 6- to 7-year-old children (SR: K–S: $P = 0.012$, S–W: $P = 0.022$; OR: K–S: $P = 0.043$, S–W: $P = 0.067$), for the 9- to 10-year-old children (SR: K–S: $P = 0.001$, S–W: $P < 0.001$; OF: K–S: $P = 0.025$, S–W: $P = 0.001$), and for the adults (SR: K–S: $P = 0.001$, S–W: $P < 0.001$; OF: K–S: $P = 0.001$, S–W: $P < 0.001$). Normality tests on the RTs did not reveal any significant results.

Two 4 (AGE GROUP) × 2 (SYNTAX) analyses of variance (ANOVAs) were computed in order to draw conclusions on the level of between-group inferential statistics; a nonparametric one on the ranked mean accuracy rates and a parametric one on the mean RTs. Post hoc, 4 separate ANOVAs were run on each sample in order to assess within-group effects driving the AGE GROUP × SYNTAX interaction revealed by accuracy rates. Within-group differences between conditions were calculated post hoc using paired *t*-tests. The *P* values of these post hoc analyses were Bonferroni corrected for the 4 tests. Only trials in which a response (correct or incorrect) was given were included into the analyses of the response accuracy means. For the assessment of the RTs, only correct answers were modeled.

FMRI Data: Acquisition

The experiment was carried out on a 3.0-T Siemens TIM Trio (Siemens AG) whole-body magnetic resonance scanner using a 12-radiofrequency-channel head coil. With the goal of investigating brain functional activity on the basis of a scanning protocol sensitive to BOLD contrasts, a T_2^*-weighted gradient-echo echo-planar imaging (EPI) sequence was applied to 26 slices with time repetition (TR) = 2 s, time echo (TE) = 30 ms, field of view (FOV) = 192 mm, matrix size = 64 × 64 voxels and voxel size 3 × 3 × 3 mm³. In order to correct for geometric distortions in EPI caused by magnetic field inhomogeneity, a field map was obtained for each dataset. For anatomical localization, T_1-weighted 3-dimensional magnetization-prepared rapid-acquisition gradient echo (MPRAGE) pulse sequences with TR = 1.480 ms, TE = 3.46 ms, TI = 740 ms, FOV = 256 × 240, matrix size = 256 × 240 × 128 and voxel size = 1 × 1 × 1.5 mm³ were acquired.

FMRI Data: Preprocessing and Second-level Analysis

Functional images were preprocessed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). First, a cubic spline interpolation algorithm was applied to the time series of individual slices to correct for time differences between slices recorded within the same scan and to resample them afterward (slice time correction). Second, images were realigned (i.e., spatially registered and transformed) to the first acquired image to correct for movement between scans, and then unwarped to correct for distortions caused by magnetic field inhomogeneities and interpolation artifacts (motion correction). Third, low-resolution functional images of each participant were mapped (i.e., coregistered) onto the corresponding high-resolution T_1-weighted structural images and subsequently normalized to a Montreal Neurological Institute (MNI) template covering the whole age range (Fonov et al. 2011) (spatial normalization) in order to provide a common space for the group comparisons. We ensured that the scaling did not affect the functional data (O'Shaughnessy et al. 2008; Yoon et al. 2009) by also normalizing them to MNI templates exactly

matching the group-specific mean age and afterward comparing both volumes against each other: This did not reveal any significant differences. Finally, data were spatially low-pass filtered using a 4 mm full width half-maximum (FWHM) Gaussian kernel. A temporal high-pass filter with a cutoff frequency of 1/120 Hz was applied in order to remove low-frequency drifts within voxel time series (spatial and temporal smoothing).

For the statistical analysis, functional whole-brain data were passed through a general linear model in order to run a least-squares parameter estimation (regression analysis), as implemented in SPM8. Design matrices were created on the basis of a hemodynamic response function and its derivatives. In order to control for head motion, the realignment parameters were included as regressors into the model.

The 3 × 3 × 3 mm³ voxel size of the fMRI data could potentially have led to significantly more partial volume effects in the cortical activation area of the younger children compared with the older children and the adults if their cortical thickness was <3 mm. A recent study, in which cortical thickness development was analyzed longitudinally in infants from birth to 18 months of age, revealed an average cortical thickness well >4 mm in 18-months-old infants in the perisylvian language areas investigated in our experiment (Li et al. 2014). Furthermore, the authors did not detect any significant decreases in cortical thickness during this developmental period suggesting that macrostructurally relevant pruning processes in the gray matter did not yet start at 18 months of age. The second study currently available suggests a cortical thickness clearly >3 mm in children aged 5 years and no decrease during further development but only increase in diameter in the same areas (Sowell et al. 2004). It is therefore considered unlikely that cortical thickness in perisylvian areas is below 3 mm during the late fourth and fifth year of life, the developmental period that is most relevant for the youngest sample (mean age: 4.4, range: 3.9–4.11). Accordingly, it can be assumed that the diameter of the cortex fully covered an entire voxel dimension in these children.

First-level contrast images were computed for all 4 conditions against the baseline (null events) and then passed to a one-sample *t*-test separately for each group. The resulting group baseline activation maps were thresholded at $P < 0.001$ (cluster size corrected to $P < 0.001$). We determined the probability of a false detection by dual thresholding of both type I error and cluster size; running 10 000 iterations of a Monte Carlo Simulation using AFNI AlphaSim (http://afni.nimh.nih.gov/afni). Only clusters extending 11 voxels were considered significant, minimizing the probability of false detection to a maximum of 0.001 for uncorrected α rates of 0.001 ($z > 3.09$), given that the voxels size was 3 × 3 × 3 mm³ and voxels were smoothed with a 4 mm FWHM kernel.

FMRI Data: Region-of-Interest Analysis

A region-of-interest (ROI) analysis was conducted using the MarsBar toolbox (http://marsbar.sourceforge.net/). In order to compute the mean percent signal changes for each group and condition, β-weights and time courses were extracted from 4 spherical ROIs ($r = 6$ mm) covering the brain regions most frequently associated with syntactic processing (IFG opercularis: −40, 14, 24; IFG triangularis: −51, 30, 6; anterior STG: −47, 14, −28; posterior STG: −58, −36, 2). The percent BOLD signal change values were normally distributed with the exception of the posterior STG in the 9- to 10-year-old children for the OR (K–S: $P = 0.005$; S–W: $P = 0.021$) and the anterior STG in the adults for the OR (K–S: $P = 0.004$; S–W: $P = 0.017$). However, Levene's tests of

equality of error variances were not significant (posterior STG: $F_{3,76} = 2.14$, $P = 0.102$; anterior STG: $F_{3,76} = 0.26$, $P = 0.853$).

The finding that the signal change values in the anterior STG are negative in the child samples might be explained by a very common caveat of fMRI analyses, namely by distortions of the signal near the sinuses which potentially degrades signal detection in the anterior temporal cortex (Bonner and Price 2013). Note, however, that although the absolute signal intensity might be altered, the relative intensity values should be preserved. Taking this into account, the validity of the argument that there is only little developmental dynamics in this ROI with respect to complex syntax is not compromised.

The task robustly activated not only the left posterior STG but also the left IFG opercularis in all age groups although differences in spatial extent remain. When lowering the threshold from $P < 0.001$ to $P < 0.005$ in the 3- to 4-year-old children, the left frontal activation cluster fully covers the IFG opercularis (see Supplementary Fig. 1). In order to rule out a potential misidentification of BOLD activity especially in the younger children, we inspected the variability of individual activation maxima for each participant and noticed no systematic differences (see Supplementary Tables 2 and 3).

In the analysis of the fMRI results presented in Figure 3 and Table 2, mean percent BOLD signal changes were compared separately for each ROI computing ANOVAs. Three models were set up; 2 models, in which SR and OR were treated as single separate factors and one model, in which both conditions were modeled as levels of one factor. Accordingly, P values were Bonferroni corrected for the 3 tests for each ROI. Within-group comparisons of BOLD signal changes between SR and OR sentences were carried out running paired-samples t-tests and Bonferroni correcting the resulting P values for the 4 tests per ROI. In all following correlation analyses including the fMRI results, BOLD signal changes of the 2 dorsally connected ROIs (IFG opercularis and posterior STG) were added together and, similarly, BOLD signal changes of the 2 ventrally connected ROIs (IFG triangularis and anterior STG) were added together. Correlations between BOLD signal change and sentence comprehension performance of the separate ROIs can be found in Supplementary Table 4.

DTI Data: Acquisition

Forty-six out of 60 children and 19 out of 20 adults who had already taken part in the fMRI study also participated in the DWI study. The time limit for participation was set to 6 weeks after the fMRI scanning. Elevan children and one adult could not be invited for the second experimental session within this time frame. Three participants were removed from the study because they exceeded the cutoff quality criterion of maximum 5 head-motion-corrupted image directions in the entire dataset (Brauer et al. 2013). We acquired 12 datasets from the 3- to 4-year-old children (8 girls, mean: 4.7, range: 3.10–4.11), 17 datasets from the 6- to 7-year-old children (9 girls, mean: 7.3, range: 9.9–7.11), 17 datasets from the 9- to 10-year-old children (7 girls, mean: 7.6, range: 6.7–7.11), and 19 datasets from the adult participants (7 female, mean: 26.4, range: 21.9–33.8). During the overall experiment time of ~20 min, participants were watching a video.

dMRIs were collected with the same hardware as the functional MR images using a twice-refocused spin EPI sequence (Reese et al. 2003) with TE = 100 ms, TR = 9300 ms, matrix size = 128 × 128 voxels, voxel size = 1.7 × 1.7 × 1.7 mm³, 65 axial slices covering the whole brain. We used 60 isotropically distributed diffusion-encoding gradient directions with a b-value = 1000 s/mm². Seven anatomical reference $b0$ images without diffusion weighting were acquired once at the beginning of the sequence and after each block of 10 diffusion-weighted images for off-line motion correction. Fat saturation was applied together with 6/8 partial Fourier imaging and generalized auto-calibrating partially parallel acquisitions with acceleration factor = 2 (Griswold et al. 2002). Random noise in the data was reduced by averaging 2 acquisitions.

DTI Data: Analysis

T_1-weighted structural images were skull-stripped and coregistered into Talairach space (Talairach and Tournoux 1988) using the LIPSIA software (http://www.cbs.mpg.de/institute/software/lipsia). We used the 14 images without diffusion-weighting ($b0$) to estimate the motion correction parameters applying a rigid-body registration algorithm (Jenkinson et al. 2002) implemented in FSL (http://www.fmrib.ox.ac.uk/fsl). Head-motion-corrupted image directions (max. 5) were removed from each dataset. Motion correction for diffusion-weighted images was combined with a global registration to the individual T_1-weighted structural image, and the gradient direction for each volume was corrected with the rotation parameters. In the registration process, the images were interpolated to the new reference frame with an isotropic voxel resolution of 1 mm and then a voxel-wise diffusion tensor was fitted to the datasets, and the FA was computed.

We used the 2 ROIs in Broca's area (IFG opercularis: –40, 14, 24; IFG triangularis: –51, 30, 6) as seeds for the tractography. The coordinates were projected onto the individual T_1-weighted structural images by applying the transformation obtained by inverse linear and nonlinear registration of the individual anatomical images to an MNI template covering the whole age range (Fonov et al. 2011) using FSL. To assure a robust seeding of the tractography in the white matter, the registered fMRI activations were projected onto the centerline of the closest white matter gyrus. Therefore, the individual FA maps were thresholded at 0.25 and then reduced to an individual white matter skeleton representing the center lines of the gyri (Smith et al. 2006), as implemented in FSL. The white matter voxel closest to the group activation coordinates served as the center-seed point for the fiber tracking. To improve the robustness of the tracking, the seed region was extended to the intersection between a sphere of radius of 5 mm around the seed point and the white matter skeleton.

Probabilistic fiber tracking was performed by computing a series of random walks (number of repetitions fixed by Monte Carlo simulations), starting from a seed point and running through surrounding voxels in order to obtain a probabilistic map based on this trajectory. Random walks are constrained by the orientation of the diffusion tensor in the seed voxel and the surrounding voxels. The number of times the random walk crossed a specific voxel determined the connectivity strength between the seed area and the specific white matter voxel (Anwander et al. 2007). One lakh random walks were started in the selected seed region. The connectivity values corresponded to the number of tracks per voxel and ranged between zero and 100 000. To reduce the dynamic range, a logarithmic transformation was applied to these values followed by a scaling between zero and one. To average the connectivity maps, the resulting single-subject connectivity maps were then normalized onto the most typical subject's map for each group. Therefore, a nonlinear deformation field to a target image was computed (Thirion 1998) and applied to the single-subject connectivity maps. The target subjects were identified by nonlinear alignment of each individual anatomical image (represented by the FA map) to every other one in order

to compute for which image spatial transformation to all remaining images within the group was minimal (Smith et al. 2006). Finally, the aligned connectivity maps were averaged for each group and thresholded at 0.25 for structural analysis and visualization. This tract reconstruction approach minimizes the possible influence of partial volume effects of adjacent CSF and other neighboring structures as well as crossing fibers since the resulting tractograms robustly represent the centers of all individual tractograms sparing their outermost boarders.

For quantitative analysis of the microstructural properties in the different fiber bundles connecting the seed areas, we measured the FA values in each individual brain within the single-subject tractography maps. The FA values for all voxels with a connectivity value >0.25 (equivalent to >18 random walks) were averaged for each tract and participant. The resulting mean FA values were normally distributed according to K–S- and S–W tests. FA values were compared separately for each tract by computing ANOVAs. Between-group differences were determined running independent-samples t-tests and Bonferroni correcting the resulting P values for the 3 tests per tract.

Correlational Analyses

Linear associations between BOLD signal change, FA as well as mean response accuracy and mean RT were assessed either running Pearson correlations or Spearman's rank correlations (for accuracy). Adjusted correlations were computed based either on partial Pearson correlations or partial Spearman's rank correlations (for accuracy). Age and verbal working memory skills, that is, digit span scores, were included as covariates of no interest in these models. All P values were Bonferroni corrected for the 3 tests for each pair of associations. We compared the differences between the corresponding correlation coefficients using either Fisher's r to z test (Dunn and Clark 1969) or Meng's z test (Meng et al. 1992). Fisher's r to z test is a statistically valid algorithm that enables testing for significant differences between correlated correlation coefficients that are independent of each other, that is, that do not share a variable they were correlated against, using Fisher's r to z transformation and accounting for the sample size employed to obtain each coefficient. Meng's z test is a statistically valid extension of Fisher's r to z test to dependent correlated correlation coefficients that have one variable in common.

Results

Sentence comprehension improved with age showing increased response accuracy and decreased RTs (Table 1). For the 2 youngest age groups, accuracy on the more complex OR was significantly worse than on simple SR (3–4: $t_{1,19} = 2.37$, $P_{corr} = 0.026$; 6–7: $t_{1,19} = 3.98$, $P_{corr} = 0.004$) (Fig. 2).

The functional brain imaging data showed that all groups activate the perisylvian cortex (left > right hemisphere) as well as the occipital cortex (Fig. 3a). ROI analyses were conducted for 4 left-hemispheric ROIs known to be major parts of the adult language network (Friederici 2011, 2012) and also shown to be active during sentence processing in children (Brauer et al. 2011; Nuñez et al. 2011): 2 dorsally connected regions relevant for complex syntax (Newman et al. 2010), and 2 ventrally connected regions supporting language meaning used as internal control regions within the language network (Humphries et al. 2006; Pallier et al. 2011). The syntax-relevant regions (Fig. 3b,c), showed a stronger BOLD activation than the ventrally connected control regions (Fig. 3d,e) and a selective increase across age (Table 2).

Table 1 Between-group effects at the behavioral level

|	Accuracy	Reaction time
Main effect of syntactic complexity	$F_{3,76} = 17.91$, $P < 0.001^*$	$F_{3,76} = 9.62$, $P = 0.003^*$
Main effect of age group	$F_{3,76} = 40.48$, $P < 0.001^*$	$F_{3,76} = 15.34$, $P < 0.001^*$
Interaction age group × syntactic complexity	$F_{3,76} = 5.70$, $P = 0.001^*$	$F_{3,76} = 0.22$

Note: The developmental disparity between the 4 age groups was evident in a significant age group × syntactic complexity interaction for accuracy ($F_{3,76} = 5.70$, $P = 0.001$) and in significant main effects of age group ($F_{3,76} = 40.48$, $P < 0.001$) and of syntactic complexity ($F_{3,76} = 9.62$, $P = 0.003$). RTs mirrored the general syntax performance increase across age reflected in significant main effects of age group ($F_{3,76} = 15.34$, $P < 0.001$) and of syntactic complexity ($F_{3,76} = 9.62$, $P = 0.003$). Asterisks indicate significant effects. In order to assess within-group effects driving the AGE GROUP × SYNTAX interaction revealed by accuracy rates, the procedure was applied again to each of the 4 samples separately. We found a significant main effect of SYNTAX after Bonferroni correction for the 4 tests for the 3- to 4-year-old children ($F_{1,19} = 5.6$, $P_{corr} = 0.029$) and for the 6- to 7-year-old children ($F_{1,19} = 15.84$, $P_{corr} = 0.003$), but neither for the 9- to 10-year-old children ($F_{1,19} = 0.03$) nor for the adults ($F_{1,19} = 0.31$).

Using probabilistic fiber tracking we reconstructed 2 white matter fiber bundles connecting the language-relevant areas: the dorsal tract, that is, the AF, seeding in the IFG opercularis and a ventral internal control tract, that is, the inferior fronto-occipital fasciculus (IFOF), seeding in the IFG triangularis. The 2 seeds were placed at the very same coordinates as those used in the functional ROI analysis. The FA values were found to progressively increase with age for both tracts (AF: $F_{3,62} = 24.24$, $P_{corr} < 0.001$; IFOF: $F_{3,62} = 30.41$, $P_{corr} < 0.001$) (Fig. 4a–c). In the AF, FA values differed significantly between 3–4 and 6–7 ($t_{1,27} = 2.55$, $P_{corr} = 0.019$), 6–7 and 9–10 ($t_{1,32} = 3.16$, $P_{corr} = 0.004$) and 9–10 and adults ($t_{1,34} = 2.22$, $P_{corr} = 0.035$). In the IFOF, FA values differed significantly between 6–7 and 9–10 ($t_{1,32} = 4.17$, $P_{corr} < 0.001$).

Correlational analyses covarying out the factors age and verbal working memory were conducted to test whether BOLD signal changes within the fronto-temporal language areas or the FA of the particular fiber tract connecting them determines the late development of full grammatical skills. In these analyses, the 2 dorsally connected regions, that is, the left IFG opercularis and the left posterior STG, were combined and, similarly, the 2 ventrally connected regions, that is, the left IFG triangularis and the left anterior STG, were combined. The rationale behind this approach was to investigate the effects of these regions within each functional language network as a whole. The functional activations of the dorsally connected regions were found to be more strongly correlated with the accuracy on syntactically complex OR than that of the ventrally connected regions ($z = 1.97$, $P_{corr} = 0.049$). The BOLD signal change in the dorsally connected regions was also more strongly related to syntax performance accuracy when additionally adjusting for the activation of the ventrally connected regions compared to the ventrally connected regions adjusted for the activation of the dorsally connected regions ($z = 4.03$, $P_{corr} < 0.001$) suggesting a unique contribution of the IFG opercularis and posterior STG to the processing of complex syntax. Crucially, the correlation of the dorsally connected regions and accuracy on OR was significantly stronger than with SR ($z = 2.06$, $P_{corr} = 0.039$) (Table 3). Correlations between BOLD signal change and sentence comprehension performance of the separate age groups are shown in Supplementary Table 5 and a corresponding discussion is provided as Supplementary Material).
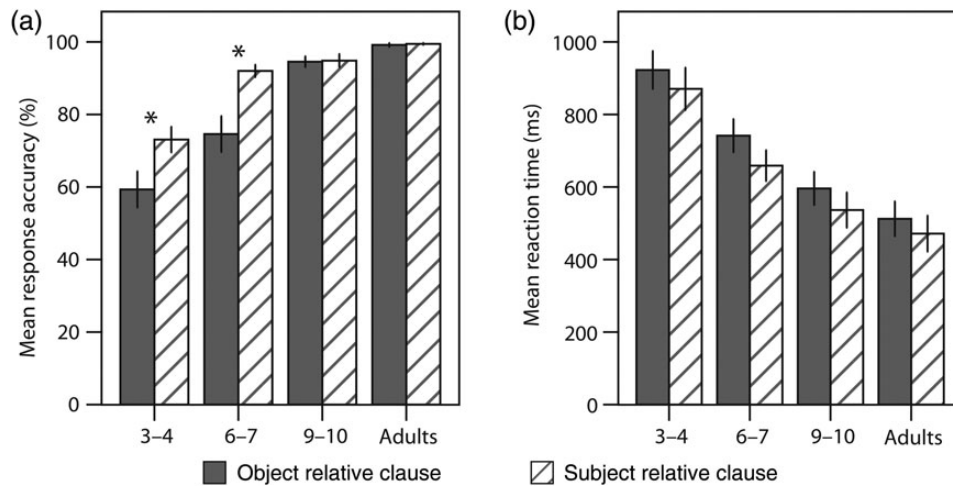
Figure 2. Behavioral results. (*a*) Mean response accuracy (%) and (*b*) mean RTs in milliseconds for the 4 groups: 3–4-year-olds, 6–7-year-olds, 9–10-year-olds and adults. The 3–4-year age group performed poorly, but nevertheless above chance (mean response accuracy: 59.33% for OR and 74.59 for SR. The 6–7-year-olds performed with a mean response accuracy of 73.08% for OR and 92.01% for SR. The 2 youngest groups showed a significantly better accuracy on the syntactically simple sentences than for those with complex syntax as indicated by the asterisks (3–4: $t_{1,19} = 2.37$, $P_{corr} = 0.026$; 6–7: $t_{1,19} = 3.98$, $P_{corr} = 0.004$). For the 9–10-year-old children and adults, accuracy was generally high (>90%) with no difference as a function of syntactic complexity observed within group (9–10: $t_{1,19} = 0.17$; Adults: $t_{1,19} = 0.56$). All P values are Bonferroni corrected. Error bars indicate standard errors of the means (SEM).
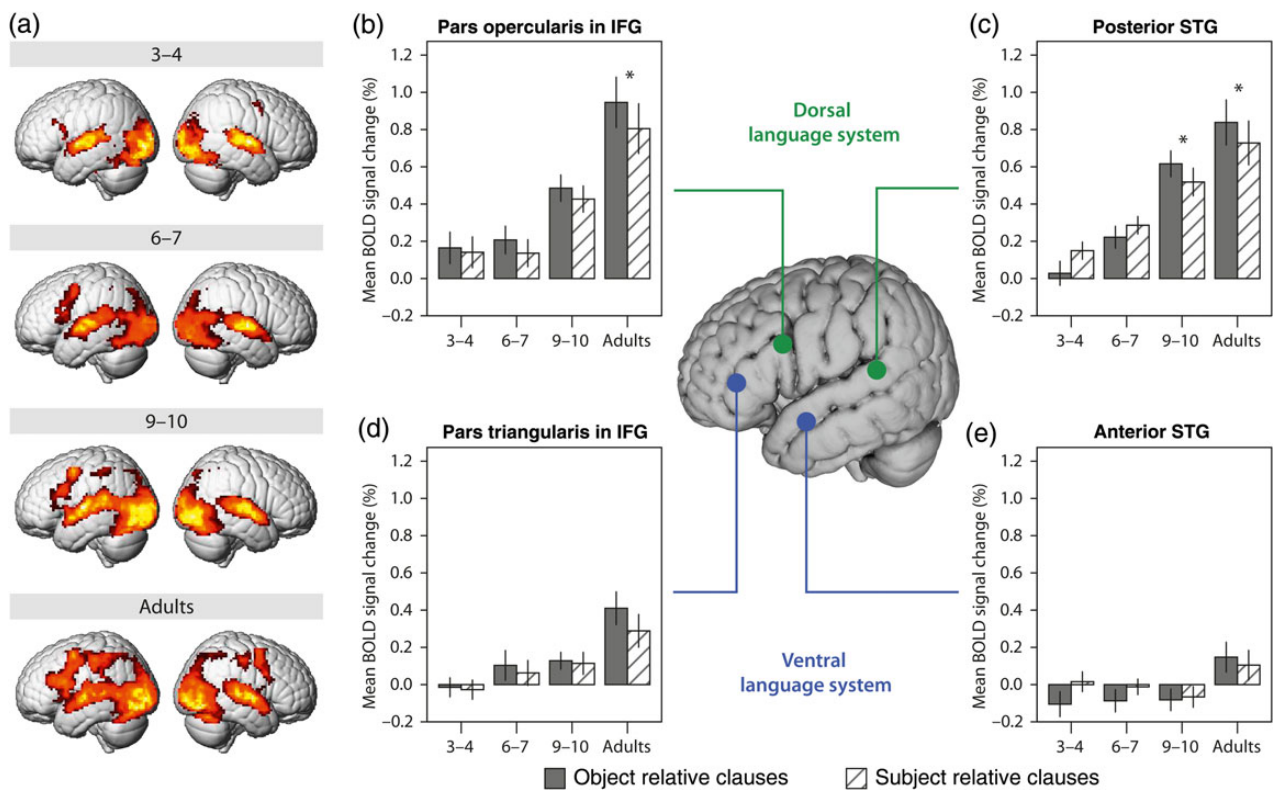
Figure 3. Functional activation at the whole-brain level (*a*) and in 4 syntax-relevant ROIs (*b*–*e*). (*a*) BOLD responses to sentence processing for all 4 age groups contrasted against baseline (null events) and thresholded at $P < 0.001$ (cluster size corrected). (*b*–*e*) BOLD responses to simple SR and complex OR in the 4 age groups. The 2 dorsally connected ROIs (*b*,*c*) revealed main effects of both SR and OR against baseline (IFG opercularis: $F_{3,76} = 11.02$, $P_{corr} < 0.001$ (SR); $F_{3,76} = 13.89$, $P_{corr} < 0.001$ [OR]; posterior STG: $F_{3,76} = 10.55$, $P_{corr} < 0.001$ (SR); $F_{3,76} = 19.25$, $P_{corr} < 0.001$ [OR]) with the IFG opercularis showing an exclusive effect for the difference between both sentence types ($F_{3,76} = 12.17$, $P_{corr} = 0.004$). In the ventrally connected ROIs (*d, e*), only a main effect of OR was detected in IFG triangularis ($F_{3,76} = 6.59$, $P_{corr} = 0.006$). Asterisks indicate significant differences between conditions. Error bars indicate standard errors of the means (SEM).

The FA values showed significantly stronger correlations of the dorsal pathway with syntax performance than those of the ventral pathway for speed (SR: $z = 2.02$, $P_{corr} = 0.044$; OR: $z = 2.1$, $P_{corr} = 0.036$) but not for accuracy (SR: $z = 1.01$; OR: $z = 1.01$). Crucially, the FA of the dorsal tract was significantly related to accuracy and RTs on OR when adjusting for the FA values of the ventral

Table 2 BOLD main effects in the 4 left-hemispheric ROIs

	IFG opercularis	Posterior STG	IFG triangularis	Anterior STG
SR	$F_{3,76} = 11.02$, $P_{corr} < 0.001^*$	$F_{3,76} = 10.55$, $P_{corr} < 0.001^*$	$F_{3,76} = 3.65$	$F_{3,76} = 1.38$
OR	$F_{3,76} = 13.89$, $P_{corr} < 0.001^*$	$F_{3,76} = 19.25$, $P_{corr} < 0.001^*$	$F_{3,76} = 6.59$, $P_{corr} = 0.006^*$	$F_{3,76} = 3.08$
OR versus SR	$F_{3,76} = 12.17$, $P_{corr} = 0.004^*$	$F_{3,76} = 0.05$	$F_{3,76} = 7.05$	$F_{3,76} = 2.34$

Note: Both regions within the dorsal language pathway (IFG opercularis and posterior STG) revealed significant activation differences between groups for both OR and SR. Furthermore, a significant activation difference between the simple and the complex sentences (OR vs. SR) was detected in the IFG opercularis but not the posterior STG. This effect was most strongly driven by the adult sample since they showed a significantly higher activation for the complex sentences than for the simple sentences in the pars opercularis ($t_{1,19} = 4.27$, $P_{corr} = 0.002$) whereas the other groups did not (3–4: $t_{1,19} = 0.79$; 6–7: $t_{1,19} = 1.11$; 9–10: $t_{1,19} = 1.62$). In the posterior STG, significant differences (OR vs. SR) were found for the adults ($t_{1,19} = 6.86$, $P_{corr} < 0.001$) and 9- to 10-year-olds ($t_{1,19} = 3.05$, $P_{corr} = 0.02$) but not for the 2 younger age groups (3–4: $t_{1,19} = 1.82$; 6–7: $t_{1,19} = 0.98$). The pars triangularis of the inferior frontal gyrus (IFG triangularis) as one part of the ventral language pathway did only show activation differences between groups for OR but neither for SR nor for OR versus SR. The anterior part of the STG (anterior STG) as the other part of the ventral language pathway did not show any activation differences across the groups. Additionally, for the ventral pathway system no OR versus SR activation difference was found, in none of the age groups. All P values are Bonferroni corrected. Asterisks indicate significant effects.
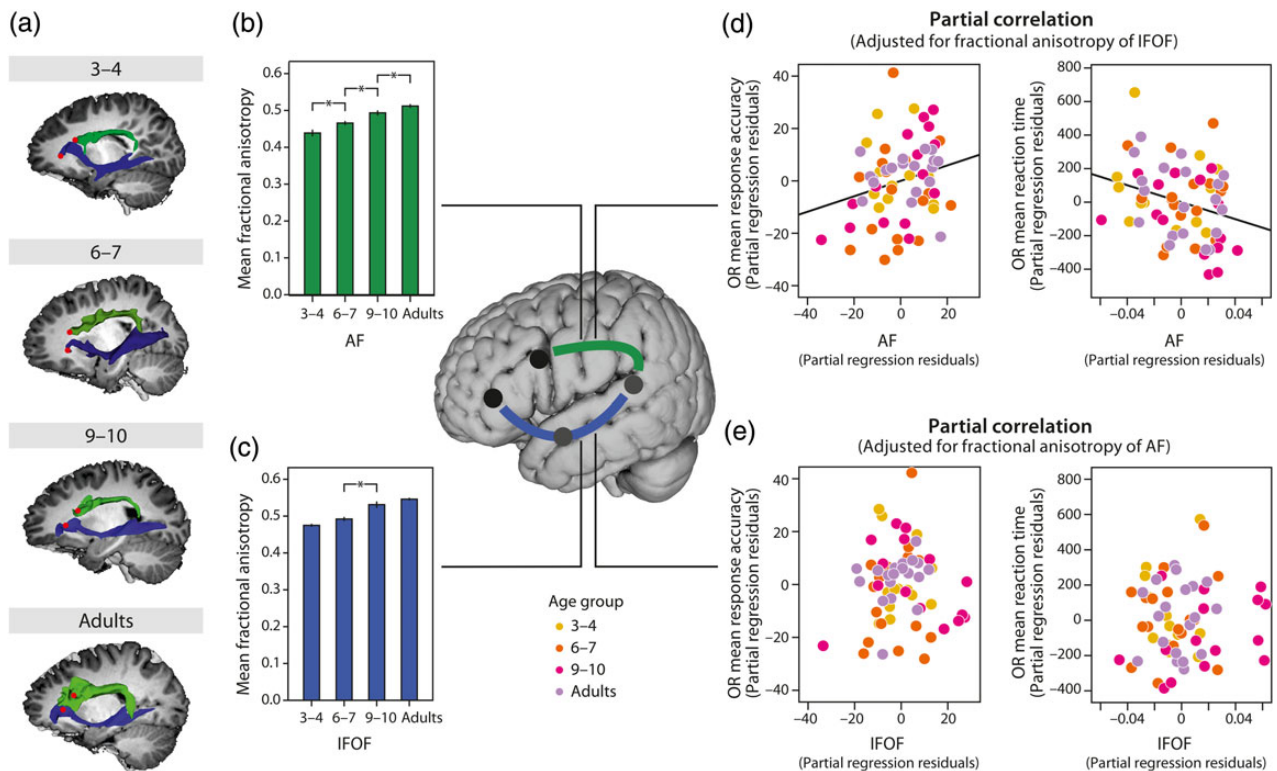
Figure 4. The maturation of the dorsal and ventral fiber pathways (a–c) and their involvement in sentence processing (d,e). (a) Seeding in the left IFG opercularis we reconstructed the left AF (green) and seeding in the left IFG triangularis we reconstructed the left IFOF (blue). Depicted are 3D renderings of all tracts for each age group. (b,c) Mean FA profiles of both tracts differing significantly between 3–4 and 6–7 ($t_{1,27} = 2.55$, $P_{corr} = 0.019$), 6–7 and 9–10 ($t_{1,32} = 3.16$, $P_{corr} = 0.004$) and 9–10 and adults ($t_{1,34} = 2.22$, $P_{corr} = 0.035$) in the AF and between 6–7 and 9–10 ($t_{1,32} = 4.17$, $P_{corr} < 0.001$) in the IFOF. (d,e) The individual FA values were significantly more strongly related to the individual OR performance in the dorsal, but not the ventral tract when adjusting for the effect of the respective other tract (RT: $r = -0.23$, $P_{corr} = 0.045$ vs. RT: $r = 0.09$, $P = 0.493$; Accuracy: $r = 0.23$, $P_{corr} = 0.047$ vs. $r = 0.025$, $P = 0.854$). Asterisks indicate significant differences between age groups. Error bars indicate standard errors of means (SEM).

tract (Fig. 4d) but not vice versa (Fig. 4e) and the correlations differed significantly from each other (RT: $z = 3.21$, $P_{corr} = 0.001$; Accuracy: $z = 1.99$, $P_{corr} = 0.046$) (see Supplementary Table 6).

Finally, to answer the question which of these neural parameters is the stronger predictor for language performance we related each neural factor to each of the behavioral measures while controlling for the influence of the respective other factor. The FA of the AF significantly predicted syntax processing speed not only in the basic model adjusting for age and memory as described above but also in the advanced model additionally adjusting for

the BOLD signal change in the IFG opercularis and posterior STG ($r = -0.22$, $P_{corr} = 0.045$). Furthermore, the correlation between the FA of the AF and syntax processing speed was significantly stronger than the correlation between BOLD signal changes in the interconnected ROIs and syntax processing speed ($z = 1.85$, $P_{corr} = 0.045$). BOLD signal changes in the IFG opercularis and posterior STG significantly predicted syntax processing accuracy for OR in the basic model adjusting for age and memory as seen above and also in the advanced model additionally adjusting for structure ($r = 0.32$, $P_{corr} = 0.004$). Crucially, however, the

Table 3 Correlations between BOLD signal change and sentence comprehension performance

	Dorsally connected ROIs (BOLD signal change %) and accuracy (%)	Dorsally connected ROIs (BOLD signal change %) and RT (ms)	Ventrally connected ROIs (BOLD signal change %) and accuracy (%)	Ventrally connected ROIs (BOLD signal change %) and RT (ms)
SR	$r = 0.13$, $P = 0.185$	$r = -0.15$, $P_{corr} = 0.133$	$r = 0.19$, $P_{corr} = 0.063$	$r = -0.1$, $P = 0.288$
OR	$r = 0.28$, $P_{corr} = 0.005^*$	$r = -0.06$, $P = 0.534$	$r = 0.07$, $P = 0.48$	$r = -0.06$, $P = 0.536$
OR versus SR	$z = 2.06$, $P_{corr} = 0.039^*$	$z = 1.22$	$z = 1.36$	$z = 0.54$

Note: The activation of the IFG opercularis and the posterior STG (dorsally connected ROIs) was significantly correlated with the processing accuracy of the syntactically complex OR. This correlation was significantly stronger than the correlation of the OR with the activation of the ventrally connected ROIs (IFG triangularis and anterior STG) ($z = 1.97$, $P_{corr} = 0.049$). The BOLD signal change of the IFG opercularis and the posterior STG was still significantly related to the processing accuracy of the syntactically complex OR when adjusting for the BOLD signal change of the IFG triangularis and anterior STG ($r = 0.28$, $P_{corr} = 0.005$) but not vice versa ($r = -0.16$, $P_{corr} = 0.11$) with the difference between the correlation coefficients reaching significance ($z = 4.03$, $P_{corr} < 0.001$). The correlation between the dorsally connected ROIs and the accuracy on syntactically complex OR was significantly stronger than the correlation with the accuracy on syntactically simpler SR within the same ROI ($z = 2.06$, $P_{corr} = 0.039$). All P values are Bonferroni corrected. Age and verbal working memory scores were entered as covariates of no interest in all models. Asterisks indicate significant effects.

Table 4 Partial correlations between BOLD signal change, FA, and sentence comprehension performance

	STRUCTURE and accuracy (%) adjusted for FUNCTION	STRUCTURE and RT (ms) adjusted for FUNCTION	FUNCTION and accuracy (%) adjusted for STRUCTURE	FUNCTION and RT (ms) adjusted for STRUCTURE
SR	$r = 0.2$, $P_{corr} = 0.067$	$r = -0.19$, $P_{corr} = 0.079$	$r = 0.21$, $P_{corr} = 0.056$	$r = -0.04$, $P = 0.751$
OR	$r = 0.21$, $P_{corr} = 0.054$	$r = -0.22$, $P_{corr} = 0.045^*$	$r = 0.32$, $P_{corr} = 0.004^*$	$r = 0.04$, $P = 0.75$
OR versus SR	$z = 0.12$	$z = 0.29$	$z = 1.13$	$z = 0.7$

Note: FA significantly predicts syntax processing speed for the OR when eliminating the effect of BOLD signal change. Furthermore, FA is a better predictor of syntax processing speed than BOLD signal change as the FA values of the SLF/AF correlate stronger with the RTs of the OR than the BOLD signal change in the IFG opercularis and posterior STG ($z = 1.85$, $P_{corr} = 0.045$). BOLD signal change significantly predicts syntax processing accuracy for the OR and FA does not but this difference is not significant ($z = 0.84$). The difference between the correlation coefficients for accuracy on OR and SR related to BOLD signal change did not reach significance when adjusting for the indirect effect of FA. All P values are Bonferroni corrected. Age and verbal working memory scores were entered as covariates of no interest in all models. Asterisks indicate significant effects.

correlation between BOLD signal change and syntax processing accuracy was not significantly stronger than the correlation between FA and syntax processing accuracy ($z = 0.84$). Accordingly, FA was related to speed and BOLD signal change did not explain considerable additional variance with respect to this relation. BOLD signal change, however, was related to accuracy but FA did explain considerable additional variance with respect to this relation. Crucially, the associations between BOLD signal change and the accuracy on sentences of each type no longer differed significantly from each other when removing the effect of FA ($z = 1.13$). Thus, BOLD signal change within the left IFG opercularis and the left posterior STG and the FA of the interconnecting AF together explain the specific effect of syntax performance accuracy (Table 4).

Discussion

The present data allow for the first time a precise specification of the tripartite relation between fronto-temporal hemodynamic cortical activity, FA of the interconnecting white matter fiber bundles and sentence comprehension performance from early language development into adulthood. This goes beyond our previous study reporting whole-brain functional responses underlying sentence comprehension of 3 age groups of children (Skeide et al. 2014). The results demonstrate a dependency of full grammar skills not only on the increased functional selectivity of Broca's area and the posterior STG, but moreover, on the maturational status of the interconnecting dorsal pathway (AF). Thus, full performance on syntactically complex sentences can only be achieved once this fiber tract is matured so that the dorsal language network can provide its full functionality. These findings show that certain cognitive abilities, here milestones

in language development, are only achieved once the brain maturational prerequisites are in place.

In accordance with earlier functional MRI studies on the language network in adults and infants (Dehaene-Lambertz et al. 2006), we observed that the language-sensitive areas along the Sylvian fissure responded in a similar way across development when listening to sentences irrespective of their complexity. Therefore, we consider it a valid approach to use independent ROIs from the adult literature to assess syntax-specific changes in functional selectivity within a developmentally stable pattern of language comprehension related neural activity. This is further supported by structural postmortem studies suggesting that the cytoarchitectonic specialization of the late language-related cortices is almost finished after 3 years of life (Judaš and Cepanec 2007; Petanjek et al. 2011), and by the general physiological comparability of the BOLD signal amplitude between young children and adults in the frontal and temporal lobe (Taki et al. 2011; Moses et al. 2013).

The functional cortical selectivity for complex syntax emerges, however, emerges slowly with age. It is only established at 9–10 years of age, as reflected in an adult-like significantly stronger BOLD signal change within the posterior STG in response to OR compared with SR, while the behavioral asymmetry between OR and SR, as reflected in a significantly lower mean response accuracy for OR compared with SR, disappears. The finding that the left IFG opercularis crucially underlies the cortical sensitivity for complex syntax in adulthood is in line with numerous studies indicating that this region forms a core computational unit of sentence processing across varying experimental designs and task demands (Caplan et al. 2008) that can be spatially disentangled from surrounding domain-general areas serving working memory, cognitive control and music processing (Makuuchi et al. 2009; Fedorenko et al. 2011, 2012).

We identified the dorsal pathway to be crucially involved in syntactic processing. In contrast to our findings, several previous studies suggested an involvement of both a dorsal and a ventral pathway in sentence comprehension (Rolheiser et al. 2011; Papoutsi et al. 2011; Griffiths et al. 2012). However, these studies either used syntax manipulations that were not controlled for possibly interfering semantic processes or reported effects that were considerably more pronounced in the dorsal tract compared with the ventral tract. These limitations were overcome in a semantic-free artificial grammar learning study (Friederici et al. 2006) and in a study on primary progressive aphasia patients with a confined damage either to the AF or the IFOF (Wilson et al. 2011) that provide converging evidence for a specific link between complex syntax and the dorsal pathway. Our observations provide strong support for the crucial role of the dorsal pathway for syntax performance. Furthermore, we extend the present knowledge about the neural implementation of complex syntax by providing direct statistical evidence that while functional activation in the 2 relevant language regions are important for performance accuracy of syntactically complex sentences, the FA of the white matter connection between these regions is not only the better predictor for the performance speed but also explains considerable additional variance in performance accuracy of syntactically complex sentences. This implies that the role of the maturing AF goes beyond increasing transmission efficiency. Finally, it is also shown that syntactic computation during sentence comprehension can be segregated from verbal working memory contributions, which is in line with a previous study in adults (Meyer et al. 2012).

Right hemisphere homologs of the left-hemispheric syntax-sensitive regions were not analyzed in the present study as they were, to the best of our knowledge, never specifically associated with the processing of complex syntactic information, particularly in right-handers. It is rather known that a bilateral organization of the language network supports the acquisition of phonological representations during late prenatal and early postnatal development (Dubois et al. 2010; Perani et al. 2011; Leroy 2011; Mahmoudzadeh et al. 2013). Moreover, Catani et al. (2007) have convincingly shown that a more pronounced bilateral connectivity can be advantageous for learning new words by semantic associations in adulthood. The right-hemispheric perisylvian areas and their interconnecting fiber tracts were also very consistently linked to the processing of suprasegmental (prosodic) features of speech (Friederici 2011, 2012). Hence there is currently no evidence for the view that a more bilateral organization of the language network might have an impact on the processing of syntactic information. Crucially, this view is supported by a previous fMRI study, in which the whole-brain data of the 3 groups of children of the current study were analyzed (Skeide et al. 2014) indicating that right-hemispheric BOLD activity seen in the baseline contrast (sentences vs. silence) entirely disappears when contrasting syntactically more complex versus simpler sentences. Additional whole-brain analyses involving the adult sample corroborate the position that the right hemisphere does not directly contribute to the processing of complex syntactic information (see Supplementary Figs 2–4).

Note, that we do not doubt that the ventral pathway is involved in syntax processing. Nevertheless, it is assumed to be primarily involved in detecting low-level syntactic features such as phrase types (Friederici 2012). Additionally, the present data do not speak against the possibility that the dorsal and ventral tracts might also underlie or even share other language-specific and language-unspecific cognitive functions. The noninvasive reconstruction of the AF, its subcomponents and their termination areas are still under intensive investigation and discussion (Catani et al. 2002; Rilling et al. 2008; Fernández-Miranda et al. 2014; Brown et al. 2014). The focus of the present study was on the subcomponent of the AF that directly connects the posterior inferior frontal and the posterior STG as the 2 regions involved in syntactic processing.

The FA changes reported in the current study most likely reflect several aspects of white matter maturity including myelination (Mukherjee et al. 2001), axon growth (Paus 2010), and fiber density (Scholz et al. 2009). With the help of the currently available methods, it is not possible to disentangle the specific contribution of each factor to the FA. The methodological challenge for future studies is to develop noninvasive tissue-sensitive indices of white matter development. Importantly, FA is also influenced by the directional coherence of fibers (Westin et al. 2002). Crossing fibers can lead to partial volume effects substantially confounding FA estimation. The tract reconstruction approach used in the present study minimizes the possible influence of partial volume effects induced by crossing fibers since the individual connectivity maps are thresholded and averaged before using them as masks for second-level statistics on the FA maps. The resulting tractograms robustly represent the centers of all individual tractograms sparing their outermost boarders where fiber crossings most frequently occur in the AF and the IFOF.

The AF is considered highly relevant to the evolution of language due to its unique specialization in humans compared to nonhuman primates (Rilling et al. 2008). Our findings complement this picture suggesting that the left AF decisively contributes not only to the phylogenetic but also to the ontogenetic emergence of sophisticated language skills. Support for this view comes from the tight functional relation between the interconnected frontal and temporal language regions that is corroborated by their common genetic signatures (Johnson et al. 2009) as well as their common receptorarchitectonic fingerprints in the left but not in the right hemisphere (Zilles et al. 2014).

To conclude, this is the first study determining the direct relation between hemodynamic functional activity, white matter FA and behavioral accuracy and speed in a higher cognitive domain across multiple developmental age groups. It is demonstrated that although the basic cortical components of the functional language network are already established at 3 years of age, the domain-specific selectivity for complex grammar within this network emerges late and so does the white matter fiber tract forming the dorsal connection of this network. The present data suggest that not fronto-temporal BOLD activity alone but BOLD signal changes within the frontal and temporal cortex and the FA of the interconnecting AF together specifically predict the ontogenetic emergence of complex grammar skills.

Supplementary Material

Supplementary Material can be found at http://www.cercor.oxfordjournals.org/online.

Funding

Notes

References

Amunts K, Schleicher A, Ditterich A, Zilles K. 2003. Broca's region: cytoarchitectonic asymmetry and developmental changes. J Comp Neurol. 465:72–89.

Anwander A, Tittgemeyer M, von Cramon DY, Friederici AD, Knösche TR. 2007. Connectivity-based parcellation of Broca's area. Cereb Cortex. 17:816–825.

Bonner MF, Price AR. 2013. Where is the anterior temporal lobe and what does it do? J Neurosci. 33:4213–4215.

Brauer J, Anwander A, Friederici AD. 2011. Neuroanatomical prerequisites for language functions in the maturing brain. Cereb Cortex. 21:459–466.

Brauer J, Anwander A, Perani D, Friederici AD. 2013. Dorsal and ventral pathways in language development. Brain Lang. 127:289–295.

Brown EC, Jeong JW, Muzik O, Rothermel R, Matsuzaki N, Juhász C, Sood S, Asano E. 2014. Evaluating the arcuate fasciculus with combined diffusion-weighted MRI tractography and electrocorticography. Hum Brain Mapp. 35:2333–2347.

Caplan D, Alpert N, Waters G, Olivieri A. 2000. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. Hum Brain Mapp. 9:65–71.

Caplan D, Chen E, Waters G. 2008. Task-dependent and task-independent neurovascular responses to syntactic processing. Cortex. 44:257–275.

Caplan D, Vijayan S, Kuperberg G, West C, Waters G, Greve D, Dale AM. 2002. Vascular responses to syntactic processing: event-related fMRI study of relative clauses. Hum Brain Mapp. 15:26–38.

Catani M, Allin MP, Husain M, Pugliese L, Mesulam MM, Murray RM, Jones DK. 2007. Symmetries in human brain language pathways correlate with verbal recall. Proc Natl Acad Sci USA. 104:17163–17168.

Catani M, Howard RJ, Pajevic S, Jones DK. 2002. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. Neuroimage. 17:77–94.

Chomsky C. 1969. The acquisition of syntax in children from 5 to 10. Cambridge: MIT Press.

Chomsky N. 1995. Language and nature. Mind. 104:1–61.

Clark EV. 2003. First language acquisition. Cambridge: Cambridge University Press.

Dehaene-Lambertz G, Hertz-Pannier L, Dubois J, Mériaux S, Roche A, Sigman M, Dehaene S. 2006. Functional organization of perisylvian activation during presentation of sentences in preverbal infants. Proc Natl Acad Sci USA. 103:14240–14245.

Dittmar M, Abbot-Smith K, Lieven E, Tomasello M. 2008. German children's comprehension of word order and case marking in causative sentences. Child Dev. 79:1152–1167.

Dubois J, Benders M, Lazeyras F, Borradori-Tolsa C, Leuchter RH, Mangin JF, Hüppi PS. 2010. Structural asymmetries of perisylvian regions in the preterm newborn. Neuroimage. 52:32–42.

Dunn OJ, Clark VA. 1969. Correlation coefficients measured on the same individuals. J Am Stat Assoc. 64:366–377.

Fanselow G. 2001. Features, theta-roles, and free constituent order. Ling Inquiry. 32:405–437.

Fedorenko E, Behr MK, Kanwisher N. 2011. Functional specificity for high-level linguistic processing in the human brain. Proc Natl Acad Sci USA. 108:16428–16433.

Fedorenko E, Duncan J, Kanwisher N. 2012. Language-selective and domain-general regions lie side by side within Broca's area. Curr Biol. 22:2059–2062.

Fernández-Miranda JC, Wang Y, Pathak S, Stefaneau L, Verstynen T, Yeh FC. 2014. Asymmetry, connectivity, and segmentation of the arcuate fascicle in the human brain. Brain Struct Funct.

Fields RD, Neuroscience. 2010. Change in the brain's white matter. Science. 330:768–769.

Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL, Brain Development Cooperative Group. 2011. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage. 54:313–327.

Fox AV. 2011. TROG-D. Test zur Überprüfung des Grammatikverständnisses. Idstein: Schulz-Kirchner.

Fox D, Grodzinsky Y. 1998. Children's passive: a view from the by-phrase. Linguist Inq. 29:311–332.

Frey S, Campbell JS, Pike GB, Petrides M. 2008. Dissociating the human language pathways with high angular resolution diffusion fiber tractography. J Neurosci. 28:11435–11444.

Friederici AD. 1983. Children's sensitivity to function words during sentence comprehension. Linguistics. 21:717–739.

Friederici AD. 2011. The brain basis of language processing: from structure to function. Physiol Rev. 91:1357–1392.

Friederici AD. 2012. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn Sci. 16:262–268.

Friederici AD, Bahlmann J, Heim S, Schubotz RI, Anwander A. 2006. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. Proc Natl Acad Sci USA. 103:2458–2463.

Friedmann N, Belletti A, Rizzi L. 2009. Relativized relatives. Types of Intervention in the Acquisition of A-bar Dependencies Lingua. 119:67–88.

Griffiths JD, Marslen-Wilson WD, Stamatakis EA, Tyler LK. 2012. Functional organization of the neural language system: dorsal and ventral pathways are critical for syntax. Cereb Cortex. 23:139–147.

Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. 2002. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Res Med. 47:1202–1210.

Guasti MT. 2002. Language acquisition: The growth of grammar. Cambridge: MIT Press.

Hahne A, Eckstein K, Friederici AD. 2004. Brain signatures of syntactic and semantic processes during children's language development. J Cogn Neurosci. 16:1302–1318.

Hirsh-Pasek K, Golinkoff RM. 1996. The origins of grammar: evidence from early language comprehension. Cambridge: MIT Press.

Hoeft F, Meyler A, Hernandez A, Juel C, Taylor-Hill H, Martindale JL, McMillon G, Kolchugina G, Black JM, Faizi A, et al. 2007. Functional and morphometric brain dissociation between dyslexia and reading ability. Proc Natl Acad Sci USA. 104:4234–4239.

Höhle B, Schmitz M, Santelmann LM, Weissenborn J. 2006. The recognition of discontinuous verbal dependencies by German. 19-Month-olds: evidence for lexical and structural influences on children's early processing capacities. Lang Learn Dev. 2:277–300.

Humphries C, Binder JR, Medler DA, Liebenthal E. 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. J Cogn Neurosci. 18:665–679.

Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage. 17: 825–841.

Johnson MB, Kawasawa YI, Mason CE, Krsnik Z, Coppola G, Bogdanović D, Geschwind DH, Mane SM, State MW,

Sestan N. 2009. Functional and evolutionary insights into human brain development through global transcriptome analysis. Neuron. 62:494–509.

Judaš M, Cepanec M. 2007. Adult structure and development of the human fronto-opercular cerebral cortex (Broca's region). Clin Linguist Phon. 21:975–989.

Leroy F, Glasel H, Dubois J, Hertz-Pannier L, Thirion B, Mangin J, Dehaene-Lambertz G. 2011. Early maturation of the linguistic dorsal pathway in human infants. J Neurosci. 31:1500–1506.

Li G, Nie J, Wang L, Shi F, Gilmore JH, Lin W, Shen D. 2014. Measuring the dynamic longitudinal cortex development in infants by reconstruction of temporally consistent cortical surfaces. NeuroImage. 90:266–279.

Mahmoudzadeh M, Dehaene-Lambertz G, Fournier M, Kongolo G, Goudjil S, Dubois J, Grebe R, Wallois F. 2013. Syllabic discrimination in premature human infants prior to complete formation of cortical layers. Proc Natl Acad Sci USA. 110:4846–4851.

Makuuchi M, Bahlmann J, Anwander A, Friederici AD. 2009. Segregating the core computational faculty of human language from working memory. Proc Natl Acad Sci USA. 106:8362–8367.

Markman EM, Wasow JL, Hansen MB. 2003. Use of the mutual exclusivity assumption by young word learners. Cogn Psychol. 47:241–275.

Melchers P, Preuss U. 2009. Kaufman assessment battery for children. Frankfurt: Pearson.

Meng XL, Rosenthal R, Rubin DB. 1992. Comparing correlated correlation coefficients. Psychol Bull. 111:172–175.

Mesulam MM, Rogalski EJ, Wieneke C, Hurley RS, Geula C, Bigio EH, Thompson CK, Weintraub S. 2014. Primary progressive aphasia and the evolving neurology of the language network. Nat Rev Neurol. 10:554–569.

Meyer L, Obleser J, Anwander A, Friederici AD. 2012. Linking ordering in Broca's area to storage in left temporo-parietal regions: the case of sentence processing. Neuroimage. 62:1987–1998.

Moses P, DiNino M, Hernandez L, Liu TT. 2013. Developmental changes in resting and functional cerebral blood flow and their relationship to the BOLD response. Hum Brain Mapp. 35:3188–3198.

Mukherjee P, Miller JH, Shimony JS, Conturo TE, Lee BC, Almli CR, McKinstry RC. 2001. Normal brain maturation during childhood: developmental trends characterized with diffusion-tensor MR imaging. Radiology. 221:349–358.

Nazzi T, Bertoncini J. 2003. Before and after the vocabulary spurt: two modes of word acquisition? Dev Sci. 6:136–142.

Newman SD, Ikuta T, Burns T Jr. 2010. The effect of semantic relatedness on syntactic analysis: an fMRI study. Brain Lang. 113:51–58.

Nuñez SC, Dapretto M, Katzir T, Starr A, Bramen J, Kan E, Bookheimer S, Sowell ER. 2011. fMRI of syntactic processing in typically developing children: structural correlates in the inferior frontal gyrus. Dev Cogn Neurosci. 1:313–323.

Oberecker R, Friederici AD. 2006. Syntactic ERP components in 24-month-olds' sentence comprehension. Neuroreport. 17:1017–1021.

Oberecker R, Friedrich M, Friederici AD. 2005. Neural correlates of syntactic processing in two-year-olds. J Cogn Neurosci. 17:1667–1678.

Oldfield RC. 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 9:97–113.

O'Shaughnessy ES, Berl MM, Moore EN, Gaillard WD. 2008. Pediatric functional magnetic resonance imaging (fMRI): issues and applications. J Child Neurol. 23:791–801.

Pallier C, Devauchelle AD, Dehaene S. 2011. Cortical representation of the constituent structure of sentences. Proc Natl Acad Sci USA. 108:2522–2527.

Papoutsi M, Stamatakis EA, Griffiths J, Marslen-Wilson WD, Tyler LK. 2011. Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. Neuroimage. 58:656–664.

Paus T. 2010. Growth of white matter in the adolescent brain: myelin or axon? Brain Cogn. 72:26–35.

Perani D, Saccuman MC, Scifo P, Anwander A, Spada D, Baldoli C, Poloniato A, Lohmann G, Friederici AD. 2011. Neural language networks at birth. Proc Natl Acad Sci USA. 108:16056–16061.

Petanjek Z, Judaš M, Šimic G, Rasin MR, Uylings HB, Rakic P, Kostovic I. 2011. Extraordinary neoteny of synaptic spines in the human prefrontal cortex. Proc Natl Acad Sci USA. 108:13281–13286.

Reese TG, Heid O, Weisskoff RM, Wedeen VJ. 2003. Reduction of Eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magn Res Med. 49:177–182.

Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, Behrens TE. 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. Nat Neurosci. 11:426–428.

Rolheiser T, Stamatakis EA, Tyler LK. 2011. Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule. J Neurosci. 31:16949–16957.

Sakai KL. 2005. Language acquisition and brain development. Science. 310:815–819.

Schipke CS, Friederici AD, Oberecker R. 2011. Brain responses to case-marking violations in German preschool children. Neuroreport. 22:850–854.

Schipke CS, Knoll LJ, Friederici AD, Oberecker R. 2012. Preschool children's interpretation of object-initial sentences: Neural correlates of their behavioral performance. Dev Sci. 15:762–774.

Scholz J, Klein MC, Behrens TEJ, Johansen-Berg H. 2009. Training induces changes in white matter architecture. Nat Neurosci. 12:1370–1371.

Silva-Pereyra J, Rivera-Gaxiola M, Kuhl PK. 2005. An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing. Cogn Brain Res. 23:247–258.

Skeide MA, Brauer J, Friederici AD. 2014. Syntax gradually segregates from semantics in the developing brain. Neuroimage. 100:106–111.

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, et al. 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage. 31:1487–1505.

Sowell ER, Thompson PM, Leonard CM, Welcome SE, Kan E, Toga AW. 2004. Longitudinal mapping of cortical thickness and brain growth in normal children. J Neurosci. 24:8223–8231.

Steinbeis N, Bernhardt BC, Singer T. 2012. Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. Neuron. 73:1040–1051.

Taki Y, Hashizume H, Sassa Y, Takeuchi H, Wu K, Asano M, Asano K, Fukuda H, Kawashima R. 2011. Correlation between grey matter density-adjusted brain perfusion and age using brain MR images of 202 healthy children. Hum Brain Mapp. 32:1973–1985.

Talairach J, Tournoux P. 1988. Co-planar stereotactic atlas of the human brain. New York: Thieme.

Thirion JP. 1998. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal. 2:243–260.

van der Lely HK, Pinker S. 2014. The biological basis of language: insight from developmental grammatical impairments. Trends Cogn Sci. 18:586–595.

Wake H, Lee PR, Fields D. 2011. Control of local protein synthesis and initial events in myelination by action potentials. Science. 33:1647–1651.

Westin CF, Maier SE, Mamata H, Nabavi A, Jolesz FA, Kikinis R. 2002. Processing and visualization for diffusion tensor MRI. Med Image Anal. 6:93–108.

Wilson SM, Galantucci S, Tartaglia MC, Rising K, Patterson DK, Henry ML, Ogar JM, DeLeon J, Miller BL, Gorno-Tempini ML.

2011. Syntactic processing depends on dorsal language tracts. Neuron. 72:397–403.

Yeatman JD, Dougherty RF, Ben-Shachar M, Wandell BA. 2012. Development of white matter and reading skills. Proc Natl Acad Sci U S A. 109:E3045–E3053.

Yoon U, Fonov VS, Perusse D, Evans AC, Brain Development Cooperative Group. 2009. The effect of template choice on morphometric analysis of pediatric brain data. NeuroImage. 45:769–777.

Zilles K, Bacha-Trams M, Palomero-Gallagher N, Amunts K, Friederici AD. 2014. Common molecular basis of language processing revealed by neurotransmitter receptor fingerprints. Cortex. 63:79–89.